

推薦論文

2文字連続筆記した平仮名空中手書き文字の分割法

鈴木 慶^{1,a)} 清水 忠昭¹ 吉村 宏紀¹ 松村 寿枝²

受付日 2013年10月29日, 採録日 2014年6月17日

概要: 本論文では空中手書き文字 (AHC) 入力システムのための自動的な文字分割手法を提案する. AHC 入力システムは, 人間が空中に文字を書く動作をコンピュータに認識させて文字入力を行うシステムである. 従来の AHC 入力システムでは, 文字と文字を自動で切り離すことができなかったため, 人間の意図的な操作により文字と文字を区切る必要があった. 提案手法では, AHC のストロークを移動ストロークと文字ストロークに判別することで, 文字を自動で分割する. 評価実験の結果, closed テストでは移動ストロークと文字ストローク双方で 93%以上の正解率となった. open テストでは文字ストロークは 88.8%, 移動ストロークは 73.3%の正解率を得た.

キーワード: 空中手書き文字, 身振りインターフェイス

A Character Segmentation Method for Aerial Handwritten Hiragana Characters

KEI SUZUKI^{1,a)} TADAAKI SHIMIZU¹ HIROKI YOSHIMURA¹ TOSHIE MATSUMURA²

Received: October 29, 2013, Accepted: June 17, 2014

Abstract: The aerial handwritten character (AHC) input system is employed in user interfaces for the recognition of a user's aerial handwriting gestures. A conventional AHC input system cannot separate sequential characters into individual characters automatically; it requires the user to perform this task. Therefore, we propose a character segmentation method for the AHC system. Our method separates an AHC trajectory into individual characters by distinguishing transition strokes from character strokes. Experiments conducted to evaluate our method showed that the detection accuracy of transition and character strokes was greater than 93% in the closed test, whereas it was 73.3% and 88.8%, respectively, in the open test.

Keywords: aerial handwritten character, gesture interface

1. はじめに

人間同士のコミュニケーションに“声”や“身振り手振り”などの動作を用いることは自然であり, 我々は意思疎通の過程を楽しむことさえできる. 一方, 高度化した家電や情報機器にユーザの意図を伝えることは難しく, その操作は多くの場合苛立つ経験である. この問題を解決するために, 人間にとって「自然な動作」によって機械と接する

ことができるインタフェース (ナチュラル・ユーザ・インタフェース: NUI) が研究・開発されている.

NUIの手法を取り入れた手軽な文字入力技術として, “手書き文字認識”が研究されている. 隠れマルコフモデルを用いて手書き文字を認識する研究 [1], [2] やオンライン手書き文字認識の認識精度を上げる研究 [3] など多数の研究事例があり, スマートフォンやペンタブレットの普及により手書き文字認識技術への需要はますます高まっている.

人間同士のコミュニケーションにおいて, 文字は通常は紙上などに書かれるが, 手元に紙や筆記具がない場合, 指

¹ 鳥取大学大学院工学研究科
Graduate School of Engineering, Tottori University, Tottori
680-8552, Japan

² 奈良工業高等専門学校
Nara National College of Technology, Yamatokoriyama,
Nara 639-1080, Japan

^{a)} suzuki.kei16@gmail.com

本論文の内容は 2012 年 10 月の平成 24 年度 (第 63 回) 電気・情報関連学会中国支部連合大会にて報告され, 支部長により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である.

で空中をなぞって文字を伝えることがある。タッチパネルなどのデバイス上で前者を模したインタフェース技術が一般的な手書き文字認識であるのに対し、後者を模したインタフェース技術を“空中手書き文字認識”と呼ぶ。

空中手書き文字認識の技術は、文字を書く動作を検出するセンサを身に付けるセンサ装着型の手法と、据え置いたカメラなどのセンサに向かって筆記動作をするセンサ非装着型の手法に分類できる。

センサ装着型の手法では、持ち運びが容易な軽量でコンパクトサイズのペン状デバイスが用いられることが多い。ペン先についた加速度センサを用いて、筆記動作の加速度方向を検出し文字を認識する方法 [4], [5], ペン先についたカメラを用いて、筆記動作時の映像の移動方向情報から文字を認識する方法 [6], ヘッドマウントディスプレイ, IR センサ, Web カメラなどを装着し、目の筆記動作パターンで文字認識する方法 [7] などが提案されている。センサ装着型の手法には、「いつでもどこでも」利用できるユビキタスを可能にする利点があるが、一方でセンサデバイスを身につける必要があり、利用者にとって負担となる。

センサ非装着型の手法として、画像処理を軽減するために手に電球を持ちビデオカメラに向かって文字を書く方法 [8] や、複数の web カメラを利用者の周りに設置し、3D 処理をして文字を書く方法 [9] などが提案されている。これらセンサ非装着型には、ハンズフリーによる自然な文字入力の可能性がある。

空中手書き文字認識によるハンズフリーで自然な文字入力のためには、文字分割の問題を解決しなければならない。我々が紙に文字を書く際には、無意識に筆を上げる動作を行い、文字と文字を切り離している。しかし、空中に文字を書く場合には、文字は一繋がり軌跡として連続して書かれる。先行研究 [8] では、文字の筆記を検出するセンサは非装着であるが、“スイッチ”となるデバイスを身に付け、ユーザの意図的なスイッチ操作によって文字分割を行っている。この方法では文字分割が確実である反面、デバイスを身に付ける必要があるうえに、その操作は「自然な動作」とはいえない。一方、先行研究 [9] では複数の web カメラを用いた 3D 処理によって、ユーザがカメラに対して手を引く動作を検出し文字分割を行う。この方法では、ユーザの周囲に複数のカメラを設置する必要があり、大規模なシステムとなってしまうだけでなく、空中手書き文字の動作としては不自然な“筆上げ”の動作を必要とする欠点がある。

この問題に対し、続けて書かれた数字についてユーザによる文字分割操作なしで文字認識を行う手法 [10] が報告されている。この手法は、続けて書かれた数字を分割することなくテンプレートマッチングする方法である。この手法は数字に関して高い認識率を示しているが、平仮名やアルファベットのように認識対象となる文字種の増加への対処に懸念がある。

本研究では、空中手書き文字認識によるハンズフリーで自然な平仮名文字入力を簡易なシステムで実現することを目標として、文字分割の問題の解決手法を提案する。提案手法では、自然に書かれた空中手書き文字から文字軌跡の特徴を抽出し、文字どうしをつなぐ“移動ストローク”を自動的に発見することにより文字分割を行う。評価実験の結果、本手法によるストロークの判別は、closed テストでは、移動ストロークと文字ストローク双方で 93%以上の正解率となった。open テストでは、文字ストロークは 88.8%, 移動ストロークは 73.3%の正解率を得た。

2. 研究目標と空中手書き文字の定義

本研究の目標は、空中手書き文字 (Aerial Handwritten Characters: AHC) によるハンズフリーで自然な文字入力システムの実現である。図 1 は提案する平仮名空中手書き文字の文字分割法を検証するために試作した AHC 入力システムのカメラがとらえた筆記者の様子であり、指先で AHC を筆記した軌跡が重畳表示されている。

図 2 に平仮名「あい」を書いた際の AHC の例を示す。AHC の筆記は紙上の筆記と異なり文字の描線を切る筆上げの動作がないため、文字は一筆書きの軌跡として書かれる。この軌跡を“文字軌跡”と呼ぶ。図 2 中の ①～⑨ に示すような文字の画と、それらをつなぐ画を“ストローク”と呼ぶ。また、①～⑤, ⑦～⑨ のように文字を構成するストロークを“文字ストローク”, ⑥ のように文字の間をつなぐストロークを“移動ストローク”と呼ぶ。

本研究では、文字認識のために AHC を 1 文字ごとに分割する手法を提案する。文字分割のためには文字軌跡をストロークに分割し、移動ストロークを取り除けばよい。したがって、AHC における文字分割の問題は AHC を構成するストローク中から移動ストロークを発見する問題に帰着できる。

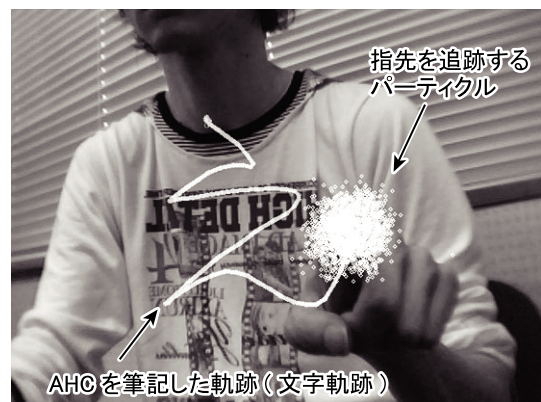


図 1 提案システムの表示画面 ~ USB カメラがとらえた筆記者の動画に AHC の軌跡が重畳表示される

Fig. 1 Display screen of the proposed system – An AHC trajectory is superimposed on a video image of the user captured using a USB camera.

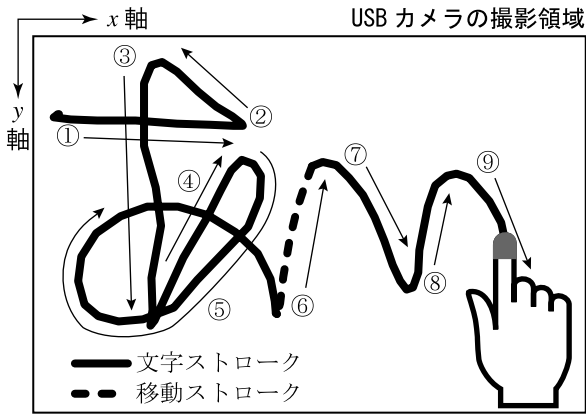


図 2 空中手書き文字の例「あい」

Fig. 2 Example of AHC operation (“ai” Japanese characters).

3. 提案手法

図 3 に提案手法による文字軌跡データの取得から文字分割までの処理の流れを示す。流れ図中の処理箱の左側に付した番号は、処理箱の中に記載された処理を詳述した節の番号を示している。本研究では、手軽な AHC 入力システムを目指しているため、表 1 に示す一般的な PC と USB カメラの組合せでシステムを構成した。

3.1 指先の追跡—文字軌跡の取得

AHC の文字軌跡は、AHC の筆記動作を USB カメラで撮影した動画において筆記者の指先を実時間で追跡することによって取得する。本研究の主題は AHC の文字分割であるため、指先追跡は、筆記者の指先に図 4 に示す赤い指サックを装着し、パーティクルフィルタ [11] を用い赤い物体を追跡する簡易な方法で実現した。パーティクルフィルタは OpenCV [12] を利用して実装した。USB カメラの解像度は、640 (pixels) × 480 (pixels) に設定し、フレームレートは $F = 33.33$ (fps) とした。

USB カメラから取得した動画の第 n フレームについて、指先を追跡するパーティクルの重心点を指先位置のサンプル点 $\mathbf{p}_n = (x_n, y_n)$ とした。得られたサンプル点の並び $\mathbf{P} = \{\mathbf{p}_n\}$, $1 \leq n \leq N$ が文字軌跡であり、 N は文字軌跡の長さである。

3.2 文字をストロークに分割する

文字軌跡 $\mathbf{P}$ をストロークに分割するために、以下の式に示す AHC 筆記の速度 v_n と加速度 α_n を用いた。

$$v_n(\text{pixel/s}) = \frac{\sqrt{(x_n - x_{n-1})^2 + (y_n - y_{n-1})^2}}{T} \quad (1)$$

$$\alpha_n(\text{pixel/s}^2) = \frac{v_n - v_{n-1}}{T} \quad (2)$$

ここで、 T は動画のフレーム周期であり、 $T(\text{s}) = 1/F$ である。また、AHC 筆記の速度と加速度は、次式の 5 点移動平均フィルタを適用して平滑化した。

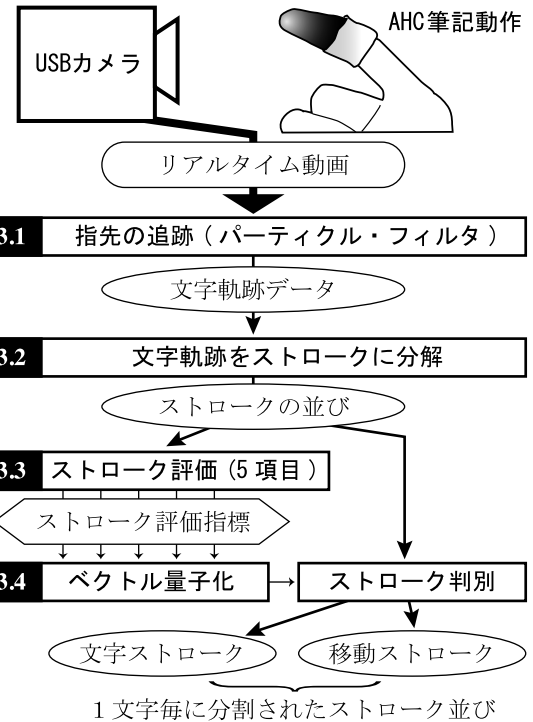


図 3 提案する AHC 分割法の概要

Fig. 3 Block diagram of the proposed AHC segmentation method.

表 1 PC と USB カメラの仕様

Table 1 Specifications of the PC and the USB camera.

PC		USBカメラ	
CPU	Intel®Core™2 Duo E7500@ 2.94GHz	品名	BUFFALO BSW13K06H
		画素数	130 万画素
メモリ	4GB	解像度	最大 1280×1024
OS	Ubuntu 10.04.1 LTS	FPS	最大約 30fps



図 4 指先追跡のための赤い指サック

Fig. 4 Fingerstall for fingertip tracking.

$$\bar{f}_n = \frac{f_{n+2} + 2f_{n+1} + 3f_n + 2f_{n-1} + f_{n-2}}{9} \quad (3)$$

図 5 に、実測した筆記加速度に基づいて、加速している部分は濃い色で、減速している部分は薄い色で文字軌跡を描画した例を示す。この例でも分かれるとおり、AHC の筆記の際、ストロークの描き始めに加速し、中間部では一定速となり、ストロークの描き終わりに減速する傾向がある。この性質を用いて、筆記動作が減速後に加速した点で文字軌跡 $\mathbf{P}$ をストローク $\mathbf{S}_m$, $1 \leq m \leq M$ に分割する。 M は分割で得られたストローク数である。

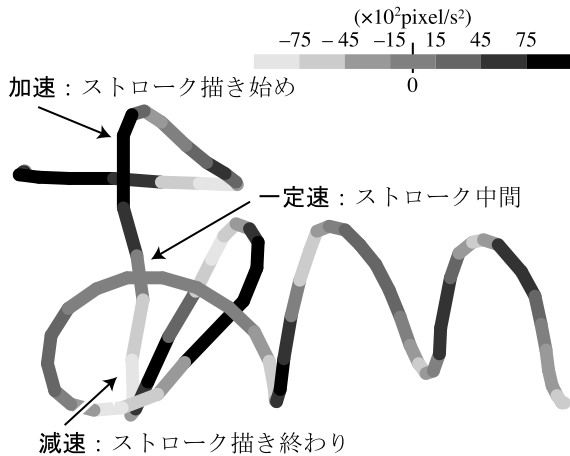


図 5 文字軌跡上における指先の加速度

Fig. 5 Acceleration of the fingertip on the character trajectory.

3.3 ストロークの評価指標

本研究では、提案手法に以下の2つの条件を課した。

- スイッチなどの特別なデバイスを持たずハンズフリーでAHC入力ができること
- AHCの筆記動作以外に文字を区切るための不自然な動作を導入しないこと

これらの条件を達成するために、AHC筆記の際の特別な文字区切り動作に頼らずAHCのストロークの性質だけから文字ストロークと移動ストロークを判別し、文字の分割を行う。以下の節では、AHCを構成するストロークの“移動ストロークらしさ”を数値化する5つの指標を提案する。なお、5つの指標の組をベクトルとして扱い、ベクトル量子化を行ってストローク判別に用いるため各指標の取る値が $[0, 1]$ となるように各指標を設計している。

3.3.1 直線性指標 I_{ss}

直線性指標は、ストロークの直線らしさを評価する指標である。移動ストロークは、文字から文字へ移る際に無意識に最短距離で描かれるため、直線である可能性が高い。このため、直線性指標はストロークの移動ストロークらしさの指標となる。

m 番目のストローク S_m に含まれるサンプル点に対して最小二乗法による回帰直線を求める。

$$\begin{aligned} \hat{y} &= a_m x + b_m \\ a_m &= \frac{\sum_{p_n \in S_m} (x_n - \mu_{xm})(y_n - \mu_{ym})}{\sum_{p_n \in S_m} (x_n - \mu_{xm})^2} \\ b_m &= \mu_{ym} - a_m \mu_{xm} \end{aligned} \quad (4)$$

式(4)において、 $\mu_m = (\mu_{xm}, \mu_{ym})$ はストローク S_m に含まれるサンプル点の重心である。ストローク S_m の直線性指標は、回帰直線の寄与率として以下の式で定義する。

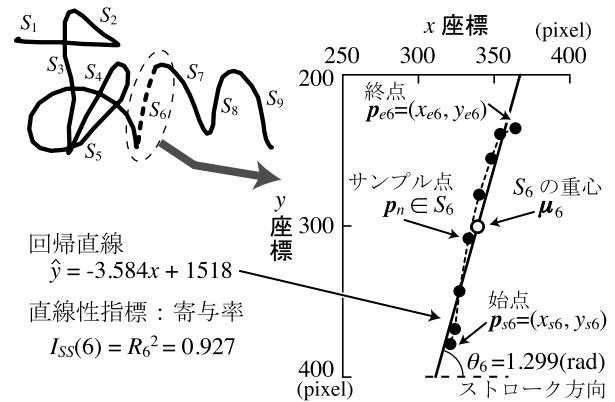


図 6 ストロークの直線回帰と直線性指標

Fig. 6 Linear regression of a stroke and stroke straightness index.

$$I_{ss}(m) = R_m^2 = \frac{\sum_{p_n \in S_m} (\hat{y}_n - \mu_{ym})^2}{\sum_{p_n \in S_m} (y_n - \mu_{ym})^2} \quad (5)$$

式(5)において、 $\hat{y}_n$ は回帰直線による y_n の近似値である。図6に直線性指標の算出例を示す。図中の文字軌跡は平仮名「あ」を空中筆記したものであり、9本のストロークに分割されている。そのうち、6番目のストローク S_6 を拡大して、直線回帰の様子を図示している。直線性指標は回帰直線の寄与率なので $[0, 1]$ の値をとり、値が大きい程ストロークが直線らしいことを示している。例示した S_6 の直線性指標 $I_{ss}(6)$ は0.927と大きな値であり、直線に近いストロークであるといえる。

3.3.2 ストローク方向指標 I_{sd}

ストローク方向指標は、ストロークが伸びる方向を評価する指標である。文字が書かれる矩形領域を上下に2等分した場合、平仮名46文字中の43文字(93.5%)は上領域から書き始められる。また、同じく43文字(93.5%)は下領域で書き終わる[13]。大部分の平仮名が上から書き始められ、下で書き終わるため、左から右へ平仮名を連続して書いた場合には、移動ストロークは右上方向を向く傾向がある。この性質を評価するために、ストローク S_m について先に求めた回帰直線の傾き a_m とストロークの始点 $p_{sm} = (x_{sm}, y_{sm})$ 、終点 $p_{em} = (x_{em}, y_{em})$ を用いて、ストローク方向 θ_m を次式で求める。

$$\theta_m = \begin{cases} \theta, & (x_{em} - x_{sm}) \geq 0 \\ \theta - \text{sgn}(\theta)\pi, & (x_{em} - x_{sm}) < 0 \end{cases} \quad (6)$$

式(6)において、 $\theta = \arctan(-a_m)$ である。

ストローク方向指標は、経験的に $-1/2\pi$ と $5/8\pi$ で最小値0.0となり、 $1/4\pi$ で最大値1.0となるように4次関数を当てはめて設定し、4次関数の値が負になる場合には、ストローク方向指標の値が0.0となるように式(7)で定義した。式(7)のストローク方向指標関数の概形を図7に示す。

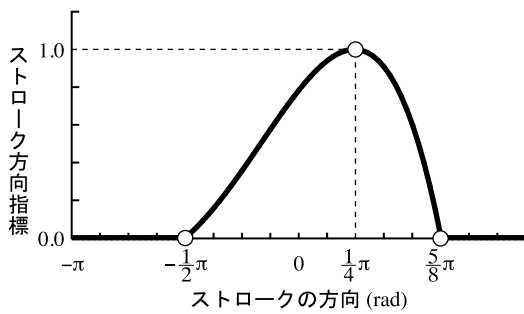


図 7 ストローク方向指標関数

Fig. 7 Stroke direction index function.

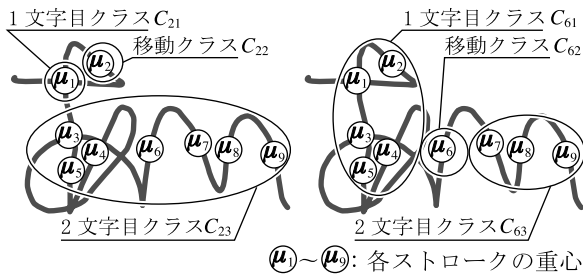


図 8 ストロークのクラス分けの例

Fig. 8 Example of class separation of strokes.

$$Isd(m) = \max \left(\sum_{i=0}^4 c_i \theta_m^i, 0 \right)$$

$$c_4 = -0.014, c_3 = -0.119, c_2 = -0.164, c_1 = 0.482, c_0 = 0.785 \quad (7)$$

3.3.3 クラスタ指標 Isc

クラスタ指標は、ストロークの集まり方を評価する指標である。一般に文字は複数のストロークで形成されているため、文字ストロークは一塊の集団になる。一方で、移動ストロークは文字の一部ではないため、その周辺にはストロークが少なく、孤立する傾向がある。右方向に2文字を書いた場合、移動ストロークを挟んで左右に文字ストロークが集まっているといえる。この性質を評価するために、クラスタ指標を以下の手順で定義する。

k 番目のストローク S_k , $2 \leq k \leq M-1$ を移動ストロークと仮定して、ストロークの重心 $\{\mu_m\}$ を以下の3つのクラスに分割する。

- 1文字目クラス: $C_{k1} = \{\mu_m | m < k\}$, ストローク数 M_1
- 移動クラス: $C_{k2} = \{\mu_m | m = k\}$, ストローク数 M_2
- 2文字目クラス: $C_{k3} = \{\mu_m | m > k\}$, ストローク数 M_3

図 8 は「あい」のストロークの重心を3つのクラスに分割した例である。

上記のようにクラス分割を行った場合、移動ストロークが正しく仮定されたときに、文字ストロークの集団がよく集まり、3つのクラスの分離度が最大になる。この性質を評価するために、移動ストロークを S_k と仮定したときのクラスタ指標 $Isc(k)$ を上記クラスのクラス内分散とクラ

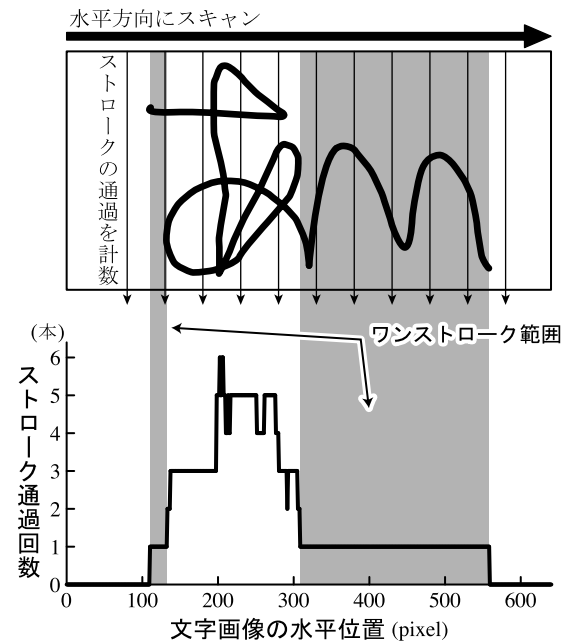


図 9 ワンストローク範囲の例

Fig. 9 Example of one stroke range.

ス間分散を用いて以下のように定義する。

クラス内分散

$$\sigma_w^2(k) = \frac{1}{M} \sum_{i=1}^3 \sum_{\mu_m \in C_{ki}} \{(\mu_{xm} - \rho_{xki})^2 + (\mu_{ym} - \rho_{yki})^2\} \quad (8)$$

クラス間分散

$$\sigma_b^2(k) = \sum_{i=1}^3 \frac{M_i}{M} \{(\rho_{xki} - \rho_x)^2 + (\rho_{yki} - \rho_y)^2\} \quad (9)$$

ただし、 $\rho_{ki} = (\rho_{xki}, \rho_{yki})$ は C_{ki} に属する μ_m の重心であり、 $\rho = (\rho_x, \rho_y)$ はすべての μ_m の重心である。

$$Isc(k) = \frac{\sigma_b^2(k)}{\sigma_w^2(k)} \quad (10)$$

3.3.4 ワンストローク指標 Ios

ワンストローク指標は、筆記面を1本だけ横切るストロークを発見するための指標である。見方を変えると、ストロークの集まり具合を評価するもう1つの指標ともいえる。

普段我々が紙に文字を書く場合、筆を上げる動作により文字を分かち書きする。AHCの筆記には筆上げ動作がなく一筆書きで文字が書かれるため、文字と文字の間隙にはストロークが1本だけ存在することになる。この性質を定量的に評価するために、図 9 のように、筆記面の水平方向1ピクセルごとに垂直方向にストロークの通過数を計数する。このとき、曲がったストロークが同じ水平位置を複数回通過した場合も通過数として計数する。これにより、図 9 下図のように水平位置を何本のストロークが通過したか知ることができる。ストロークの通過数が1回だけの水

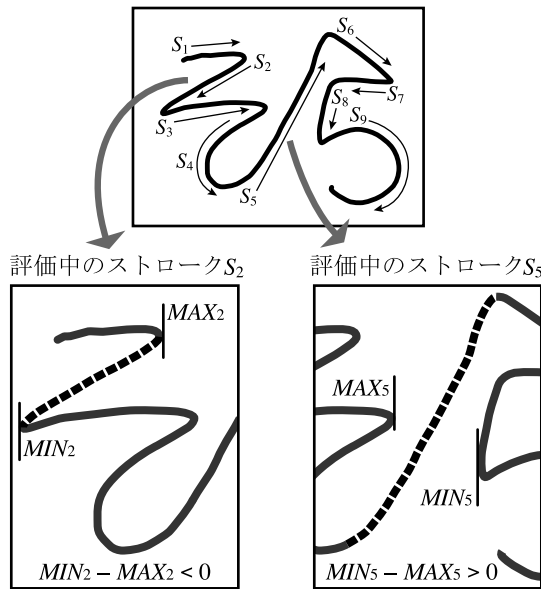


図 10 MAX_m と MIN_m の例
Fig. 10 Example of MAX_m and MIN_m .

平位置の範囲をワンストローク範囲と呼ぶ。

m 番目のストローク S_m がワンストローク範囲を通過する水平方向の長さをワンストローク長 o_m (pixel) とし、ワンストローク長の中の最大値を $o_{\max} = \max\{o_m\}$ とする。このとき、ワンストローク指標 $Ios(m)$ を以下で定義する。

$$Ios(m) = \frac{o_m}{o_{\max}} \quad (11)$$

ワンストローク指標 $Ios(m)$ は、筆記面上で最大のワンストローク長を持つストロークで最大値 1 となり、ストロークの全長がワンストローク範囲外の場合は最小値 0 となる。

3.3.5 MAXMIN 指標 Imm

MAXMIN 指標は、評価するストロークに対する先行ストロークと後続ストロークの配置を評価する指標である。

m 番目のストローク S_m について、その先行ストローク群 $\{S_i | i < m\}$ に含まれるサンプル点の x 座標の最大値を MAX_m とする。後続ストローク群 $\{S_i | i > m\}$ に含まれるサンプル点の x 座標の最小値を MIN_m とする。

MAX_m と MIN_m の実際例を図 10 に示す。図 10 は「それ」と書かれた AHC の文字軌跡の例であり、 S_1 から S_9 の 9 本のストロークに分割されている。

図 10 右下の例では、移動ストロークである S_5 について、 $MIN_5 - MAX_5$ の値は正となっている。評価対象となるストロークが移動ストロークである場合、先行ストロークは先行文字を構成し、後続ストロークは後続文字を構成するため、両者は評価対象の左右に分かれて配置される。このため、先行ストロークの最大値 MAX_m よりも後続ストロークの最小値 MIN_m が大きくなり、 $MIN_m - MAX_m > 0$ となる。また、この値は 2 つの文字がはっきり離れているほど大きな値となる。

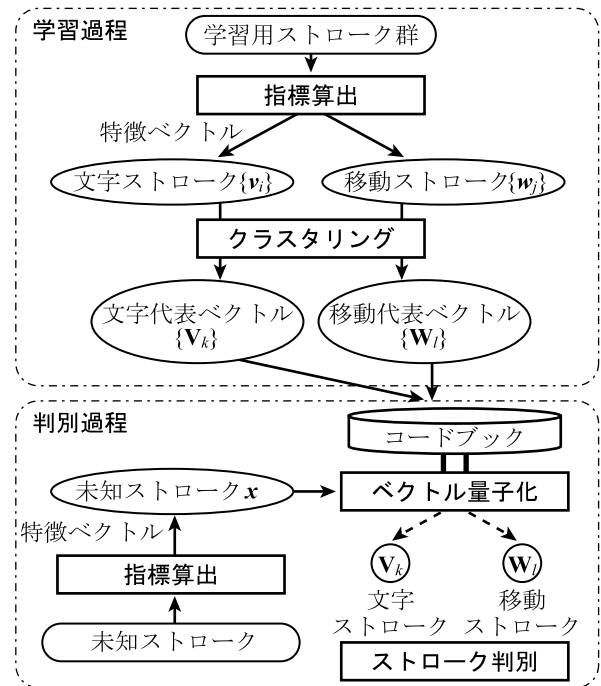


図 11 ベクトル量子化とストローク判別のブロック図
Fig. 11 Block diagram of VQ and stroke distinction.

一方、図 10 左下に示した文字ストローク S_2 については、 $MIN_2 - MAX_2$ の値が負になっている。評価対象となるストロークが文字ストロークである場合、その前後のストロークは評価対象のストロークと同一文字を構成する 경우가多く、それらに含まれるサンプル点の x 座標は近い値を取る。このため、 $MIN_m - MAX_m$ は比較的小さな値を取り、負の値になりやすい。

上述のような性質を持つ $MIN_m - MAX_m$ を用いて、ストローク S_m の MAXMIN 指標 $Imm(m)$ を以下のように定義した。

$$Imm(m) = \frac{d_m}{d_{\max}} \quad (12)$$

式 (12) において、 $d_m = \max(MIN_m - MAX_m, 0)$ 、 $d_{\max}$ は $\{d_m\}$ の最大値である。

3.4 ベクトル量子化とストロークの判別

AHC の文字軌跡を分割して得られた各ストロークについて、3.3 節で定義した 5 つの指標の値を並べて 5 次元の特徴ベクトルとし、図 11 に示す手順でベクトル量子化を応用してストロークの判別を行う。

学習過程では、学習用の文字軌跡データを分割したストローク群から、文字ストロークの特徴ベクトル $\{v_i\}$ と移動ストロークの特徴ベクトル $\{w_j\}$ を得る。 $\{v_i\}$ と $\{w_j\}$ を、k-means 法により別々にクラスタリングして、文字代表ベクトル $\{V_k\}$ と移動代表ベクトル $\{W_l\}$ を得る。この両者をまとめてベクトル量子化のコードブックとする。

ストロークの判別過程では、筆記された AHC から得られた未知ストロークの特徴ベクトル x を先のコードブック

を用いてベクトル量子化する．このとき， $\mathbf{x}$ が文字代表ベクトル $\mathbf{V}_k$ にコーディングされれば未知ストロークを文字ストロークと判別し，移動代表ベクトル $\mathbf{W}_l$ にコーディングされれば未知ストロークを移動ストロークと判別する．

4. 評価実験

4.1 文字軌跡データの収集

4.1.1 文字軌跡データ収集環境

提案手法の評価実験のために，図 12 に示す実験環境で文字軌跡データを収集した．被験者はディスプレイとカメラから約 1m の距離で椅子に座った状態で，ディスプレイに表示された AHC の文字軌跡を確認しながら文字を書く．ディスプレイ表示は文字軌跡のみであり，文字の分割結果の表示は行わない．分割結果を表示して文字軌跡データ収集を行うと，被験者が提案システムの動作に慣れてシステムが分割しやすい文字を書くようになってしまうためである．文字の分割結果を表示しないことにより，システムの動作に適応していない筆記者の自然な筆記の文字軌跡データを収集することができる．

4.1.2 文字軌跡データ収集の手順

文字軌跡データの収集では，① 実験者が被験者に対してシステムの説明を行い，② AHC 筆記の手本を見せ，③ 被験者に 2, 3 回程度の練習筆記をさせた後，④ 文字軌跡データの収集を行った．筆記する文字については口頭と張り紙による掲示で指示し，筆記開始のタイミングは被験者に任せた．

4.1.3 収集した文字軌跡データの詳細

ベクトル量子化のコードブックを作成する学習用データとして，AHC 入力システムに慣れた 20 代男性 1 名が表 2 に示す 2 文字 50 組を書いた文字軌跡データを収集した．表 2 の 50 組は，乱数プログラムによって無作為に生成した平仮名の組から国語辞典に掲載されている単語を残す方法で作成した．この 50 組には，文字ストローク 386 本，移動ストローク 50 本が含まれる．

一方，提案手法の open テストのための評価用データとして，AHC 入力システムを初めて使用する 20 代から 40 代の男性 10 名，女性 5 名の計 15 名が書いた文字軌跡データを収集した．筆記文字は，1 名につき無作為に選んだ 2 文字 10 組であり，文字ストローク $86 \times 15 = 1,290$ 本，移動ストローク $10 \times 15 = 150$ 本である．

4.2 ストローク分割の評価

提案手法を用いて文字軌跡をストロークに分割した一例を図 13 に示す．分割結果がよく分かるように，ストローク分割されるごとに文字軌跡に交互に色付けをして図示した．図 13 に示すように，分割失敗の多くは ① 書き淀み，② 書き迷い，③ ストローク分断，④ 移動ストローク消失，⑤ ストローク連結の 5 種類に分類できる．



図 12 AHC データの記録風景

Fig. 12 AHC data recording.

表 2 乱数によって生成した平仮名 50 組

Table 2 50 sets of Hiragana characters generated randomly.

あい, あか, あさ, あし, あな, あり, いか, いし, いす, いぬ
いも, えき, えん, おす, おと, かき, かた, かめ, かわ, きく
くし, さか, さし, しろ, すす, すな, たい, たて, たに, てん
なし, ぬの, ねこ, のり, はり, はん, ひく, ひる, ふく, へい
へた, ほね, まつ, まる, みそ, むり, めん, もも, るす, ろし

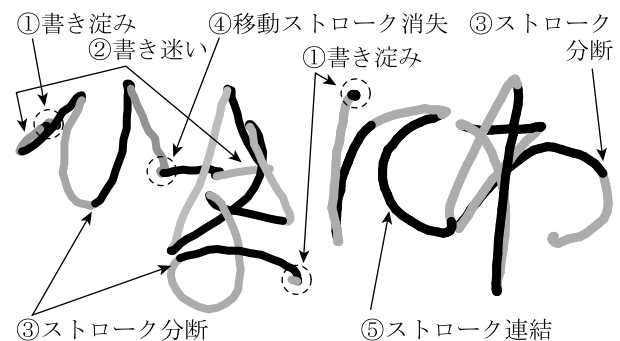


図 13 ストローク分割失敗の例—失敗を 5 つに分類した

Fig. 13 Failure case examples of stroke separation – Failures are classified into 5 categories.

① 書き淀みは，AHC の書き始めや書き終わりで起こりやすく，ごく短い余剰ストロークが生じる．② 書き迷いは，AHC の筆記中に筆記者が次の画を書く前に筆運びを迷ったために余剰ストロークが生じる．③ ストローク分断は，本来 1 本であるべきストロークが複数のストロークに分断される現象であり，ストロークが曲線で筆記途中に減速加速する場合に起こりやすい．④ 移動ストローク消失は，文字の組合せや筆記者の書き癖により，前後の文字の画が移動ストロークを介さず直接接続する現象である．⑤ ストローク連結は，ストロークの切れ目が発見できず，2 本のストロークがつながる現象である．

表 3 と表 4 にストローク分割の結果を示す．表中の“理想本数”は平仮名の正式な画とその間をつなぐ軌跡を文字ストロークとし，移動ストロークが存在するとして数えたストローク数である．“分割結果”は本手法により文字軌跡

表 3 AHC 軌跡から分割されたストロークの内訳 (学習用データ)

Table 3 Breakdown of stroke separation from AHC trajectories (training data).

ストローク種別		文字(本)	移動(本)	全体
理想本数		386	50	436
分割結果	総数	442	49	491
	成功	360(93.3%)	49(98.0%)	409(93.8%)
	失敗	82	0	82
	①②③	27	1	27
	④	—	0	0

表 4 AHC 軌跡から分割されたストロークの内訳 (評価用データ)

Table 4 Breakdown of stroke separation from AHC trajectories (evaluation data).

ストローク種別		文字(本)	移動(本)	全体
理想本数		1290	150	1440
分割結果	総数	1898	156	2054
	成功	1266(98.1%)	141(94.0%)	1407(97.7%)
	失敗	632	15	647
	①②③	24	2	24
	④	—	9	9

をストロークに分割した結果である。“総数”は分割の結果得られたストロークの数である。“成功”は正しく分割されたストローク数であり、“失敗”は先に述べた分割失敗の分類別に集計した。

① から ③ の分割失敗のために、余剰なストロークが生じ、“理想本数”に対して“総数”が多くなっている。① および ② の分割失敗で生じた余剰ストロークは文字ストロークとして扱い、③ ストローク分断で生じた余剰ストロークは、分断されたストロークと同じ種別として扱った。このような取扱いをすることで、「移動ストロークを発見して AHC を文字分割する」という本研究の目的に関しては、④ 移動ストローク消失以外の分割失敗は大きな影響がない。

表 3 と表 4 に見るように、ストローク分割は 90%以上の高い成功率となっている。表 3 の分割成功率が表 4 の分割成功率より低くなった原因は、書き癖の個人差と、システムに慣れた筆記者の場合筆記速度が上がり分割成功率がやや低下する傾向があるためである。

4.3 コードブックの作成

学習用データから得たストローク (文字ストローク 442 本, 移動ストローク 49 本) の特徴ベクトルを用い, k-means 法によってベクトル量子化のためのコードブックを作成した。クラスタ数は、文字ストローク 16, 移動ストローク 16 とした。このため、コードブックの代表ベクトル数は 32 である。また、初期乱数を変えて k-means 法を 10 回繰り返して、学習用データに対して移動ストロークと文字ストロークの平均正解率が最も高いものをコードブックとして採用した。

表 5 移動ストロークと文字ストロークの判別正解率

Table 5 Detection accuracy of transition and character strokes.

ストローク種別	closed テスト (学習用データ)			open テスト (評価用データ)		
	ストローク(本)		正解率 (%)	ストローク(本)		正解率 (%)
	正解	資料		正解	資料	
移動	重複なし	理想本数	96.0	重複なし	理想本数	73.3
	48	50		110	150	
	重複あり	総数	98.0	重複あり	総数	74.4
	48	49		116	156	
文字	411	442	93.0	1685	1898	88.8

4.4 ストローク判別の評価

提案手法の性能を評価するためストローク判別実験を行った。学習用データによる closed テストと評価用データによる open テストの結果を表 5 に示す。

文字ストロークについて、① から ③ の分割失敗で生じた余剰ストロークが誤って移動ストロークと判別されると、AHC が正しく文字に分割されない。⑤ ストローク連結によって文字ストロークが失われても、文字分割の結果に悪影響を与えない。このため、文字ストロークの正解率は、表 3 および表 4 の総数に対する正解の比率として算出した。

一方、移動ストロークについて、④ および ⑤ の分割失敗でストロークが失われると文字分割は失敗する。移動ストロークが ③ ストローク分断された場合、分断されたストロークの 1 本でも移動ストロークと判別されれば、文字分割としては成功する。このため、移動ストロークの正解率は、理想本数に対する重複正解を省いた正解の比率として算出している。また、文字ストロークの評価と条件を揃える意味で、総数に対して重複正解を許す場合の正解率も併記した。

文字ストロークについては、学習用データによる closed テストと評価用データによる open テストともに 90%程度の高い正解率となった。一方、移動ストロークについては、closed テストでは 96.0%と高い正解率となったが、open テストでは 73.3%の正解率となった。

open テストでの移動ストロークの正解率の悪化の一因は、ストローク分割で起こる分割失敗 ④, ⑤ により移動ストロークが失われることである。移動ストロークの正解率を理想本数に対してではなく分割成功に対して計算すると、open テストでも 78.0% (110/141) と 8 割近い高い正解率となる。したがって、提案した 5 つの指標はうまく機能しているといえ、ストローク分割において移動ストロークを失わないようにする工夫が必要である。しかし、④ 移動ストローク消失については、文字の組合せや筆記者の書き癖に原因があるため、提案システムの枠内では回避できない。この点については、提案システムとは異なるアプローチが必要である。

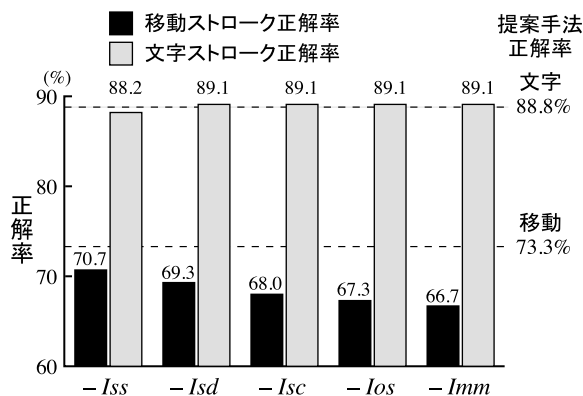


図 14 1つの指標を省略した場合のストローク正解率

Fig. 14 Stroke detection accuracy when one index is omitted.

4.5 各指標のストローク判別への貢献度

提案手法では、5つの指標 (I_{ss} , I_{sd} , I_{sc} , I_{os} , I_{mm}) を用いてストロークの判別を行っている。これら5つの指標のストローク判別に対する貢献度を調べるために、1つの指標を除いた4次元の特徴ベクトルを用いて提案手法と同じ方法でストローク判別を試みる実験を行った。以後、5つの指標のうち I_{ss} を除いて構築したシステムを“ $-I_{ss}$ ”と表記し、他の場合についても同様に表記する。

図 14 に、1つの指標を除いた4次元の特徴ベクトルを用いたシステムについて、open テストのストローク正解率を示す。文字ストロークについては、どの1つの指標を除いても提案システムと変わらない正解率が得られた。一方、移動ストロークについては、6.6% ($-I_{mm}$)~2.6% ($-I_{ss}$) 低下した。この結果から、 $I_{mm} > I_{os} > I_{sc} > I_{sd} > I_{ss}$ の順でストローク判別に関与していると考えられる。

I_{ss} や I_{sd} はストロークの直線性や方向などストローク単独の性質を反映した指標であり、 I_{sc} や I_{os} , I_{mm} は複数のストロークの集まり方や位置関係を反映した指標である。前者より後者の貢献度が高い傾向にあることから、ストローク正解率の向上のためには、ストロークどうしの関係に着目した指標の導入が有効であると考えられる。

5. おわりに

本研究では、AHC の文字軌跡をストロークに分割し、文字どうしをつなぐ“移動ストローク”を自動的に発見することによって、連続に書かれた AHC の文字分割を行う手法を提案した。提案した手法によるストローク判別について評価実験を行った結果、closed テストでは、移動ストロークと文字ストローク双方で 93%以上の正解率となった。open テストでは、文字ストロークは 88.8%，移動ストロークは 73.3%の正解率を得た。open テストに AHC 入力システムを初めて使用する利用者の筆記文字を使ったことを考慮すると高い正解率を得られたといえる。

一方で、AHC として書かれる文字の組合せや利用者の筆記の癖により、移動ストロークが消失する例など、問題点

も明らかとなった。この点を解決できれば、移動ストロークの正解率を 80%程度にできる可能性がある。提案システムの実用化に向けて、細部の調整とともに今後の課題としたい。

参考文献

- [1] 高橋賢一郎, 安田英史, 松本 隆: Hidden Markov Model を用いたオンライン手書き文字認識, 電子情報通信学会技術研究報告, PRMU96-211, pp.143-150 (1997).
- [2] 片山喜規, 内田誠一, 迫江博昭: 座標特徴と方向特徴の選択的利用に基づくオンライン文字認識 HMM, 電子情報通信学会論文誌, Vol.J91-D, No.8, pp.2112-2120 (2008).
- [3] 小林 充, 真崎晋哉ほか: オンライン手書き文字認識アルゴリズム RAV, 情報処理学会論文誌, Vol.41, No.9, pp.2536-2544 (2000).
- [4] 中井 満, 米澤久光: 加速度センサを用いた空中手書き文字認識, 第 8 回情報科学技術フォーラム講演論文集, H-019, 第 3 分冊, pp.133-134 (2009).
- [5] 中井 満, 草島広志: 筆記の向きや筆記具の持ち方が自由な空中手書き文字認識, 第 11 回情報科学技術フォーラム講演論文集, H-031, 第 3 分冊, pp.185-186 (2012).
- [6] 西村好宏, 苗村昌秀: ビデオカメラを利用した空中非目視手書き文字入力方式, 電子情報通信学会技術研究報告, ITS2005-67, pp.119-124 (2006).
- [7] 園田智也, 村岡洋一: 空中での手書き文字入力システム, 電子情報通信学会論文誌, Vol.J86-D-II, No.7, pp.1015-1025 (2003).
- [8] 大倉 充, 森田展行ほか: 空中に描かれた数式を計算するシステムの試作, 映像情報メディア学会誌, Vol.57, No.10, pp.1367-1372 (2003).
- [9] 保呂 毅, 稲葉雅幸: 複数カメラを用いた手書き文字認識システム, WISS2006, pp.121-122 (2006).
- [10] 大倉 充, 真鍋亮介, 島田英之, 島田恭宏: 空中に指先で描かれた続け書き数字の認識アルゴリズム, 情報処理学会論文誌, Vol.52, No.2, pp.910-916 (2011).
- [11] 樋口知之: 粒子フィルタ, 電子情報通信学会誌, Vol.88, No.12, pp.989-994 (2005).
- [12] OpenCV.jp: OpenCV.jp (オンライン), 入手先 (<http://opencv.jp/>) (参照 2014-04-09).
- [13] こどもの習い事と家庭学習: ひらがな書き取り練習プリント, こどもの習い事と家庭学習 (オンライン), 入手先 (<http://www.kids-points.com/drill/hiraganap.html>) (参照 2014-04-09).

推薦文

情報処理学会中国支部表彰規定に則り、平成 24 年度 (第 63 回) 電気・情報関連学会中国支部連合大会で発表された中から特に優秀であることが認められた優秀論文発表賞を授賞した論文である。

(情報処理学会・元中国支部長 三池秀敏)



鈴木 慶 (正会員)

2012 年鳥取大学工学部知能情報工学科卒業。2014 年鳥取大学大学院工学研究科博士前期課程修了。手書き文字認識をはじめとするジェスチャー・インタフェイスの研究に従事。



清水 忠昭

1987 年大阪大学基礎工学部生物工学科卒業。同年鳥取大学工学部助手。2002 年博士（工学）。現在、鳥取大学大学院工学研究科准教授。音声信号処理とインタフェイス関連の研究に従事。



吉村 宏紀 (正会員)

1993 年鳥取大学工学部知能情報工学科卒業。1997 年、鳥取大学工学研究科博士後期課程情報生産工学専攻修了。博士（工学）現在、鳥取大学工学研究科情報エレクトロニクス専攻、助教。画像、音声のデジタル信号処理の研究に従事。電子情報通信学会、情報処理学会の会員。



松村 寿枝

1998 年鳥取大学大学院工学研究科博士前期課程修了。2000 年まで大日本スクリーン製造株式会社勤務。2003 年鳥取大学大学院工学研究科博士後期課程修了。博士（工学）。財団法人鳥取県産業振興機構勤務。2004 年より奈良工業高等専門学校情報工学科助手を経て現在、同准教授。音声情報処理の研究に従事。