

空中手書き文字認識におけるペン先方向の加速度特徴の有効性

Effectiveness of Acceleration toward the Pen-tip Direction in Aerial Handwritten Character Recognition

大坪 由香利†

中井 満†

Yukari Otsubo

Mitsuru Nakai

1. はじめに

空中に筆記する場合も文字はある平面に書かれていると思われ、文字認識の特徴量も筆記面の 2 次元特徴で十分と考えてきた。我々はこれまでに加速度センサを用いて空中に書いたひらがな文字の認識を行っており、2 次元特徴の方が良いという結果を得た [1]。一方、数字の「0」と「6」のようにペンの動きが似ている文字は、加速度特徴で識別することは困難である。しかし、木村らの研究では、それらを正しく識別できており [2]、その手法では 3 次元特徴が用いられていた。そこで本稿では、ペン先方向の加速度特徴の有効性について報告する。

2. ペン先方向の加速度特徴

本研究では、筆記具として任天堂の Wii リモコン (加速度センサ内蔵) を使用し、図 1 のように空中に文字を書く。センサの座標系は図 2 の通りである。xz 平面が筆記面、y 軸がペン先方向となる。

3 次元特徴で「0」と「6」が識別できた理由として、筆記具の姿勢の違いが考えられる。腕を水平に伸ばして文字を書くと、筆記終了時のペン先方向が「0」は上向きになるのに対し、「6」は水平な姿勢となる。図 3 に「0」と「6」を筆記したときの加速度信号を示す。静止状態でも、ペン先 (y 軸) 方向には重力加速度の成分がかかっており、これは傾きによって変化する。2 つの加速度信号を比較すると、筆記終了時に違いがあることがわかる。このことから、ペン先方向の加速度特徴が文字の縦方向の位置座標情報の代わりになると考えられる。



図 1: 筆記風景
(ただし、筆跡は見えない)

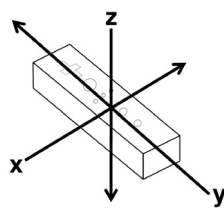


図 2: 筆記具の座標系

3. 空中手書き文字認識の構成

筆記時の加速度信号を 10 ミリ秒間隔でサンプリングする。1 文字毎にこの加速度信号の平均を原点に移動し、筆記面 (xz) とペン先方向 (y) のそれぞれについて、加速度の大きさの平均が 1 となるように正規化を行う。この信号を特徴量とし、一筆書き文字 HMM (Hidden Markov Model) を用いて尤度計算を行う。尤度が最大の文字を認識結果とする。

本研究では、ペン先方向の有効性を調査するため、特徴量に

† 富山県立大学, Toyama Prefectural University.

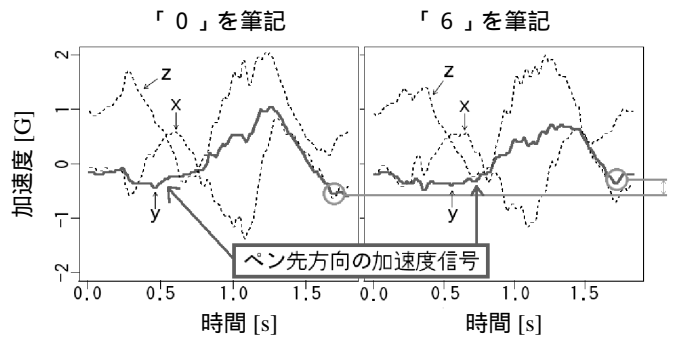


図 3: 「0」と「6」の 3 軸加速度信号

重みを付けられるマルチストリーム HMM (MS-HMM) でモデル化する。特徴量を筆記面 (xz) とペン先方向 (y) の 2 つのストリームにわけ、それぞれの対数尤度の重み付き線形和を識別関数とする。従来法では識別関数は $L_S(\omega) = \log P(x, y, z|\omega)$ と表される。ここで、 ω は認識候補の文字種である。 $P(\cdot|\omega)$ は ω の HMM から各信号が出力される確率である。これを筆記面とペン先方向が独立であると仮定して 2 つに分けると、

$$\log P(x, y, z|\omega) = \log P(x, z|\omega) + \log P(y|\omega)$$

となり、ペン先方向の重み係数を α として、

$$L_M(\omega) = (1 - \alpha) \log P(x, z|\omega) + \alpha \log P(y|\omega)$$

を MS-HMM の識別関数とする。 $\alpha = 0.0$ の時に 2 次元特徴、 $\alpha = 0.5$ の時に従来の 3 次元特徴と等価となる。

4. ペン先方向に重み付けした文字認識実験

認識対象は、数字 10 字種 (「0」～「9」) またはひらがな 71 字種 (「あ」～「ん」、濁点半濁点を含む) である。数字・ひらがな各々に対して各字種 1 文字ずつ筆記したものを 1 セットとした。数字の収集では筆記具の持ち方として、リモコン持ち (図 4) とペン持ち (図 5) の 2 種類を指定した。同一の筆記者でも持ち方によって筆記具の姿勢や腕の振り方が変わると予想され、特徴量の有効性に影響があると考えられる。ひらがなの収集では持ち方を指定せず、筆記者の自由とした。また、図 3 の筆記では腕を水平にしており、ペン先方向の加速度特徴が効きやすい条件であったが、より使用環境に近い条件で実験を行うために、数字・ひらがな共に振り方を指定しなかった。書き順は統一した。これらの条件で 5 名の筆記者から各 10 セットずつ収集した。

実験条件として筆記者 5 名のうち 4 名のデータを学習した後、学習外の 1 名で評価を行った。これを組み合わせを変えて 5 回試行した。ペン先方向の尤度の重み α を 0.0 から 1.0 まで 0.1 ずつ変化させて認識率を比較した。

図 6 に重みと認識率の関係を示す。数字は持ち方に依らず $\alpha = 0.5$, ひらがなは $\alpha = 0.4$ のとき認識率が最大となり, 数字とひらがな共にペン先方向の加速度特徴は有効であった。これは, 従来法 [1] と異なる結論であり, 2 つのストリームに分けて正規化, モデル化したことが効いたと思われる。また数字について持ち方別に比較すると, リモコン持ち・ペン持ち共に最適な重みは $\alpha = 0.5$ であり, 持ち方による影響は無かった。しかし, リモコン持ちに比べてペン持ちの方が認識率が低くなった。これは, ペン持ちの方が個人差が大きいためだと考えられる。

ペン先方向の特徴が影響した文字対の例を表 1 に示す。2 節で述べたように「0」と「6」がペン先方向の加速度特徴を使用することで識別できることが確認できた。ひらがなの「か」と「や」は, 1 画目が似ており, 2 画目と 3 画目も縦に振る動作である。「か」では「|」の後に「>」を書くため, ペン先を中位置で止める。それに対し「や」では「>」の後に「|」を書くため, 下まで振り下ろす動作となる。また「ふ」は最後の 2 画の振りが濁点と誤認識しやすいが, 「ふ」は下方で振るのに対し, 濁点は上方で振る動作となる。「0」と「6」の対と同様に, ペン先方向の加速度特徴が縦方向の位置座標情報の代わりとなり, 識別できたと考えられる。しかし「4」と「9」のようにペン先方向の特徴を付加することで逆に識別が困難になる文字対があった。理由は筆記面では異なる動きが, 上下方向だけに着目すると似た挙動をするためである。したがって, 文字種毎に適切な重みは異なっていると考えられる。



図 4: リモコン持ち



図 5: ペン持ち

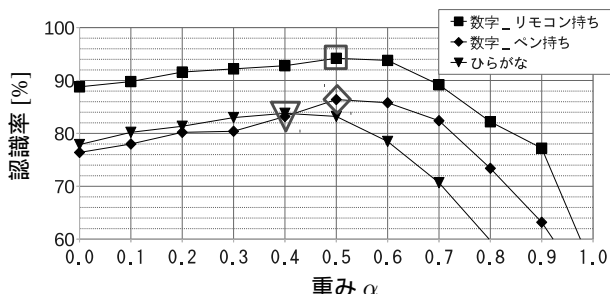


図 6: ペン先方向の重みと認識率 (%) の関係

表 1: ペン先方向の特徴が影響した文字対

辞書	可能となった例	困難となった例
数字	0 ↔ 6, 6 ↔ 9	0 ↔ 8, 4 ↔ 9
ひらがな	か ↔ や, た ↔ を, ふ ↔ ぶ	お ↔ ず, ね ↔ ぬ, ね ↔ ほ

5. サンプル毎に重みを調節した認識率の上限

図 7 は同じ筆記者が筆記した数字「6」の 10 サンプルについて識別できた重みの範囲を示している。このように同じ筆記者, 同じ文字種であってもサンプル毎に識別可能な重みは異なっている。そこで, サンプルを認識する前に筆記の環境(持ち方や筆記者, あるいはその他の要因)から最適な重みが推定できるものと仮定し, 認識率の上限について調査した。いずれかの重みで認識できていればそのサンプルは認識可能として, 認識率を算出した。図 7 の例では $\alpha = 0.5$ の場合なら 60% であるが, いずれかの重みで認識できているため, 重みを調節すれば 100% 認識できる。重み 0.0 の場合, 重み固定の場合, 重みを調節した場合の結果を図 8 に示す。重み固定の場合で約 5 割, 可変にして約 8 割の認識誤りを削減できると期待できる。

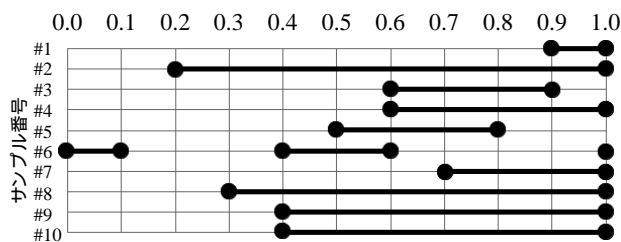


図 7: 同一筆記者の「6」10 サンプルの識別可能な重み

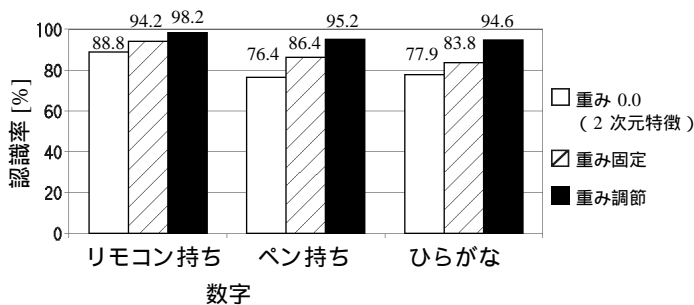


図 8: 認識率 (%) の比較

6. まとめ

ペン先方向の加速度特徴は数字・ひらがな共に有効であった。「0」と「6」のような筆記面の加速度特徴だけでは識別困難な文字対でもペン先方向の加速度特徴を付加することで識別できた。また, サンプル毎に適切な重みが推定できれば最大で 8 割の認識誤りを削減できる見込みを得た。

謝辞 本研究は科研費(24500151)の助成を受けたものである。

参考文献

[1] 中井, 米澤, “加速度センサを用いた空中手書き文字認識,” 情報科学技術フォーラム (2009-09)
 [2] 木村 他, “3 軸加速度センサを用いた動作確認に関する基本検討,” 信学総大 (2009-03)