

空中手書き文字入力インタフェースのための非筆記動作の除去

Removal of Non-Writing Motion for Aerial Handwriting Character Input Interface

吉田 大樹[†]中井 満[†]

Taiki Yoshida

Mitsuru Nakai

1. はじめに

加速度センサを内蔵した筆記具を用いて、空中に手書きした文字を認識する。これまで、筆記具を振る加速度信号のパワーの閾値を設定して筆記の開始と終了を検出した [1]。このため、筆記以外の動作の信号が文字として誤入力されないように、静かに筆記具を持つ、あるいはボタン操作で筆記の on/off を切り替える必要があった。これに対し、本研究ではボタン操作が不要な自然なインタフェースを目指す。本稿では加速度信号のフーリエ記述子を特徴抽出し、筆記動作と非筆記動作を識別する手法を提案する。

2. 加速度信号のフーリエ記述子

動作認識の特徴量として一次元の加速度信号のフーリエ変換が用いられることが多い [2]。しかし、文字を書く加速度信号は二次元の信号である。そこで、筆記動作と非筆記動作を識別する特徴量としてフーリエ記述子 [3] を用いる。フーリエ記述子は、筆点を複素平面上の点として、その時系列を離散フーリエ変換したものである。二次元平面上の曲線形状の概形を低次の特徴で近似できることから、文字認識 [4] や筆記者識別 [5] にも利用されている。

3 軸加速度センサを内蔵した任天堂の Wii リモコンを用い、図 1 のように空中に文字を書く動作の加速度信号を 10 ミリ秒間隔でサンプリングする。センサの座標系は図 2 の通り、xz 面が筆記面である。この xz 面の加速度信号の軌跡を複素平面上の時系列 $z_t (t = 0, 1, \dots, L - 1)$ とする。L は時系列の長さである。このときフーリエ記述子 c_k は

$$c_k = \frac{1}{L} \sum_{t=0}^{L-1} z_t \exp(-2\pi i t k / L),$$

$$k = -(L - 1)/2, \dots, -1, 0, 1, \dots, (L - 1)/2$$

で表される。

図 3 は文字「あ」を筆記したときの加速度信号（外接矩形を 0 から 1 に正規化）とそのフーリエ記述子のパワースペクトルである。筆記時間は約 2.3 秒であり、228 点の時系列である。図 4 は「縦に（3 回）振る」動作の場合である。「縦に振る」動作は 1 次元的な図形のためパワースペクトルがほぼ左右対称になる。一方、文字を書く場合は図形に右回り、左回りの成分があるため非対称になる。このような違いから、筆記動作と非筆記動作を識別できると考えた。

3. 筆記動作と非筆記動作の種類

収集した動作を表 1 に示す。空中手書き文字認識の対象はひらがな 71 字種（濁点、半濁点を含む）なので、空中にひら

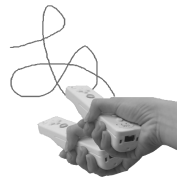


図 1: データの収集風景

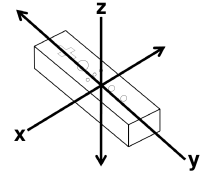


図 2: 筆記具の座標系

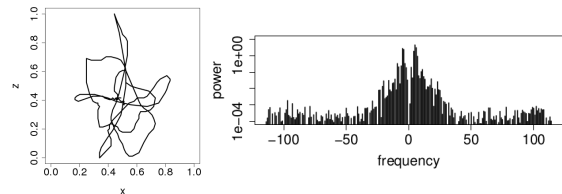


図 3: 「あ」の筆記面の加速度信号（左）とパワースペクトル（右）

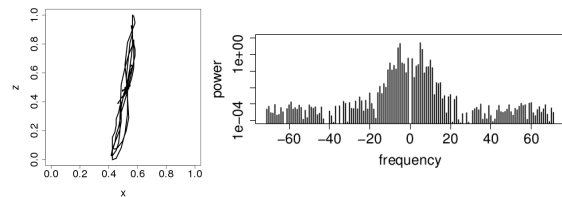


図 4: 「縦に振る」動作の筆記面の加速度信号（左）とパワースペクトル（右）

表 1: 収集した動作

筆記動作	71 種	空中にひらがな（「あ」～「ん」）を書く
非筆記動作	5 種	置く、持ち上げる、縦に振る、横に振る、歩く

がなを書くことを筆記動作とした。一方、筆記具を「置く」、「持ち上げる」、筆記具を 3 回「縦に振る」、「横に振る」、筆記具を手に持って「歩く」といった 5 動作を非筆記動作とした。筆記具を「置く」、「持ち上げる」動作は筆記前後に多い動作であり、信号のパワーの閾値を用いた検出法ではこのような短い動作でも文字として誤入力してしまう。筆記具を「縦に振る」、「横に振る」動作は 3 回振ることで、文字を筆記する時間長とほぼ同じになり、筆記動作と誤識別するかもしれないと考えたからである。また、本研究はいつでもどこでも文字入力を行えるシステムを目指しているため、筆記具を手に持って「歩く」場合も考慮した。

4. 筆記動作と非筆記動作の識別実験

筆記動作として、1 名によるひらがな 71 文字 6 セットの計 426 サンプルを収集し、非筆記動作として、表 1 の 5 動作 100 セットの計 500 サンプルを収集した。それぞれの半数を学習

[†]富山県立大学, Toyama Prefectural University.

に、残りの半数を評価に使用した。特徴量には、図5のようにフーリエ記述子の低域の20次 ($k = -20$ から 20) を用いた。ただし、 c_0 は用いない。識別手法にガウシアンカーネルのSVM (Support Vector Machine) を用いて、筆記動作と非筆記動作の2クラスに識別した。結果を表2に示す。実験の結果、筆記動作と非筆記動作を97.5%識別できた。誤識別された動作を表3に示す。「し」「つ」「置く」「持ち上げる」といった短い動作を誤識別したことがわかる。これらの動作は1次元的な動作になるので互いに誤識別したと考える。しかし、「れ」、「縦に振る」といった長い動作も誤識別されている。このサンプルと同じ動作の他サンプルとを比較したが、違いは見られなかった。

識別に必要な次数を調査するため、図6のような6通りの帯域の特徴量で比較した。それぞれの識別率を表4に示す。結果より、20次から特徴量を減らすと識別率が低下することがわかった。特に、負の側の特徴量を除くと筆記動作の識別率が著しく低下した。負の側の特徴量は時計回りの回転動作を表す。表5は全てのひらがなを、正と負のパワーの大小で2つに分類したものであり、負の回転を持つ文字の方が多い。表6はどちらか片側の特徴を用いた場合に非筆記動作に誤識別された文字種であり、正の特徴のみを用いた場合には負の回転を持つ文字が、負の特徴のみを用いた場合には正の回転を持つ文字が非筆記動作に誤ることがわかる。

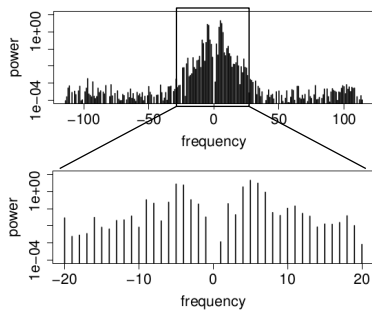


図5: フーリエ記述子特徴量

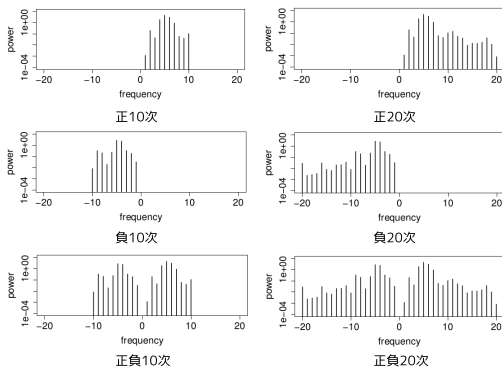


図6: 特徴量の種類

5. まとめ

フーリエ記述子特徴量の次数を変え、筆記動作と非筆記動作の識別を行った。実験結果から、正負20次の特徴を用いた

表2: 筆記動作と非筆記動作の識別結果

入力 \ 識別結果	筆記動作	非筆記動作	識別率 (%)
筆記動作	209	4	98.1
非筆記動作	8	242	96.8
合計	-	-	97.5

表3: 誤識別の内訳 (誤識別数 / サンプル数)

筆記動作	非筆記動作
「し」 2/3	「置く」 3/50
「つ」 1/3	「持ち上げる」 3/50
「れ」 1/3	「縦に振る」 2/50

表4: 特徴量の次数による識別率 [%] の比較

特徴量	筆記動作		非筆記動作	合計
	正	負		
10	0	84.0	94.4	89.2
20	0	85.0	94.4	89.6
0	10	86.3	94.8	90.5
0	20	89.2	96.0	92.6
10	10	90.6	96.8	93.7
20	20	98.1	96.8	97.5

表5: 文字ごとの正と負のパワーの違い

正の方向の回転が多い文字 ($ 正 > 負 $)	負の方向の回転が多い文字 ($ 正 < 負 $)
いきこさしせ そたてとひへ むもをぎござ じぜぞだでど びべび	あうえおかくけすち つなぬねのはふほ まみめやゆらりる れるわんがぐずじづ ばぶぼばぶべぼ

表6: 筆記動作に誤識別される文字

正のみ	か, の, く
負のみ	し, て, へ, も

場合の識別率が最も良く、筆記動作と非筆記動作を97.5%識別することができた。今後は、被験者やデータ数を増やして実験をする予定である。

謝辞 本研究はJSPS 科研費 24500151 の助成を受けて行った。

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] 嶋谷, “スマートフォンを用いたコンテンツ検索支援のための動作認識,” 信学技報 PRMU (2012-2)
- [3] 河村, 横田, “開曲線に適した新たなフーリエ記述子「I型フーリエ記述子」,” 信学論 (D-II) (2005-10)
- [4] 大友, 原, “P型フーリエ記述子を用いたオンライン手書き漢字認識,” 情処学論 (1993-02)
- [5] 中川, 小館, “フーリエ記述子を用いた平仮名の筆者識別,” 信学技報 IE (2011-04)