

短時間フーリエ記述子特徴量を用いた空中手書き文字認識

Aerial Handwritten Character Recognition using Short-Time Fourier Descriptors

吉田 大樹[†]中井 満[†]

Taiki Yoshida Mitsuru Nakai

1. はじめに

空中に書いた文字の認識システムとして、筆記具に内蔵したセンサの信号を用いる方式 [1] と外部センサで筆跡を検知する方式 [2] がある。我々は筆記具に内蔵した加速度センサの信号から文字を認識する手法において、加速度時系列信号の短時間フーリエ記述子特徴量を用いる方が良いことを報告した [3]。本稿ではカメラで取得した空中での筆跡の位置、速度、加速度の軌跡に対し、短時間フーリエ記述子の有効性を検証する。

2. 空中手書き信号の取得

空中手書き文字認識の 2 つの方式の比較のため、筆記具にかかる加速度情報と空間の筆跡情報を同時に取得する実験環境を構築した。加速度情報の取得には Wii リモコンを用い、内蔵の加速度センサの信号を 100Hz でサンプリングして Bluetooth でパソコンに送信した。同時に Kinect のハンドトラッキングツールを利用して、図 1 のようにモコンの先端部を追跡し、筆跡情報を 30Hz でサンプリングした。図 2 の Wii リモコンの座標系に合わせて、本稿ではペン先方向を y 、筆記面を xz 面 (水平方向を x 、垂直方向を z) とする。

3. 短時間フーリエ記述子特徴量 (STFD)

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 [4]。図 3 は 65 点からなる「あ」の筆跡信号全区間のフーリエ記述子のパワースペクトルのうち低域 10 次を示したものである。このようにひらがな文字はおよそ低域の 6 次までに特徴が現れている。

短時間フーリエ記述子の場合には、まず、筆跡 (位置あるいは、速度あるいは、加速度) の大きさの影響をなくすため、1 文字毎に xz 面の信号の外接矩形の長辺が等しくなるように拡大・縮小する。 x 軸を実軸に、 z 軸を虚軸にとり、時刻 t の信号を $p_t = x_t + iz_t$ とする。短時間フーリエ記述子では時刻 t の特徴として、区間 W の信号 ($p_{t-W/2} \sim p_{t+W/2}$) を分析する。一般にこの図形は閉曲線となるので、区間の始点と終点が原点に対して点対称となるように平行移動し、180 度回転した図形を重ねて閉曲線にする。以上の処理により、長さ $2W$ となる図形を $q_1 q_2 \dots q_{2W}$ とおく。このとき、短時間フーリエ記述子 c_k は

$$c_k = \frac{1}{2W} \sum_{s=1}^{2W} q_s \exp(-\pi i s k / W),$$

で表される。図 4 は $W = 14$ (約 0.4 秒)、フレームシフト 0.09 秒で分析したスペクトログラムである。筆跡の一区間に



図 1: データの収集風景

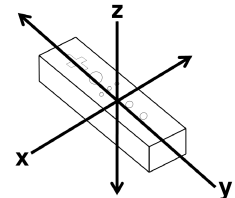


図 2: 筆記具の座標系

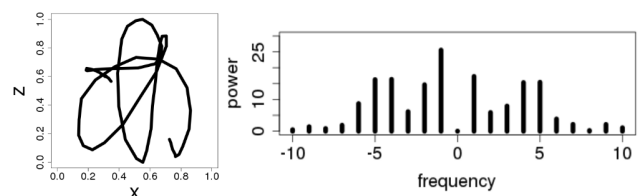


図 3: 「あ」(位置信号)の全区間フーリエ記述子のパワースペクトル

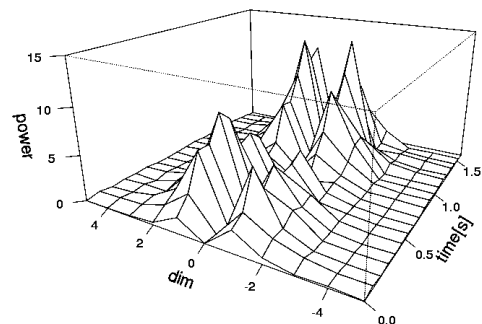


図 4: 「あ」(位置信号)の短時間フーリエ記述子のパワースペクトログラム

対するフーリエ記述子なので低域の 2 次程度までに特徴が現れている。全区間のフーリエ記述子 (図 3) の場合は、1 文字筆記する時系列の長さを L として、 $W = L$ の処理となる。

4. 空中手書き文字の認識実験

ひらがな 71 字種の認識実験を行った。各字種を一文字ずつ筆記したものを 1 セットとし、各被験者から 10 セットを収集した。被験者 33 名のデータを 17 名と 16 名のデータに分けて交差確認を行った。空中手書きの筆跡の位置信号とそれを微分または 2 回微分して求めた速度、加速度信号のいずれかを用いて文字認識実験を行った。手法 1 は、位置、速度、加

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

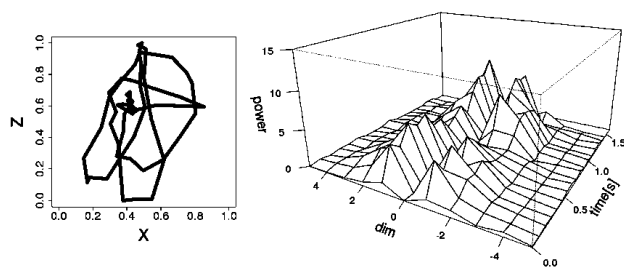


図 5: 「あ」(速度信号)の短時間フーリエ記述子のパワースペクトログラム

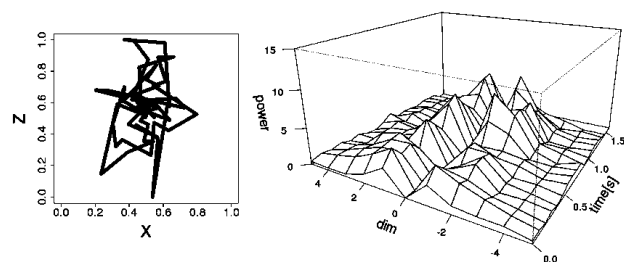


図 6: 「あ」(加速度信号)の短時間フーリエ記述子のパワースペクトログラム

速度信号をそのまま特徴量に用いた。なお、先行研究 [1] では筆記具に内蔵したセンサの信号を用いたので、センサの回転補正等の前処理が必要だったが、この実験では Kinect で取得した空間座標系の信号なので、前処理はしなかった。手法 2 は、位置、速度、加速度の全区間のフーリエ記述子の低域 10 次 ($k = -10 \sim 10$, 位相情報も含む) までは特徴量に用いた。手法 3 は、位置、速度、加速度の短時間フーリエ記述子の低域 3 次までを特徴量に用いた。なお、手法 2 では 1 状態の GMM (Gaussian Mixture Model), 手法 1 と手法 3 では時系列パターンを扱うので文字種毎に状態数の異なる HMM (Hidden Markov Model) を識別器に用いた。

手法 2 の結果を表 1 に示す。全区間の信号を用いた場合は位置、速度、加速度のいずれも低域 6 次までを用いた場合の認識率が最も高い。これは前節でも述べたが、ひらがな文字の特徴がおおよそ低域の 6 次までに現れていたからだと考えられる。

手法 3 の結果を表 2 に示す。低域 2 次までを用いた場合の認識率が最も良い。図 5, 図 6 はそれぞれ「あ」の速度または加速度の短時間フーリエ記述子のパワースペクトログラムである。図 4 と同様に、低域 2 次までに特徴が現れていることが分かる。

表 3 は 3 つの手法の最高認識率をまとめたものである。短時間フーリエ記述子を用いた手法が位置、速度、加速度のいずれにおいても最も認識率が高い。全区間のフーリエ記述子を用

表 1: 全区間フーリエ記述子を用いた場合 [%]

入力 次数	位置	速度	加速度
低域 2 次	44.3	38.8	29.3
低域 4 次	72.6	72.2	64.2
低域 6 次	76.3	76.1	71.3
低域 8 次	73.1	74.0	69.8
低域 10 次	68.0	63.6	62.6

表 2: 短時間フーリエ記述子を用いた場合 [%]

入力 次数	位置	速度	加速度
低域 1 次	84.3	82.4	78.0
低域 2 次	87.2	84.5	80.9
低域 3 次	84.5	82.0	78.7

表 3: 3 手法の比較 [%]

手法	入力	位置	速度	加速度	Wii 加速度
1. 信号		62.9	65.6	66.3	63.9
2. 全区間		76.3	76.1	71.3	69.7
3. 短時間		87.2	84.5	80.9	79.7

いる手法でも、原信号をそのまま特徴にするより認識率が高い。なお、表 3 に Wii リモコンの加速度信号を用いた場合の結果を併記した。Kinect は実空間座標系, Wii リモコンは筆記具の座標系という違いはあるが、どちらも同じ傾向にあり、短時間フーリエ記述子が最も良いという結果であった。

5. おわりに

フーリエ記述子を用いて、空中手書き文字の認識を行った。フーリエ記述子を用いることで、原信号をそのまま特徴にする手法より高い認識率を得ることができた。また、全区間のフーリエ記述子より、短時間のフーリエ記述子を用いた方が認識率は良いことが分かった。ところで、手法 1 のように時系列信号を特徴量とする先行研究 [1, 5] では不特定筆記対策のための前処理や、3 軸加速度と 3 軸角速度を合わせた特徴量の高次元化によって、認識率を約 90%まで改善している。今後はそれらの前処理と特徴量を活かして改善していく予定である。

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] 園田, 村岡, “空中での手書き文字入力システム,” 信学論 (D-II) (2003-7)
- [3] 吉田, 中井, “フーリエ記述子を用いた空中手書き文字認識,” 信学総大 (2014-3)
- [4] 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)
- [5] 大坪, 中井, “空中手書き文字認識における重力加速度成分の除去,” 信学総大 (2014-3)