

Encoder-Decoder モデルによるオフライン筆跡からの動的特徴の推定 Estimation of Dynamic features from Offline Handwriting Using Encoder-Decoder Model

細江 麻梨子^{†‡}
Mariko Hosoe

山田 智輝[‡]
Tomoki Yamada

加藤 邦人[‡]
Kunihito Kato

1. はじめに

筆跡学において、筆跡とは書字運動の一部が紙面に固定されたものと考えられている[1]。筆跡は単に機械的に印刷された文字とは異なり、人の運動特性が反映されていることから、筆跡から筆者を識別することが可能であるということが筆跡鑑定の根拠として存在する[2]。筆跡と書字運動の関係についての研究としては、ペンタブレットや筋電図から得られた書字運動データをもとに運動のモデル化を目的としたものや、文字認識精度の向上のために書字運動を考慮した辞書を作成することを目的としたものがある[3][4][5]。

筆跡鑑定の現場においては、オフラインの画像情報から得られる字画構成と字画形態に加え、実体顕微鏡等を用いて画線の交錯や筆記具のかすれ等から推定される筆順や運筆方向及び紙面の凹み具合から推定される筆圧等の動的情報が鑑定人により検査される[6]。また、英語圏の筆跡鑑定支援システムとして提案されている WANDA[7]においても、鑑定対象となるオフライン筆跡を検査者がペンタブレットでなぞり書きをすることで疑似的に時系列情報を付加し、その運筆方向の特徴が検査に用いられている。

動的特徴は鑑定人の詳細な検査により得られるものであるが、動的特徴とオフライン筆跡にはある一定の関係があることが先行研究で述べられている[8][9]。また、何らかの筆の動きを想定してすでに描画されたオフライン画像から、その描画過程を深層強化学習により推定する方法も提案されている[10]。本研究では筆跡鑑定のための基礎的研究としてオフライン筆跡と動的特徴の関係性を調べることを目的とし、オンラインデータから得られる動的特徴とオフライン筆跡の関係性を Encoder-Decoder 構造でモデル化し、オフライン筆跡から書字運動が推定可能か否かを検討した。

2. 提案手法

書字運動を把握するために、ペンタブレットを用いてオンラインデータを収集するとともに、一般公開されているオンラインデータベースである UJI データ[11]を用いた。収集したデータについては、実際に記載された用紙からオフライン画像を取得可能であるが、画像を使うにあたり、枠線の除去やオンライン情報の解像度とオフライン画像のサイズ調整等の前処理が必要となる。また、UJI データは文字を書くときのペン先座標 (x,y) 情報のみしか含まれておらず、紙に書かれたオフライン画像は存在しない。

本研究の目的は、オフライン筆跡の形態情報から動的特徴を推定することであるため、ペン先座標情報のみから生成される筆跡画像をオフライン筆跡として扱うこととした。また、動的情報である筆圧、速度、時間成分については、カラーマップの運動情報画像として相対筆圧分布画像、相対速度分布画像、時系列画像を生成し、オフライン筆跡と運動情報画像の関係性を図 1 に示す U-Net[12]構造の Encoder-Decoder モデルによるピクセル単位の対応付け学習を行った。

2.1 書字運動データの収集

人が文字を書く過程を調べるためには、ペンタブレット等により書字運動データを取得する必要がある。本研究では、ひらがな文字を記載するときのオンラインデータを収集するとともに、一般に公開されている UJI データのラテン文字及び数字を記載するときのオンラインデータを用い、字種の違いによる影響についても検討した。

2.1.1 ひらがな文字データ

ペンタブレット (Wacom 製 Intuos4) を用いて、筆者 15 名から、10 字種のひらがな文字を記載するときのオンラインデータを収集した (図 2 参照)。各筆者は、ボールペン

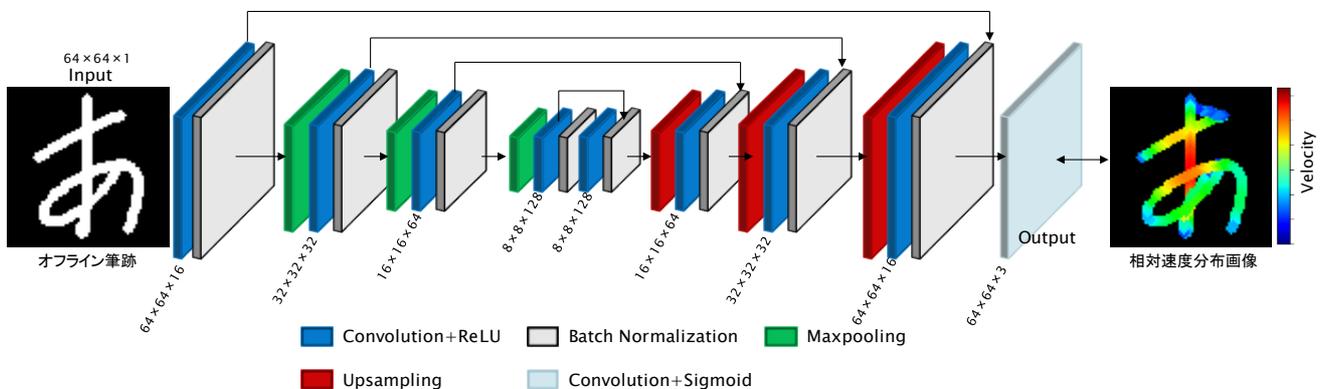


図 1 動的特徴を推定するための Encoder-Decoder モデル

[†] 岐阜県警察本部刑事部科学捜査研究所

[‡] 岐阜大学工学部

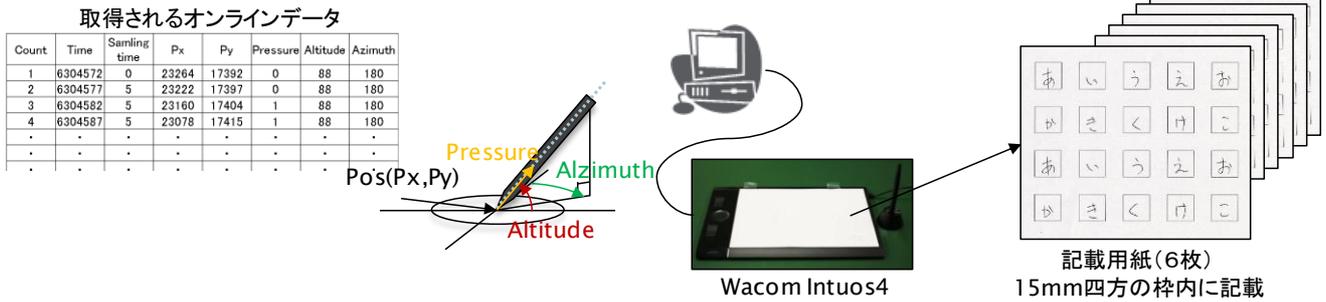


図 2 ひらがなデータの収集状況

芯を装着した intuos4 専用のインクペンにより実際にペンタブレット上に置かれた用紙に文字を記載するとともに、ペンタブレットが接続された PC により Wintab 対応の自作ソフトウェアを用いて運動情報 (ペン先座標、筆圧、ペンの方位角、ペンの仰角) を取得した。本研究で用いたペンタブレットの分解能は縦横ともに 2000Line/cm であり、サンプリングレートは 0.05ms である。また、ペンの情報は電磁誘導方式により取得されるものであり、ペンタブレット本体からペン先が約 1cm 離れた位置の情報も取得可能である。1 文字を記載するとき、1cm を超えてペンを紙面から離す筆者はみられず、空筆を含む運筆情報が取得された。

各筆者は 1.5cm×1.5cm の記載枠に「あいうえおかきくけこ」の 10 字種を、6 回は普通で速く記載するように教示し、6 回を速く記載するように教示してデータを収集した。いずれの筆者においても速く書くように教示した場合には、普通で速く書くように教示した場合よりも記載時間が短く、教示通りにデータが採取された。同教示により、少数の筆者からではあるが、筆者の個人内の変動が捉えられるオンラインデータが収集された。

2.1.2 UJI データ

UJI データは、ペンタブレット入力対応の PC により取得されたラテン文字 (大文字・小文字)、数字等を記載するときのオンラインデータである。各筆者はパソコン画面上の記載枠に従って、各字種を 2 回記載したデータが含ま

れる。UJI データには初期に収集された筆者 11 人のみのデータと、その後に収集された 49 人の筆者のデータに分けられ、モデルの学習には初期に収集された筆者 11 人中の 10 人のデータを用い、モデルの評価実験に筆者 49 人のデータを用いた。なお、UJI データはひらがなデータとは異なり、筆圧情報及び空筆部の情報は収録されておらず、運動情報はペン先座標のみである。

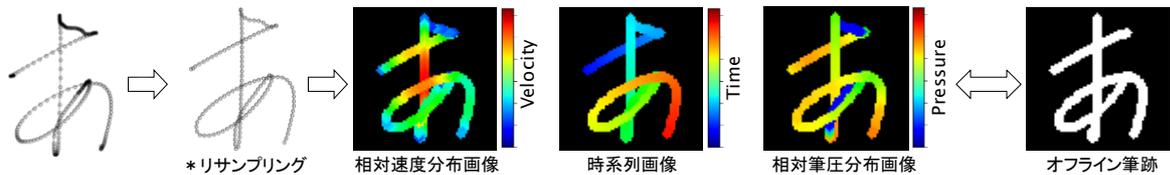
2.2 前処理

ひらがな文字データと UJI データについて、各サンプリング点間の距離に基づき速度計算するとともに、オンラインデータのペン先座標を基にペン先の軌道を画像化し、縦横比を保持した大きさの正規化(64×64pixels)、位置の中心化を行った。なお、ひらがなデータについては空筆部の情報も取得されていることから、等間隔 (約 5pixel 長) に線形補間でリサンプリングを行った。上記の前処理により、オフライン筆跡及び 1 文字記載時の相対速度、相対筆圧、時系列情報をカラーマップにより表現した運動情報画像を生成した (図 3 参照)。

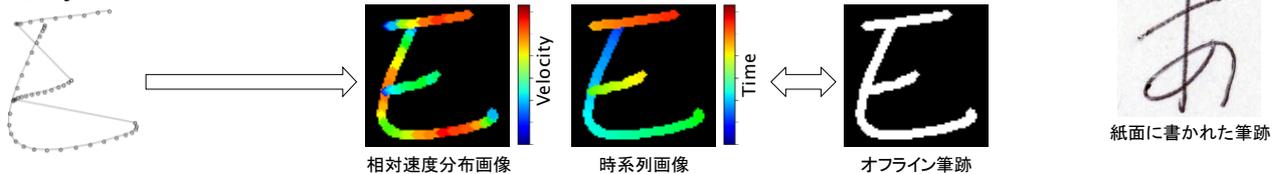
2.3 Encoder-Decoder モデルによる推定

オフライン筆跡と運動情報画像の対応関係をモデル化するために、図 1 に示す Encoder-Decoder モデルを構築した。ペン先座標のみから得られる二値化されたオフライン筆跡

(a) ひらがなデータ



(b) UJI データ



①前処理 (速度計算・位置と大きさの正規化)
ひらがなデータは空筆情報も含むため、リサンプリングしたデータを利用

②運動情報画像、オフライン筆跡の生成

図 3 運動情報画像及びオフライン筆跡生成の流れ

表 1 データ利用の内訳

対象	モデル学習用	テスト用
ひらがなデータ	13人×10字種×12回×5パターン	2人×10字種×12回×5パターン
UJIデータ	10人×62字種×2回×5パターン	1人×62字種×2回×5パターン

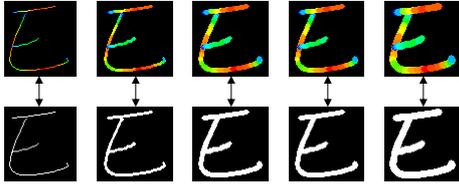


図 4 線幅変化によるデータ拡張

を入力画像として用い、各運動情報画像とモデルから推論される生成画像との誤差が小さくなるように学習を行った。3種類の運動情報画像は、いずれも同一の構造を用い、相対速度分布推定モデル、相対筆圧分布推定モデル、時系列情報推定モデルを求めた。

学習に使用したデータの内訳を表 1 に示す。文字を記載するときに使用するペンの太さや文字の大きさが変わったとしても、運動情報の分布には変化がないと考えられている[13]ことから、図 4 に示すように線の太さを 1~5pixel の 5 パターンに変化させて各画像を生成したものをモデル学習用及びテスト用データとして用意した。損失関数は、生成画像と実際の運動情報画像間の平均絶対値誤差とし、モデルの最適化には Adam を用い、ミニバッチサイズを 100 とした。UJI データについては epoch が 100 を上限、ひらがなデータについては epoch が 200 を上限として学習するように設定し、学習途中であってもモデル評価用データ（モデル学習用データの 20%）に対する損失が連続して 10 回減少しなくなったときに学習を打ち切ることとした。

3. 結果・考察

テスト用データに対して提案手法により動的特徴を推定するとともに、ETL7 手書き文字データベース[14] (ETL7

データ)に対しても提案モデルにより動的情報を推定した。また、提案手法によるオフライン筆跡からの動的特徴推定モデルの推定精度について、筆者 49 人の UJI データを用いて検証した。

3.1 テスト用データに対する推定

テスト用データに対してオフライン筆跡から動的特徴を推定した結果を図 5 に示す。オフライン筆跡から観察される形態と運動情報には関係があり、転折部では速度が遅い状態、湾曲部では曲率が大きいほど速度が速く筆圧が低い状態がみられた。これらは、人の書字運動で描かれる軌道が、躍度最小化理論やトルク変化最小化理論等の最適化原理に基づいたモデルによって近似的に生成可能である、という指摘と一致することを支持しているものといえる[15][16]。

3.2 ETL7 データに対する推定精度

ETL7 データは筆者 175 名が枠の大きさが異なる 2 枚の用紙にひらがな 46 字種及び濁点、半濁点の記号を記載したものをスキャナにより電子化し、文字単位で切出したオフライン筆跡のデータベースである。16 階調の 64×63pixels の原画像に対して、ノイズ除去、位置と大きさの正規化を行い 64×64pixels の 2 値画像として提案モデルに入力することで動的特徴を推定した(図 5-(c)参照)。

今回学習に用いたデータは筆者 13 人が 6 回記載した 10 字種のみであり、網羅的に運動を推定するためには少数のデータではあるが、基本的な筆順に準じた軌道及び運動制御理論に沿った相対筆圧推定、相対速度推定、時系列情報推定がなされているものと考えられる。

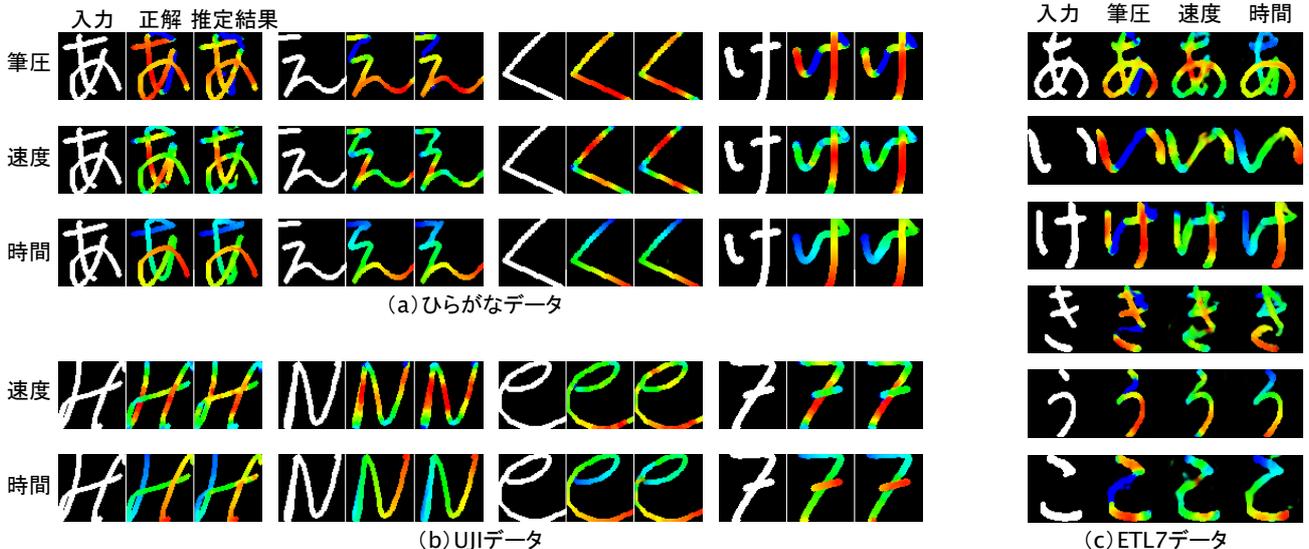


図 5 オフライン筆跡からの動的特徴の推定結果

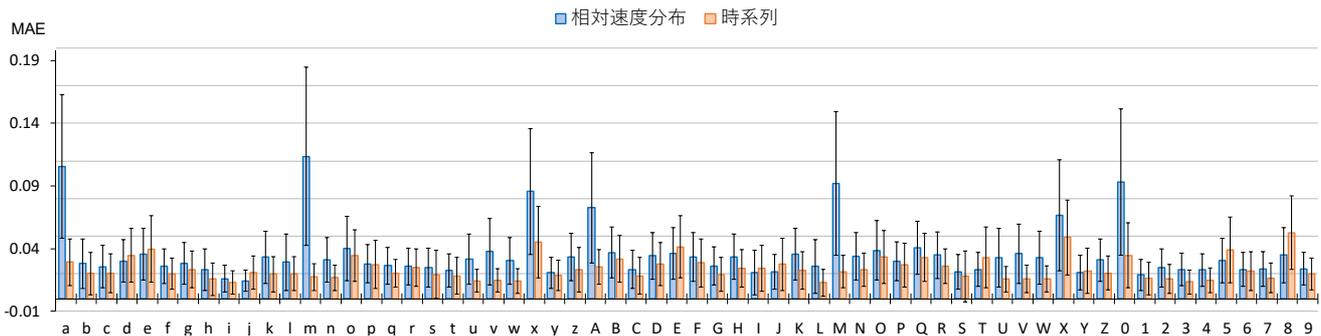


図6 推定結果と正解画像との誤差

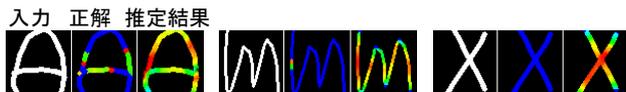


図7 相対速度分布の推定

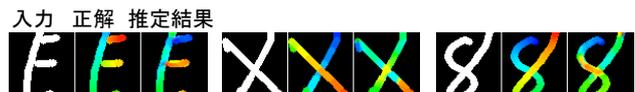


図8 時系列情報の推定

3.3 UJI データに対する推定精度

UJI データの筆者 49 名が 62 字種を 2 回記載したときのオフライン筆跡から、提案モデルにより生成される相対速度分布画像及び時系列画像と正解画像の平均絶対値誤差を各データについて求め、字種単位で平均した結果を図 6 に示す。

一部の字種 (a, m, x, A, M, X, 0 等) の相対速度分布画像で誤差が大きくなったことについては、図 7 に示すように元データの不備により相対速度分布画像が正常に作成されなかったことに起因する。しかしながら、それらのデータに対しても学習により得られた人の書字運動を反映した相対速度分布が推定されていることが確認された。時系列画像については、1 ストロークで書かれるものについては概ね正しい推定が行われた。2 ストローク以上で複数の筆順で書かれる可能性のある字種 (x, E, X 等) や、1 ストロークであっても始筆部が人によって様々であり、複数の運筆が想定される字種 (o, O, 0, 8 等) では正しく推定されない場合がみられた (図 8 参照)。

字種により推定精度に違いはみられるが、オフライン筆跡という形態情報のみからであっても、本提案モデルにより複数筆者の文字の書き方を学習することで、人の書字運動がモデル化可能であることが示された。

4. まとめ

提案する Encoder-Decoder モデルにより、オフライン筆跡からその文字を書くときの動的特徴である相対速度分布、相対筆圧分布、時系列情報が推測可能であることが示され、筆跡とその背後にある書字運動の関係性を見出した。本実験では、少数筆者のひらがなデータ及び UJI データにより運動推定モデルを学習したものであり、筆跡の多様性に対応するためには、より多くのデータを用いた実験が必要である。また、同一人が複数回、様々な条件でどのような筆跡を書くかといった個人内変動の把握が筆跡鑑定では重要であるため、各条件 (例えば、記載時間、圧の加え方、記載時の姿勢、筆記具の違い) 等をパラメータとして、それらの条件に依存しない筆者固有の書字運動特性の推定するモデルの構築を今後の課題とする。

参考文献

- [1] 黒田 正典, “書の心理—筆跡心理学の発達と課題”, 誠心書房 (1964).
- [2] 高澤 則美, “筆跡鑑定”, 科学警察研究所報告法科学編, Vol.51, No.2, pp.43-53 (1998).
- [3] S. Edelman and T. Flash, “A model of handwriting”, *Biological Cybernetics*, Vol.57, pp.25-36 (1987).
- [4] 保原 信, “手書き文字の合成モデル”, *情報処理*, Vol.11, No.10, pp.584-591 (1998).
- [5] 春日 裕之, 住田 桂一, 大川 慶, 和田 安弘, “書字運動モデルと遺伝的アルゴリズムによる変形文字の生成”, *情報処理学会論文誌*, Vol.44, No.9, pp.2363-2373 (2003).
- [6] 吉田 公一, “文書鑑定の基礎と実際”, 立花書房 (1983).
- [7] K. Franke, L. Schomaker, C. Veenhuis, L. Vuurpijl, M. van Erp, and I. Guyon, “WANDA: A common ground for forensic handwriting examination and writer identification”, *ENFHEX news – Bulletin of the European Network of Forensic Handwriting Experts*, vol.1, pp.23-47 (2004).
- [8] 押木 秀樹, 清水 陽一郎, “書字における書きやすさの重要性と書字動作に関する基礎的研究”, *書写書道教育研究*, No.21, pp.48-57 (2007).
- [9] P. Viviani and M. Cenzato, “Segmentation and coupling in complex movements”, *Journal of Experimental Psychology: Human Perception and Performance*, Vol.11, No.6, pp.828-845 (1985).
- [10] Y. Ganin, T. Kulkarni, I. Babuschkin, S. M. A. Eslami and O. Vinyals, “Synthesizing Programs for Images using Reinforced Adversarial Learning”, arXiv:1804.01118[cs,stat] (2018).
- [11] D. Llorens, F. Prat, A. Marzal, J. M. Vilar, M. J. Castro, J. C. Amengual, S. Barrachina, A. Castellanos, S. Espana, J. A. Gomez, J. Gorbe, A. Gordo, V. Palazon, G. Peris, R. Ramos-Garijo and F. Zamora, “The UJIPenchars database: A pen-based database of isolated handwritten characters”, *Proc. of the 6th International Conference on Language Resources and Evaluation* (2009).
- [12] O. Ronneberger, P. Fischer and T. Brox, “U-net: Convolutional networks for biomedical image segmentation”, *Medical Image Computing and Computer – Assisted Intervention*, 9531:pp.234-241 (2015).
- [13] 仁平 義明, “汎用運動プログラムをめぐる運動記憶研究の展開 — 書字の問題を中心に”, *心理学評論*, Vol.34, No.3, pp.358-382 (1991).
- [14] 電子技術総合研究所, Japanese Technical Committee for Optical Character Recognition, ETL 文字データベース (1973-1984)
- [15] T. Flash and N. Hogan, “The coordination of arm movements: An experimentally confirmed mathematical model”, *Journal of Neuroscience*, Vol.5, pp.1688-1703 (1985).
- [16] Y. Uno, M. Kawato and R. Suzuki, “Formation and control of optimal trajectory in human multijoint arm movement – minimum torque change model”, *Biological Cybernetics*, Vol.61, pp.89-101 (1989).