

LSTM を用いた空中一筆書き文字の筆画切り出し

LSTM-Based Stroke Segmentation of Characters Written in the Air with Single-Stroke

荒木 奏瑠[†]中井 満[†]

Kanaru Araki

Mitsuru Nakai

1. はじめに

近年の働き方改革により、リモートワークを導入する企業が増えている。その増加に伴い VR 会議に注目が集まっている。VR 会議では仮想空間上で身振り・手振り・声・文字などを用いてコミュニケーションをとる。久木元らの研究 [1] では、仮想空間において手書き注釈を多く用いることで説明時間が短くなり、効率的な意図の伝達が行えることを実証している。一方、VR 空間において一画ずつ文字を書くためにはボタン操作やピンチ動作が必要となり不自然かつ筆記に時間がかかってしまう。そこで、一筆書き文字から筆画間の不要箇所を推定することで、自然な動作で素早く文字を筆記することができると考えた。本稿では空中に一筆書きした文字の不要区間を LSTM で取り除く手法を提案する。

2. 筆画切り出しシステム

マルチタスクで学習した RNN(Recurrent Neural Network) の筆画切り出しシステムを提案する。本システムでは図 1 のように各時刻において pen up/down の筆記状態を識別することをメインタスクとし、pen down が連続する区間を筆画として切り出す。また、サブタスクとして文字の識別も同時に行う。筆跡を読みやすくする上で文字を識別することは不要であるが、文字を識別する特徴を隠れ層で抽出することによって、筆画の切り出し性能が向上すると考える。本システムでは一文字筆記することにピンチ動作を行うことを想定している。これは画間と比べて文字間で間をおくことは自然な動作と考えられるからである。したがってピンチ動作から次のピンチ動作までを一文字として、各文字の筆画切り出しを行う。

3. データ

空中の筆跡は右手人差し指先端の座標を追跡して取得する(図 2)。指先の追跡には Leap Motion コントローラを用い、120fps でサンプリングする。平仮名の多くは文字上部から書き始め、文字下部で書き終わる。そのため連続して文字を筆記しようとする、文字間に左下から右上へと移動する動きが見られる。文字間を移動する軌跡も同様に切り除くために前の文字が書き終わった点から次の文字を書き始める点までの軌跡を筆跡データとして用いる。実験では 1 人で平仮名 46 字種を 20 字ずつ筆記し、計 920 サンプルを収集した。また、データには各時刻において筆記後に文字の筆画の区間かそれ以外の区間か、視察によりラベル付けをした。

3.1 前処理

空中に文字を書くとき、位置やサイズ、筆記時間は毎回異なる。そこで位置やサイズが異なる場合の対処として文字のス

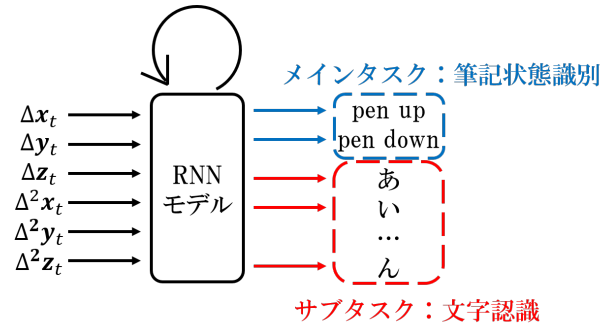


図 1: 筆画切り出しシステムの概要

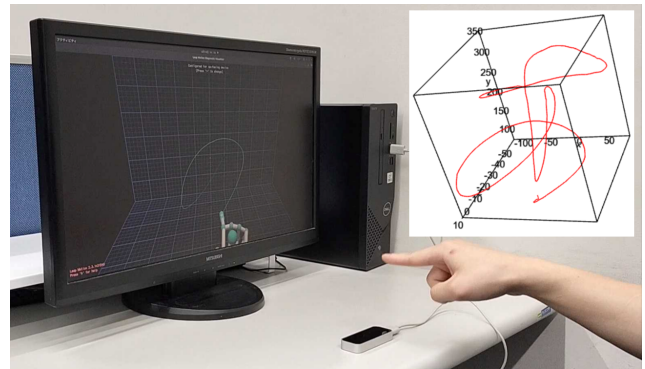


図 2: データ取得環境

ケールの正規化を行い、 x 軸、 y 軸、 z 軸上の 0 から 1 の範囲に変換する。変換後の時刻 t の筆点を $o_t=(x_t, y_t, z_t)$ とし、筆跡を $O=o_1o_2o_3\dots o_T$ とする。時刻 t と時刻 $t-1$ の筆点から $\Delta o_t=o_t-o_{t-1}$ を求め、これを移動ベクトルとする。また、移動ベクトルから $\Delta^2 o_t=\Delta o_t-\Delta o_{t-1}$ を求め、これらを時刻 t で抽出した特徴量 $f_t=(\Delta o_t, \Delta^2 o_t)$ とする。

3.2 学習

RNN を構築するにあたり、中間層には長期記憶を保持できる LSTM(Long Short Term Memory) モデルを用いる。前処理で得た特徴量 f_t を入力として与え、筆記状態識別部は毎時刻 pen up/down の 2 次元のワンホットベクトルを教師ラベルとし、文字認識部は平仮名 46 字種に対応する 46 次元のワンホットベクトルを教師ラベルとして学習を行う。

3.3 評価

筆画切り出しでは各時刻での筆記状態の識別を行う。識別された筆記状態を時系列順に並べた推定結果と、視察で付与した教師ラベルを比較し、2 つのラベルの一致率(図 3)を求める。

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

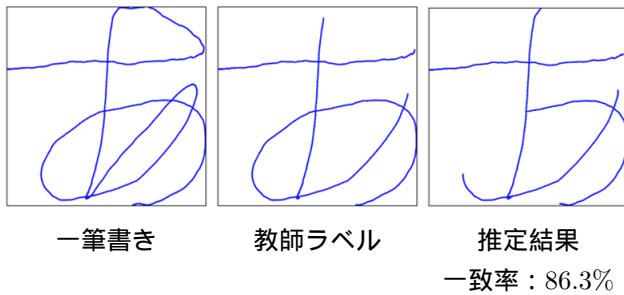


図 3: 一致率評価

表 1: LSTM のパラメータと学習条件

ノード数	128, 256*, 512
中間層	2*, 3
バッチサイズ	64*, 128, 256

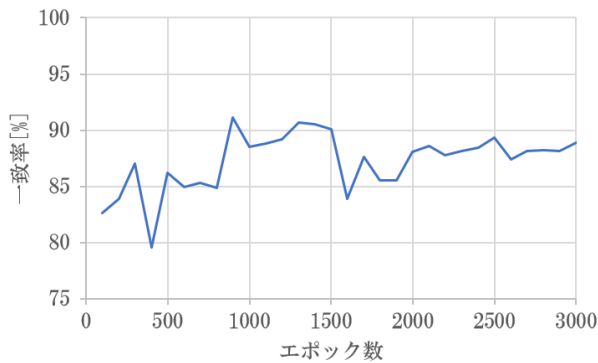


図 4: 学習回数と一致率の関係

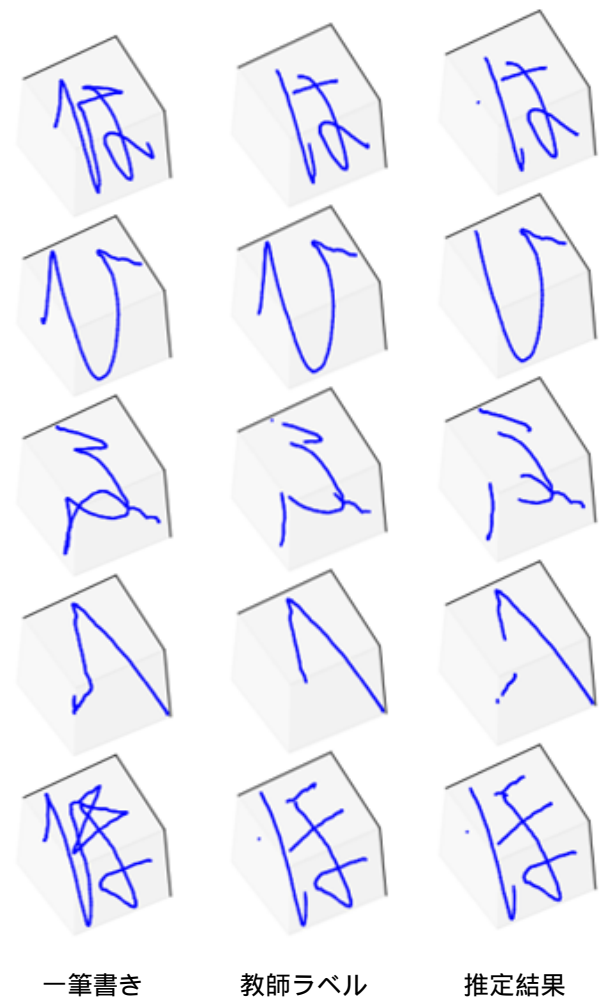


図 5: 筆画推定結果

4. 実験

LSTM の各種パラメータを変えて平仮名の学習を行った。各文字 20 サンプルの計 920 サンプルを使用した。これを各文字均等に 4 セットに分割し、3 セットを学習に、残りの 1 セットを評価に交差検証した。用いたパラメータを表 1 に示す。学習回数を 3000 回とし、100 エポックごとに学習モデルを保存した。表中の*のパラメータのときの学習回数と一致率の関係を図 4 に示す。学習回数が 900 回のときに、一致率が最も高くなり 91.1% となった。エポック数を 900 より大きくすると一致率が低下し、これ以上増やす必要はないと考える。このときの筆画切り出しを行った筆跡の一部を図 5 に示す。多くの文字に対して適切な切り出しを行うことができたと考える。ただし、「へ」のような文字では左下から右上へと上がる書き始めの筆跡が文字間を移動する軌跡であると認識されてしまい取り除かれてるという問題が見られた。この問題を解決するためには文字認識の結果を利用し、筆記開始からどの文字を書いているのかという情報を与えて処理することで「へ」や「ひ」のような文字を改善できると考える。

5. まとめと今後

LSTM を用いて空中に一筆書きした平仮名の筆画切り出しを行った。個人が筆記した文字のみを用いて学習し、一致率を求めた結果 91.1% の精度で筆画の切り出しができることが分かった。問題として文字の左下から右上へと移動する動きが文字間の不要な動きとして認識されてしまうことがあったため、文字認識結果も利用して処理を行う必要があると考える。今回の実験では個人で筆記した平仮名 46 字種各 20 サンプルとサンプル数が少なく、個人による偏りが生じて致率が高く出ていると考えられる。今後は複数人の筆記データを取得し、データを回転、ノイズを加えるなどデータ数を増やして学習、評価を行う必要があると考える。

謝辞 本研究は JSPS 科研費 21K11998 の助成を受けて行った。

参考文献

- [1] 久木元伸如, 江原康生, 古川雅人, 小山田耕二 “没入型共有 VR 空間での遠隔協調作業における手書き注釈付与を用いた思考支援に関する実験的検証,” 情処学論, Vol.48, No.6, pp.2153-2163, 2007.