

## LSTM を用いた空書き文章からの筆跡抽出

## Handwriting Trajectory Extraction from Air-Written Sentences Using LSTM

荒木 奏瑠<sup>†</sup>中井 満<sup>†</sup>

Kanaru Araki

Mitsuru Nakai

## 1. はじめに

近年の働き方として、リモートワークを導入する企業が増加しており、対面に代わる会議形式として VR 空間を活用している。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。一方で、VR 空間における文字筆記には 1 画ずつ文字を書く場合、ボタン操作やピンチ動作を毎回挟まなくてはならないという操作性の課題があり、現実世界での筆記に比べて不便である。既発表 [2] では、文字単位で一筆書きし、文字間ではペンオフ動作を入れた。本発表では文章単位で一筆書きする手法を提案する。

## 2. データ収集・作成

既発表 [2] で収集した平仮名 46 字種のデータを用いる。データには時刻  $t$  の筆点  $\mathbf{o}_t = (x_t, y_t, z_t)$ 、筆記状態、文字種情報が含まれている。平仮名 46 字種をシャッフルして並び替えを行い、文字の書き終わりと次の文字の書き始めを連結し、46 字の文字列を作成した。文字の筆跡は等時間間隔でサンプリングしており、サンプル毎に筆点数は異なる。ミニバッチ学習をするため 46 文字を連結した文字列サンプルから、図 1 のように筆点数  $L$  の筆跡を  $L/2$  点シフトしながら切り出して学習サンプルとする。すなわち、文字列を  $R = \mathbf{o}_1 \mathbf{o}_2 \mathbf{o}_3 \cdots \mathbf{o}_T$  とすると、 $s = 2$  のサンプルは  $\mathbf{o}_{L/2+1} \cdots \mathbf{o}_{3L/2}$  である。総筆点数  $T$  の文字列から  $N$  個の学習サンプルを作成する場合、 $L = 2T/(N + 1)$  である。

## 3. 筆跡抽出システム

## 3.1 前処理

空中に文字を筆記するとき、その位置は毎回異なる。そこで時刻  $t$  の筆点座標  $\mathbf{o}_t$  から速度  $\mathbf{v}_t = \mathbf{o}_t - \mathbf{o}_{t-1}$ 、加速度  $\mathbf{a}_t = \mathbf{v}_t - \mathbf{v}_{t-1}$  を求め、特徴量を  $\mathbf{f}_t = (\mathbf{v}_t, \mathbf{a}_t)$  とする。

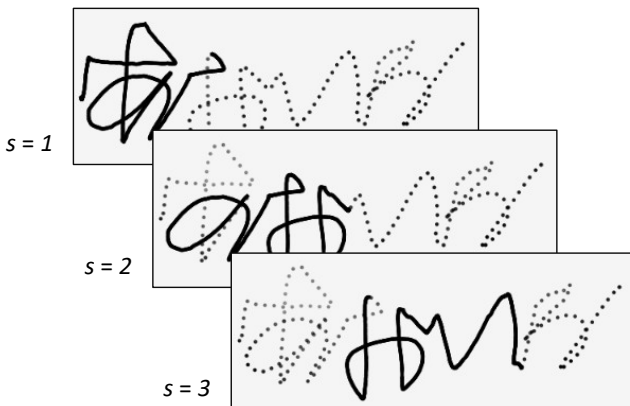


図 1: 学習サンプルの作成

## 3.2 筆記状態識別

RNN(Recurrent Neural Network) の構築には、長期依存関係の学習ができる LSTM(Long Short Term Memory) ユニットを中間層に用いる。入力に特徴量  $\mathbf{f}_t$  を与え、pen on/off の出力に softmax 関数を適応し、pen on の値が閾値を超えたときに pen on と判定し、閾値に満たない場合は pen off と判定する。連続する 2 点が pen on として識別された時、その 2 点間は筆記状態にあるとみなし、pen on が連続する区間を筆跡として抽出する。

## 3.3 学習

メインタスクである筆記状態の識別に加え、サブタスクとして文字種の識別を同時に行う。筆記状態の学習には pen on/off のワンホットベクトルを教師ラベルとし、文字の学習では文字種に対応する 46 次元のワンホットベクトルを教師ラベルとする。損失関数には交差エントロピー誤差を用い、メインタスクの損失  $\mathcal{L}_1$  とサブタスクの損失  $\mathcal{L}_2$  を組み合わせて、全体の損失  $\mathcal{L} = w\mathcal{L}_1 + (1 - w)\mathcal{L}_2$  を定義した。

## 4. 筆跡抽出実験

## 4.1 実験条件

予備実験より、LSTM の層数を 3、一層あたりのノード数を 128、バッチサイズを 46 として設定した。学習には 2 節の  $N = 300$  とし、46 文字の文字列 16 セットから、計 4800 サンプルを作成した。学習のエポック数を最大 500 とし、10 エポックごとに学習モデルを保存した。評価には 46 文字の文字列 4 セットをそのまま ( $N = 1$ ) で使用した。損失  $\mathcal{L}$  の係数  $w$  と、筆記状態の閾値を少しずつ変えた。このときの pen on と判定した線分のうち実際に pen on である線分の長さの割合を適合率、実際に pen on である線分のうち pen on と判定した線分の長さの割合を再現率とし、F 値が最大となる閾値を決定した。

## 4.2 実験結果

損失関数  $\mathcal{L}$  の係数  $w$  を 0.977、エポック数を 70 としたとき文字列で学習したモデルで最も精度が良くなった。そのときの正しくあるいは誤って判定された線分の長さの総和を表 1 に示す。適合率が 0.950、再現率が 0.978 となり、その結果 F 値が 0.964 になった。文字列で学習したモデルで「はへひほふ」を処理したときの文字認識結果を表 2 に、筆記状態の識別結果を図 2 に示す。表 2 では、筆記が進むにつれて文字認識候補が変

表 1: 混同行列

		予測		計
		pen on	pen off	
正解	pen on	95548.29	2101.30	97649.59
	pen off	5005.22	45169.61	50174.83

<sup>†</sup>富山県立大学, Toyama Prefectural University

表 2: 「はへひほふ」筆記中のサブタスクの文字認識結果の推移

時刻	305					489		647		1064								
正解	は					へ		ひ		ほ							ふ	
1 位	あ	し	は	お	ね	へ	へ	そ	も	ふ	け	ほ	そ	ほ	よ	ま	ふ	ふ
2 位		あ	ね	よ	は	れ	ふ	ふ	ひ	ひ	れ	ま	も	せ	な	ほ	な	
3 位		も	ほ	ま	ふ	ひ	ぬ	く	し	ぬ	わ	は	き	ま	ほ	ふ		
4 位			ま	ま	ま	ん	め	へ		お			ほ	ま	ま	な		
5 位			ほ	ぬ	ぬ	ね		ぬ					て	よ	は			

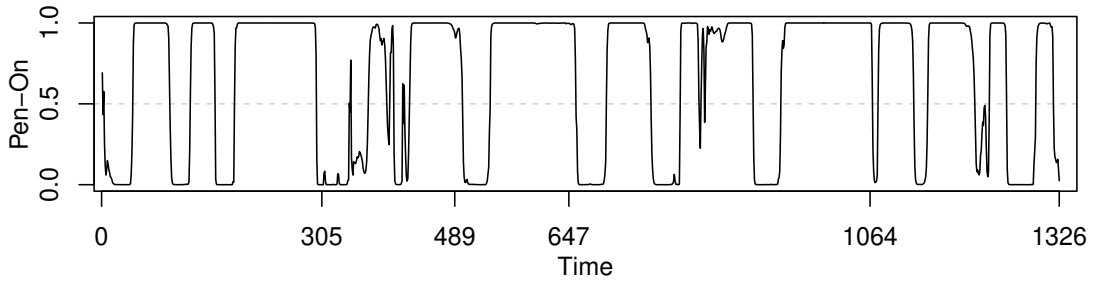
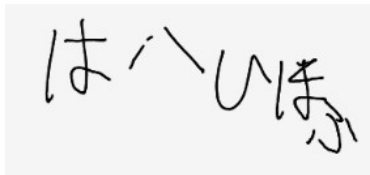


図 2: 「はへひほふ」筆記中のメインタスクの pen on 出力値の推移

文字列で学習



文字単位で学習

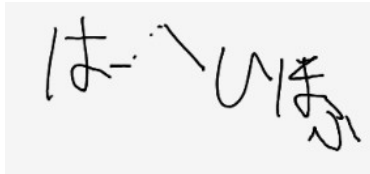


図 3: 筆跡抽出結果

わる様子を第 5 位まで示している (softmax 出力値が 0.05 に満たないものは除外している)。図 2 より、文字認識の候補に正解が含まれる区間では、筆記状態は常に 0 や 1 に近い値を取っていることが分かる。このことから、pen on を判定する閾値を変えても差は生じなかったので閾値は 0.5 で評価した。

### 4.3 考察

「はへひほふ」を筆跡抽出した結果を図 3 に示す。文字列を学習した場合の方がより正確な筆跡抽出を行えたことが分かる。文字から文字へのつながりの pen off 区間を学習することができ、精度が良くなったと考える。また、表 2 の結果と照らし合わせると、「は」は「あ、し、は、お、ね」として認識されていたが「は」の 1 画目は「し」と形状が類似し、「お」の 2 画目の前半部分、「ね」の 2 画目の後半部分は「は」の 3 画目と形状が類似する。そのため、「は」は誤認識されていたが、それでも適切な筆跡抽出を行うことができたと考える。「へ」の分析時刻をより細かく分けたときの文字認識結果を表 3 に示す。「へ」は左下から右上へと向かう動きが画間や文字間のつながりの区間として誤認識してしまい、誤って取り除いてしまったと考える。「ほ」は「は」と同様に類似した形状の文字に認識さ

表 3: 「へ」筆記中のサブタスクの文字認識結果の推移

正解	へ						
1 位	ね	は	へ	ゆ	り	へ	へ
2 位	ふ	ね	れ	り		ふ	ふ
3 位		ほ	ひ			ぬ	

れたことで多くの部分を適切に処理できたが、「ほ」の 2 画目から 3 画目へのつながりの動きが「そ」として認識されたことで正しく取り除くことができなかったと考える。

### 5. まとめ

空中に一筆書きした平仮名文字列から 0.978 の再現率と 0.950 の適合率で筆跡を抽出できた。サブタスクとして行っていた文字認識結果を確認すると、「は、ほ、け」のように形状が類似した文字と誤認識しても正しく筆跡抽出ができるが、「ほ、そ」のように形状が異なる文字と誤認識されることで筆記状態を誤認識することが分かった。文字を誤認識したときに、筆記状態が正しく識別されないことが多いため、今後はメインタスクの筆跡抽出のみでなく、サブタスクの文字認識の精度も上げることで、筆跡抽出の精度を更に向上させることが可能だと考える。

謝辞 本研究は JSPS 科研費 24K15050 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.
- [2] 荒木奏瑠, 中井満 “LSTM を用いた空中一筆書き文字の筆画切り出し,” 第 23 回情報科学技術フォーラム, 第 3 分冊, 385 - 386, 2024.