

HMMに基づくフリックキーボード入力方式

Probabilistic Flick Keyboard Based on HMMs

萩谷 俊幸†
Toshiyuki Hagiya加藤 恒夫†
Tsuneo Kato

1. はじめに

近年、タッチパネルを備えた端末の増加に伴い、ソフトキーボードでの文字入力一般的になりつつある。文字入力方式の一つとして、指をスライドさせることで文字入力を行うフリック入力普及しているが、入力誤りが多いという問題がある。これは、タッチパネルが検知するタッチ位置やスライド方向がユーザの入力意図と異なることが要因である。解決策の一つとして、ユーザの入力座標分布を考慮して、キーの検知領域を変更することが考えられる。本報告では、フリック入力での入力誤りを削減するため、ユーザの入力座標を用いた HMM に基づく入力方式について検討した結果を述べる。

2. 従来研究

ソフトキーボードの入力補正方法は、言語モデルによるものとタッチモデルによるものに分けられる。言語モデルによる入力補正方法として、Ahmet らは入力された文字列と辞書中の各単語との類似度を比較し、類似度が高い単語を候補として提示する手法を提案している [1]。Khaloudun らは、入力中の文字列と辞書中の単語との類似度を比較し、次に入力する確率の高いキーのサイズを大きくする手法を提案している [2]。しかし、言語モデルによる方法では性能が辞書に大きく依存し、固有名詞など未登録の単語は考慮されないという課題がある。

一方、タッチモデルによる補正方法として、Johan らは 10 キーでの数字入力において、各ボタンのタッチ座標の重心を用いて、キーの位置・形状を変える方法を提案している [3]。また、言語モデルとタッチモデルの両者を用いた方法としては、入力中の文字と辞書中の単語とを比較し、キー領域を変化させるといったものが Asrla ら [4] や Dmitry ら [5] により提案されている。しかし、上記のように 1 タッチが 1 文字の入力に対応する文字入力において確率モデルを適用した研究はなされているが、フリック入力において確率モデルを適用している研究は見られない。

3. 提案手法

本研究では、フリック入力において、ユーザがソフトキーボードを押してから離すまでの一連の入力座標時系列を確率モデルにより評価し、文字を決定する手法を提案する。本手法により、キー領域とフリック方向が確定的な入力方法では誤りが生じやすい、キーの境界やフリック方向判定の境界における誤りの低減が可能と考えられる。確率モデルには、文字毎に複数の状態を定義できる HMM を用い、状態毎に自己遷移を含む L-R モデルを用いる。特徴量には、一定のサンプリング周波数で取得した 2 次元入力座標時系列を用いる。HMM の各状態は対角共分散行列を持つ混合正規分布であり、モデル間の接続は、自由に連鎖するエルゴディックな接続とする。上記モデルに対し、キー領域とフリック方向が確定的なキーボードによりデータを収集し、そのデータを基に学習・評価を行うことで性能を評価する。

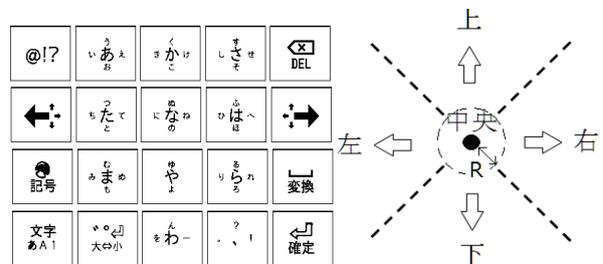


図 1 キー配置

図 2 フリック方向

4. 実験内容

4.1 データ収集

被験者には、椅子に座った状態で、片手で持ち、親指で入力してもらった。その際に、2 次元入力座標時系列および入力に対応する文字をサンプリング周波数 50Hz で取得した。データ収集に用いたキーボードは、図 1 に示すように矩形領域にキーが定義され、図 2 に示すように押下座標と離し座標の 2 点間の距離が R 未満ならば中央、R 以上なら 90° 間隔でフリック方向が定義されるフリックキーボードを用いた。単語は Juman 辞書から被験者ごとに異なる 1000 単語を抽出した。被験者は、100 単語の入力を 1 セットとし、十分に時間をおいて 10 セット行った。被験者は 25~28 歳までの男性 3 名、女性 2 名の計 5 名で、全員右手入力であった。携帯端末は Nexus S (サイズ: 4.0 inch, OS: Android4.0) を用いた。

4.2 学習および評価方法

4.1 節で得られたデータを基にモデル学習を行った。特徴量は、ユーザがソフトキーボードを押してから離すまでの一連の入力座標時系列を用いた。確率モデルには、被験者自身のデータから ML (Maximum Likelihood) 学習を行った特定ユーザモデルと、異なる被験者のデータから ML 学習を行った不特定ユーザモデル、不特定ユーザモデルを基に、被験者自身の一部のデータを用いて MLLR (Maximum Likelihood Linear Regression) により個人適応した適応ユーザモデルの 3 種を用いる。HMM の状態数は図 2 の中央に相当する文字には 2 状態、それ以外には 3 状態を割り当て、混合正規分布の混合数は 2 とした。

性能評価は、特定ユーザモデルでは、各被験者ごとに、収集データ 10 セットのうち 9 セットで学習し、残りの 1 セットを評価する CV (Cross Validation) により行った。不特定ユーザモデルと適応ユーザモデルでは、4 名のデータで学習し、残り 1 名を評価する CV により評価した。MLLR には評価用データとは異なる単語 100 個の入力データを用いる。評価には以下の式で表される誤り削減率 ERR を用いた。

$$ERR = (E_0 - E_{pm}) / E_0 \times 100 \quad (1)$$

ここで、 E_0 はキー表示領域と検知領域が等しい場合での入力誤り率、 E_{pm} は確率モデルでの認識誤り率である。

† 株式会社 KDDI 研究所 KDDI R&D Laboratories Inc.

4.3 実験結果

4.3.1 入力分布の解析

まず、キーの位置による入力座標分布の差異について述べる。全被験者の押下座標分布中心とキー中心座標の差分を示す図 3 を見ると、押下座標分布中心は x 方向は正方向に、y 方向はキーボード上部のキーでは負方向、下部では正方向に偏っている。この要因として、右手での入力であるため、ユーザが所望のキーを見易いように指を中心より右へずらしていることや、手のホームポジションからの距離が小さくなるように指を動かしていることが考えられる。

次に、全被験者、全キーにおける、フリック方向ごとの押下位置から離し位置までの平均移動量を表 1 に示す。所望の方向に対して、上方向と右方向では第一象限側に角度が傾き、下方向と左方向では第三象限側に傾く傾向が見られた。これは、片手操作では、指が円弧を描くような軌跡を描くためと考えられる。

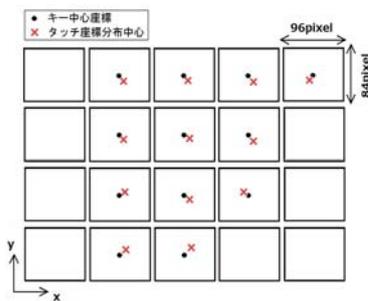


図 3 押下座標分布

表 1 フリック方向ごとの押下位置から離し位置までの平均移動量

フリック方向 (角度[°])	Δx [pixel]	Δy [pixel]	角度差分 [°]
上 (0)	25.2	53.6	+25.2
右 (90)	92.3	21.8	-13.3
下 (180)	-8.58	-108	+4.54
左 (270)	-74.7	-5.2	-3.98

4.3.2 HMM に基づく入力方式の性能評価

4.3.1 項からフリック方向ごとに、押下位置から離し位置までの平均移動量の傾向が異なることが確認できた。そこで、押下座標と離し座標の 2 点を HMM の状態に割り当てる状態モデル A と、提案手法である入力座標時系列を特徴量として用い、フリック操作で中央に相当する文字には 2 状態、それ以外には 3 状態を割り当てる状態モデル B とを特定ユーザモデルを用いて比較した結果を示す。入力精度の全被験者の平均を示す表 2 を見ると、 E_{pm} の平均は、状態モデル A で 1.25%、状態モデル B で 0.95% だった。全被験者の入力誤り率の平均 1.31% と比較すると、ERR はそれぞれ 4.58%、27.5% となり、両者とも精度向上が見られた。これは、被験者自身の入力座標を用いることで、確率モデルでの入力精度が向上することを示している。また、状態モデル A と比べて、B での精度向上が大きく、Welch の t 検定 (有意水準 5%) においても有意差が確認できた。これは、押下座標と離し座標だけでなく、フリック中の入力座標も用いることで、指の動きに即したモデルとなるためである。

表 2 状態モデルごとの入力精度の平均値

	誤り率 E_{pm} [%]	誤り削減率 ERR [%]
状態モデル A	1.25	4.58
状態モデル B	0.95	27.5
入力誤り率 E_0	1.31	-

次に、提案手法である状態モデル B での、特定ユーザモデルでの誤り傾向について述べる。表 3 に特定ユーザモデルでの推定結果の全被験者の平均値の分類表を示す。計測時に誤りだった 62 字のうち 37.1% を正解にした一方、計測時に正解だった 4775 字の 0.3% を誤認識していた。そのため、全体の ERR は 27.5% になった。特に、ユーザのフリックの初動が早く、フリック先のキー上で滞留する場合に本手法において誤認識が起こっていた。原因としては、モデル化の際、押下時、フリック時などと状態の分類をしていないため、押下時の情報に重み付いたモデルでないからと考えられる。

表 3 特定ユーザモデルによる推定結果の分類表

		推定結果 (字数)		
		正解	誤り	総数
計測結果	正解	4671	14	4685
	誤り	23	39	62

最後に、各ユーザモデルでの入力精度について述べる。

4.3.1 項での分布の傾向を基に MLLR の回帰木は 4 とした。ユーザモデルごとの誤り削減率を示す表 4 を見ると、どのモデルにおいても、ERR は正の値を示し、確率モデルを用いることの有効性が確認できる。また、不特定ユーザモデルでも 12.8% と高い誤り削減率を示した。これは、被験者に関わらず、類似したタッチ座標分布を示しているためと考えられる。特に、本研究では、全被験者が右手で入力しており、被験者の違いが及ぼす影響よりも、入力方法の共通点による影響が大きかったと考えられる。同様に、適応ユーザモデルでも 21.6% と高い誤り削減率を示し、MLLR による個人適応の有効性を確認できた。

表 4 ユーザモデルごとの誤り削減率

確率モデル	誤り削減率 ERR (%)
特定ユーザモデル	27.5
不特定ユーザモデル	12.8
適応ユーザモデル	21.6

5. おわりに

本論文では、フリック入力において、入力座標時系列を用いた HMM に基づく入力手法を提案し、評価を行った。その結果、キー領域およびフリック方向が確定的なキー入力での誤り率に対し、誤り削減率が向上した。特に、特定ユーザモデルでは、27.5% の誤り削減率を示し、本報告での提案手法の有効性が確認できた。

今後はより精度の高い判別のために特徴量の解析を行うとともに、実際の携帯端末で、本方式での入力精度および入力速度の向上を調べることで、ユーザビリティの評価を行う予定である。

参考文献

- [1] Ahmet Cunyed Tantug, "A Probabilistic Mobile Text Entry System for Agglutinative Languages", IEEE Transactions on Consumer Electronics, Vol. 56, No.2, May 2010.
- [2] Khaldoun Al Faraj et al, "BigKey: A Virtual Keyboard for Mobile Devices", HCI2009, LNCS 5612, pp.3-10, 2009.
- [3] Johan Himberg et al, "On-line Personalization of a Touch Screen Based Keyboard", IUI2003.
- [4] Asrla Gunawaradana et al, "Usability Guided Key-Target Resizing for Soft Keyboards", IUI2010
- [5] Dmitry Rudchenko, et al, "Text Text Revolution: A Game That Improves Text Entry on Mobile Touchscreen Keyboards", Pervasive 2011, LNCS 6696, pp. 206.213, 2011.