

## HMMに基づくフリックキーボード入力方式の評価

萩谷 俊幸<sup>†</sup>加藤 恒夫<sup>†</sup>株式会社 KDDI 研究所<sup>†</sup>

## 1. はじめに

近年、タッチパネルを備えた携帯端末が普及しており、指のスライド操作で文字を入力するフリック入力が一般的になりつつある。タッチパネルでの操作において、タッチの検知位置とユーザが所望した位置とが異なることが操作誤りにつながるが、フリック入力ではタッチ位置の誤りのみならず、スライド方向の誤検知が入力誤りにつながる。解決策の一つとして、ユーザがソフトキーボードをタッチ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の入力として、入力文字を判別する方法が考えられる。そこで、本報告では、入力座標時系列を用いた HMM に基づく入力方式と少量データを用いた MLLR (Maximum Likelihood Linear Regression) によるユーザ適応の検討結果について述べる。

## 2. 関連研究

ソフトキーボードに関する研究として、ユーザの入力座標を用いてキー領域を変更する方法が Johan ら [1] や Khaldoun ら [2] により提案されている。また、確率モデルを用いた方法として、タッチ分布を GMM によりモデル化し、それから得られる確率と、入力中の文字列から入力可能な辞書内の単語の言語確率を用いてキー領域を変化させる方法が Dmitry ら [3] や Tyler ら [4] により提案されている。一方、ユーザや操作方法によってタッチ分布が異なることが指摘されており、Shiri ら [5] や筆者ら [6] により、操作方法によるタッチ分布の差異の分析やそれらに対する適応方法が提案されている。しかし、上記の研究は *qwerty* やテンキーなど 1 タッチが 1 文字の入力に対応する文字入力に対して行われており、フリック入力に対する確率モデルの適応は検討されていない。

## 3. 提案手法

本研究では、フリック入力において、ユーザがタッチしてから離すまでの一連の入力座標を HMM により評価し、文字を決定する手法を提案する。HMM は連続的かつ時間的に伸縮し得る信号列のパターン認識に適しているため、タッチしてから離すまでの時間が変動し得るフリック入力には適していると考えられる。本手法では、HMM は文字毎に定義され、自己遷移を含む L-R モデルを用いる。文字はフリックが終了した時点で各 HMM の最終状態にある仮説の尤度を比較することで決定される。各状態は対角共分散行列を持つ混合正規分布を用い、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。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  $(X_t, Y_t)$  と、タッチ開始位置からの変位  $(\delta X_t, \delta Y_t)$  の計 4 次元の時系列を用いる。

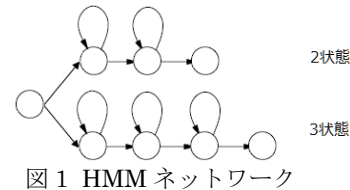


図1 HMM ネットワーク

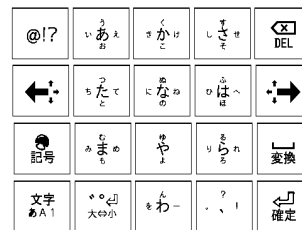


図2 キー配置

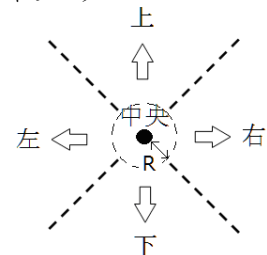


図3 フリック方向

## 4. HMMに基づくフリックキーボードの評価

本節では、本入力方式の性能をオフラインで評価する。特に、特定ユーザの入力データを大量に集めることは現実では困難なため、不特定多数のユーザの入力データから予め学習した HMM モデルを、ユーザ自身の少量のデータを用い MLLR (Maximum Likelihood Linear Regression) により適応学習することによる効果について述べる。

## 4.1 HMM モデル学習用データ収集

被験者には、姿勢および入力の統制をせず文字入力をしてもらった。その入力は、サンプリング周波数 50Hz で 2 次元入力座標の時系列データとして記録される。データ収集に用いたキーボードは、図 2 に示すように矩形領域にキーが定義され、図 3 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  $90^\circ$  間隔でフリック方向が定義される通常のフリックキーボードを用いた。入力単語は形態素解析機 Juman の辞書から被験者ごとに異なる 600 単語 (平均文字数 4.25) 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 6 セット行った。被験者は 25~35 歳までの男性 6 名、女性 3 名の計 9 名で、そのうち男性 5 名、女性 1 名は右手での入力であった。携帯端末は Nexus S を用いた。

## 4.2 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。HMM 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、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 適応した適応ユーザモデルの 3 種のモデルを用いた。MLLR 適応のためのデータ数は 0~100 語の範囲で変化させ、回帰木のクラスは 8 とした。HMM の状態数は図 3 のフリック方向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、混合正規分布の混合数は 4 とした。

性能評価は、特定ユーザモデルでは、被験者ごとに、収集データ 6 セットのうち 5 セットで学習し、残りの 1 セットを評価する CV (Cross Validation)で行い、不特定ユーザモデルと適応ユーザモデルでは、8 名のデータで学習し、残り 1 名を評価する CV で評価した。評価指標には ER (Error Rate) [7]と、以下の式で表される誤り削減率 ERR (Error Reduction Rate)を用いた。

$$ERR = (E_0 - ER) / E_0 \times 100 \quad (1)$$

ここで、 $E_0$  はキー表示領域そのままを検知領域とした場合のベースラインの誤り率である。

### 4.3 実験結果

まず、入力精度の全被験者の平均を示す表 1 を見ると、特定ユーザモデルでの ER は 2.41%，不特定ユーザモデルでの ER は 2.96%だった。全被験者のベースライン誤り率  $E_0$  の平均 3.36%と比較すると、それぞれ 28.2%，11.5%の誤り削減率を示し、両者とも精度向上が見られた。不特定ユーザモデルでも精度向上が見られる理由として、異なるユーザでも入力座標分布は類似した傾向があることと、持ち手や姿勢などによる差異に対して混合ガウス分布によりモデルが対応できていることが考えられる。

次に、図 4 に適応ユーザモデルのための適応単語数と、全被験者の ER の平均・標準偏差との関係を示す。適応単語数 0 は不特定ユーザモデルの ER を意味する。図 4 より、適応単語 5~20 語での適応において大きく ER が減少し、それより適応データが多い場合では特定ユーザモデルの ER に漸近する様子が見られる。少量のデータで ER が削減できており、ユーザのタッチ分布の適応に MLLR が有効だと言える。

表 1 確率モデルによる平均誤り率の削減(括弧内は標準偏差)

HMM モデル	誤り率 ER [%]	誤り削減率 ERR [%]
特定ユーザ	2.41 (0.46)	28.2 (4.92)
不特定ユーザ	2.96 (0.41)	11.5 (8.41)
入力誤り率 $E_0$	3.36 (0.53)	-

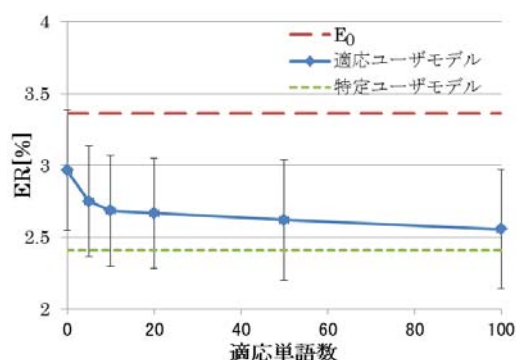


図 4 適応データ数による ER の変化

## 5. HMM フリックキーボードの端末上での評価

本節では、携帯端末上に実装した HMM フリックキーボードで精度と入力効率の評価を行う。

### 5.1 実験方法

被験者には、60 単語の入力を 1 セットとし、4.1 節でデータ収集に用いた通常のフリックキーボード A と HMM を用いたフリックキーボード B により各々 6 セット入力してもらい、比較評価をする。評価は、入力精度

の評価尺度である ER と、入力速度の評価尺度である WPM (Word Per Minute) [7] を指標として用いた。キーボードの提示順序は全 12 セットの中からランダムに決定した。HMM モデルには、4.3 節で得られた不特定ユーザモデルを基に、各セットの冒頭 10 単語の入力のうち正しく入力された文字に対する入力データを用いて携帯端末上で MLLR 適応した適応ユーザモデルを用いた。

評価対象は各セットの冒頭の 10 単語を除く 50 単語 6 セットの入力を対象とした。被験者は前章の被験者以外の 24~57 歳までの男性 6 名、女性 2 名の計 8 名で、そのうち男性 4 名、女性 1 名は右手での入力であった。

### 5.2 実験結果

表 2 に入力方法ごとの ER と WPM の平均を示す。従来のキーボード A では、ER 3.78%，WPM 17.5 にに対し、HMM によるキーボード B では、ER 3.16%，WPM 19.2 を示した。相対的に、キーボード A と比較して B の方が ER を 16.5%削減、WPM を 10.5%増加しており、B による性能向上を確認できた。WPM の向上の理由としては、入力誤りの削減により、入力の訂正にかかる時間を削減できたためと考えられる。

両者の評価尺度とも、Welch の T 検定 (有意水準 5%) においても有意差が確認でき、実際の携帯端末上での入力においても、提案手法が有効であると言える。

表 2 入力方法ごとの ER と WPM の平均 (標準偏差)

入力方法	ER [%]	WPM
キーボード A	3.78 (0.94)	17.5 (4.83)
キーボード B	3.16 (0.87)	19.2 (4.99)

## 6. おわりに

本報告では、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固定の従来方式に対し、誤りを削減できた。特に、特定ユーザモデルでは、28.2%の誤り削減率を示し、適応ユーザモデルでは、少量のデータでも MLLR の効果があることが確認できた。また、携帯端末上での操作においても、入力精度と速度にそれぞれ 16.5%，10.5%の向上が見られ、提案手法の有効性を確認することができた。

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、端末形状による差異を検討する予定である。

### 参考文献

- [1] Johan Himberg et al, "On-line Personalization of a Touch Screen Based Keyboard", IUI2003.
- [2] Khaldoun Al Faraj et al, "BigKey:A Virtual Keyboard for Mobile Devices", HCI2009, LNCS 5612, pp.3-10, 2009.
- [3] Dmitry Rudchenko et al "Text Text Revolution: A Game That Improves Text Entry on Mobile Touchscreen Keyboards", Pervasive 2011.
- [4] Tyler Baldwin et al, "Towards Online Adaptation and Personalization of Key-Target Resizing for Mobile Devices", IUI 2012.
- [5] Shiru Azenkot et al, "Touch Behavior with Different Posture on Soft Smartphone Keyboard", MobileHCI 2012.
- [6] Toshiyuki Hagiya et al, "Probabilistic Keyboard Adaptable to User and Operating Style Based on Syllable HMMs", ICPR 2012.
- [7] Ahmed Sabbir Arif et al, "Analysis of Text Entry Performance Metrics", TIC-STH 2009.