

注視点情報を用いた確率的フリック入力方式の提案

Probabilistic Flick Keyboard Incorporating Gaze Point Information

萩谷 俊幸†
Toshiyuki Hagiya加藤 恒夫†
Tsuneo Kato

1. はじめに

近年、多くのモバイル機器でソフトウェアキーボードを用いた文字入力が一般的となっている。文字入力の際、入力済みの文字列や候補単語を確認するなどキーボードと異なる領域を注視することは少なくないが、それが入力誤りの原因となりうる。解決策の一つとして、タッチ位置やフリック角度の誤差範囲の期待値に応じて候補単語の探索空間の大きさを調整することが考えられる。本報告では、文字入力効率の向上を目的に、タッチ目標点と注視点間の距離と、タッチ目標点に対するタッチ位置およびフリック角度の関係をモデル化し、同モデルを確率モデルに基づくキーボードに組み込むことで、候補単語の探索に注視点情報を反映する方法を提案する。

2. 従来研究

ソフトウェアキーボードに確率モデルを適用した方法として、GMM や HMM でモデル化されたタッチモデルから得られる確率と、入力中の文字列から選択可能な辞書内の単語の言語確率との積を最大化する文字を提示する方法が提案されている [1, 2]。また、ユーザや操作方法等のコンテキストによってタッチ分布が異なるため、それぞれに対しタッチ分布を適応させる方法が提案されている [3, 4]。一方、注視点とユーザ操作を組み合わせることにより操作支援をする方法として、Zhai らにより、マウスを用いオブジェクトを選択する際、注視点情報を用いて精度を高める MAGIC Pointing が提案されている [5]。また、タッチ操作において、Stellmach らにより、複数のオブジェクトが重なっている際、注視点に最も近いオブジェクトを選択する方法が提案されている [6]。しかしながら、視線とタッチ位置やフリック操作の関係を文字入力に用いた研究は報告されていない。

3. 提案手法

本研究の特徴は二点ある。一点は、注視点とタッチ位置間の距離と、タッチ位置およびフリック角度の誤差分布との関係を実験的に調べ、注視点-タッチモデルとしてモデル化すること。もう一点は、同モデルを確率モデルに基づくフリック入力に組み込み、算出される文字の生起確率を用いて、候補単語の探索空間を動的に変化させることである。図 1 のブロック図に沿って説明すると、まず、注視点-タッチモデルは、タッチ座標 (x_t, y_t) と注視点座標 (x_g, y_g) の入力に対して、タッチ誤差とフリック角度誤差の分散を出力する。一方、文字毎の生起確率はタッチモデルと言語モデルの確率の積で算出される。タッチモデルはタッチ座標 (x_t, y_t)

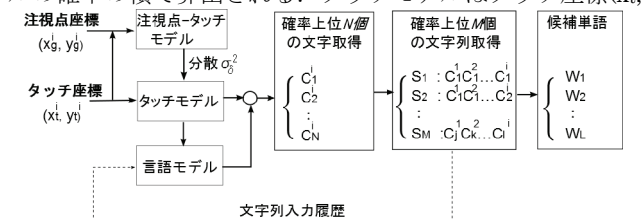


図 1 提案手法のブロック図

$y_t)$ を特徴量とする GMM と、タッチ座標と離れた座標が成す角と距離 (θ, d) を特徴量とする GMM とを連結した 2 状態の状態遷移モデルで定義され、言語モデルは、仮名 3-gram を用いる。ここで、注視点の影響を考慮するため、各状態の GMM の分散を注視点-タッチモデルからの分散に置き換える。上記手法により、入力に対し、注視点-タッチモデルで補正したタッチモデルと言語モデルの積の上位 N 個の文字が出力される。次に、入力文字列において、起こりうる組み合わせのうち確率が上位 M 個の文字列を取得し、最後に、それらの文字列に対し候補単語の検索を行い、前方一致により補完された単語候補がユーザに提示される。

4. 注視点-タッチモデルの構築

4.1 データ収集

被験者は、図 2 に示す計測アプリケーションにおいて、注視目標点を注視しながら、タッチ目標点を基準にして、上下左右にフリック操作をした。1 操作毎に、注視目標点は固定のままタッチ目標点とフリック方向を変更するという試行 8 回を 1 セットとし、注視目標点を変えて 30 セット行った。注視目標点は、半径が 40pixel (約 3mm) の黒の丸で、タッチ目標点は、一辺 60pixel の赤の正方形で描画され、フリック方向は注視目標点上に表示される矢印で指示された。端末は、各目標点とタッチ座標 (x_t, y_t) 、タッチ座標と離れた座標が成す角と距離 (θ, d) を記録した。被験者は 27~60 歳の男性 5 名、女性 3 名の計 8 名で、全員右利きであり、左手で端末を保持し右手で操作をした。端末は、Nexus 7 (Android 4.0, 7.0inch, 解像度 1920 × 1200pixel) を用い、入力には、12 キー配列のフリック入力を用いた。

4.2 注視点-タッチモデルの決定

まず、図 3 に示すように、注視目標点とタッチ目標点の距離 L に対しフレームサイズ 300pixel とフレームシフト長 200pixel のフレーム内のタッチ誤差 δ_x, δ_y と角度誤差 δ_θ について、分散 $\sigma_{\delta_x}^2, \sigma_{\delta_y}^2, \sigma_{\delta_\theta}^2$ を求める。次に、図 4 に示すような L と分散 $\sigma_{\delta_\theta}^2$ の関係に対して、線形回帰、指数回帰、ロジスティック回帰の 3 種の回帰関数を当てはめることで注視点-タッチモデルのモデル化を行い、それぞれの決定係数 R^2 を比較することによりモデルの

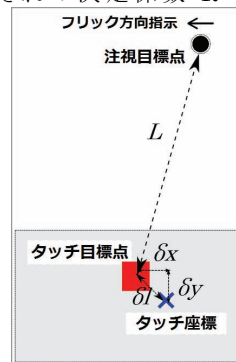
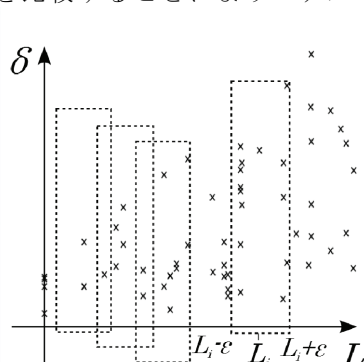
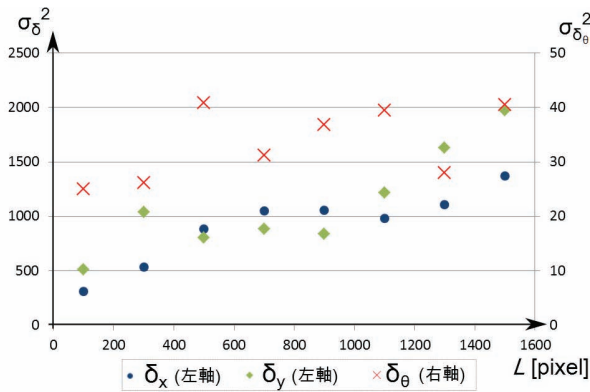


図 2 計測画面

図 3 注視目標点とタッチ目標点距離 L とタッチ誤差 δ の関係例

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

図4 Lとタッチ誤差と分散 σ_{δ}^2 の関係図表1 回帰関数毎の決定係数 R^2

	線形	指数	ロジスティック
δ_x	0.84	0.89	0.89
δ_y	0.77	0.77	0.75
δ_{θ}	0.26	0.38	0.28
δ_d	0.10	0.08	0.06

当てはまりの良さを比較した。Lと σ_{δ}^2 の関係を示す図4を見ると、Lが大きくなるにつれてタッチ誤差に関する分散は大きくなる傾向が確認できる一方、フリック角度に関する分散には強い相関が確認できなかった。回帰関数毎の決定係数を示す表1を見ると、どの回帰関数においてもLとタッチ誤差に関する分散には強い相関が見られたものの、フリック角度の分散には弱相関しか見られなかった。また、フリック距離dについても同様に調べたが、ほとんど相関が見られなかった。これらの結果より、注視点とタッチ位置の距離は、タッチのばらつきへの寄与は大きいですが、フリック操作への寄与は小さいことを示している。

一方、回帰関数の影響に着目すると、決定係数に多少の差異はあるもののANOVA検定($\alpha=0.05$)では有意差が見られなかった。それゆえ、最大の値を示した指数回帰でそれぞれをモデル化する。

5. 注視点情報を用いたフリック入力方式の評価

5.1 モデル学習

確率モデルに基づくフリック入力での言語モデルには、京都大学テキストコーパス Ver 4.0 から最尤法により学習された仮名 3-gram を用いる。一方、タッチモデルには、今回の被験者と異なる6名(30~42歳、男性4名、女性2名)により入力された合計360文の入力データから最尤法により予め学習されたものを用いる。タッチモデルのGMMの分散は、1状態目では、注視点-タッチモデルから推定される分散 $\sigma_{\delta_x}^2, \sigma_{\delta_y}^2$ に、2状態目では、 $\sigma_{\delta_{\theta}}^2$ に置き換えた。3節で述べた入力に対して出力される文字候補の数Nは最大の確率を示す文字との確率との比が0.5以上を示す文字数とし、文字列取得の最大値Mは6とした。

5.2 実験・評価方法

被験者には、独自に収集したEメールコーパス20文の入力を1セットとし、通常のフリック入力(B1)、従来の確率モデルを用いたフリック入力(B2)と提案手法の3つのキーボードにより各々3セット入力してもらい、比較評価を行った。評価指標として、入力精度はTER (Total Error Rate)を、入力速度はWPM (Words Per Minute)を用いた。キーボードの提示順序は全9セットの中からランダムに決定した。被験者は座ったまま操作し、提示された文と等し

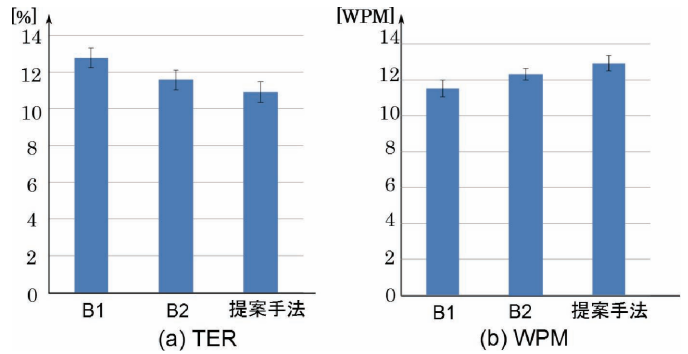


図5 入力方法毎のTERとWPMの平均(標準誤差)

くなるよう入力した。その際、端末は (x_t, y_t) と (θ, d) 、 (x_g, y_g) を記録した。視線認識器はSnapDragon SDKを基に作成し、被験者毎に各セットの施行前にキャリブレーションを行った。評価対象は各セットの冒頭の1フレーズを除く19フレーズ3セットの入力を対象とした。被験者は26~41歳までの日本人男性4名、女性2名の計6名で、全員、右手での入力であった。

5.3 実験結果

図5(a)(b)に、入力方法毎のTERとWPMの平均を示す。提案手法はB1、B2と比較してTERをそれぞれ14.7%、5.8%削減し、WPMを12.0%、5.0%向上させた。その理由として、提案手法の方が、不正確な入力に対しても候補単語を適切に提示することができ、訂正のためのキー入力数と入力誤りの低減を行うことができたためと考えられる。

上記のように、両提案手法と比べて速度と精度の向上が見られた一方で、検定により評価をしたところ、ANOVA検定($\alpha=0.05$)では有意差があることが確認できたものの、Tukeyの多重検定($\alpha=0.05$)では、いずれもB2と提案手法の間に有意差が確認できなかった。この理由として、被験者数が不十分であったことが考えられる。

6. おわりに

本報告では、注視点情報を確率モデルに基づくフリック入力に用いるために、まず、タッチ目標点と注視点間の距離と、タッチ誤差およびフリック角度誤差との関係をモデル化した。次に、同モデルから算出されるタッチ誤差およびフリック角度誤差の分散を確率モデルに基づくフリック入力に反映することで候補単語の探索空間を動的に変更する手法を提案し、評価を行った。その結果、提案手法は、通常のフリック入力、注視点を用いない従来の確率モデルに基づくフリック入力と比較して、入力誤りは14.7%、5.8%削減し、速度は12.0%、5.0%向上し、提案手法の有効性を確認することができた。

参考文献

- [1] Tyler Baldwin et al, "Towards Online Adaptation and Personalization of Key-Target Resizing for Mobile Devices", IUI 2012.
- [2] Toshiyuki Hagiya et al, "Adaptable probabilistic flick keyboard based on HMMs", IUI 2013.
- [3] Shiri Azenkot et al, "Touch Behavior with Different Posture on Soft Smartphone Keyboard", MobileHCI 2012.
- [4] Ying Yin et al. "Making Touchscreen Keyboards Adaptive to Keys, Hand Postures, and Individuals - A Hierarchical Spatial Backoff Model Approach", CHI 2013
- [5] Shumin Zhai et al. "Manual and gaze input cascaded (MAGIC) pointing", CHI 1999.
- [6] Sophie Stellmach et al. "Looking: Investigating Seamless Gaze-supported Selection, Positioning, and Manipulation of Distant Targets", CHI 2013.