

注視点情報を用いた確率的ソフトウェアキーボード入力方式

萩谷 俊幸[†]

加藤 恒夫[†]

株式会社 KDDI 研究所[†]

一々に提示される。

1. はじめに

近年、スマートフォンやタブレットのみならず多くの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、入力した文字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しながら入力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実際のタッチ位置の分散との関係をモデル化し、その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. 関連研究

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 (Gaussian Mixture Model) により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入力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ことが指摘されており、コンテキストに応じて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、マウスを用いた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、Sophie 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、視線とタッチの分散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、タッチ位置に対する誤差分布の分散との関係を注視点-タッチモデルとしてモデル化すること。もう一点は、同モデルを確率的なキーボードに組み込み、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 に示すブロック図に沿って説明すると、まず、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、注視点-タッチモデルからタッチの分散が出力される。タッチモデルは文字毎に GMM として定義されるが、注視点の影響を考慮するため、GMM の分散を注視点-タッチモデルの分散に置き換える。ある文字の生起確率は、タッチモデルからの確率と文字 n-gram 言語モデルからの積で算出され、上位 N 個の文字が取得される。ここで N は可変で、最大の確率を示す文字との確率との比がある閾値 ϵ 以内を示す語の数である。次に、入力列において、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

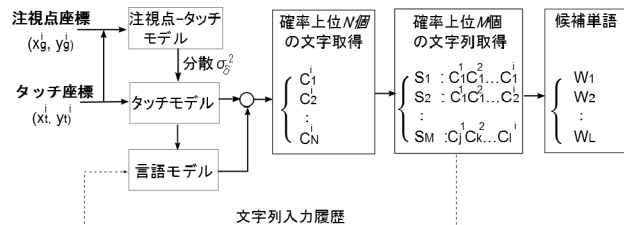


図 1 提案手法のブロック図

4. 注視点-タッチモデルの構築

本節では注視点とタッチ位置間の距離と、タッチ位置に対する誤差分布の分散との関係のモデル化について述べる。

4.1 データ収集

被験者には、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、タッチ目標点をタッチしてもらった実験を行った。1 タッチ毎に、注視目標点は固定のままタッチ目標点のみ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 10 セット行った。タッチ目標点と注視目標点は、それぞれ半径が 20pixel (約 4mm) の赤と黒の丸で表示された。端末は、各目標点と、タッチ座標 (x_t, y_t) を記録した。被験者は 25~59 歳の男性 10 名、女性 6 名の計 16 名で、そのうち男性 8 名、女性 5 名は右利きだった。被験者は両手で端末を保持し利き手で操作をした。端末は、AQUOS PAD SHT21 (Android4.0, 7.0inch, 解像度 1280×800pixel) を用い、キーボードの配列は QWERTY 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 200pixel とフレームシフト長 100pixel 内のタッチ誤差 δ_1 について分散 σ_δ^2 を求める。次に、図 4 に示すような L と分散 σ_δ^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 種の回帰式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当てはまりの良さを比較した。

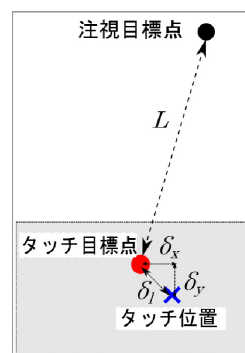


図 2 計測画面

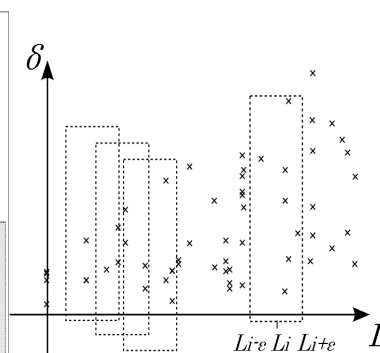


図 3 注視点とタッチ目標点の距離 L とタッチ誤差 δ の関係例

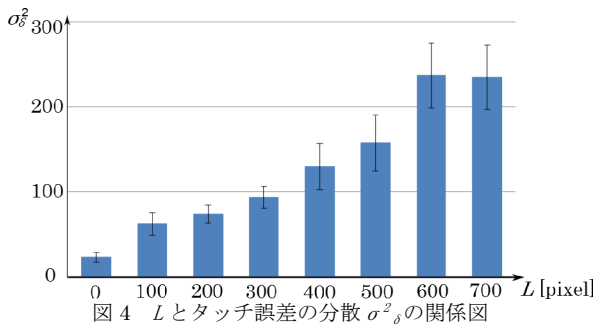


図4 Lとタッチ誤差の分散 σ^2_d の関係図

表1 回帰関数毎の決定係数

回帰関数	線形	指数	ロジスティック
R^2	0.79	0.77	0.81

L と σ^2_d の関係を示す図 4 を見ると、L が大きくなるにつれて σ^2_d も大きくなる傾向が確認できる。次に、回帰式毎の決定係数を示す表 1 を見ると、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、以降の実験では最も単純なモデルである、線形回帰 ($\sigma^2_d = aL + b$ ($a=0.32, b=14.7$)) を用いる。

5. 注視点情報を用いた確率的ソフトウェアキーボード入力方式の評価

本節では、注視点-タッチモデルを用いた確率的ソフトウェアキーボード入力方式の性能評価について述べる。

5.1 モデル学習

確率的キーボードで用いられる文字 3-gram で定義された言語モデルは、Brown Corpus [8] から最尤法により学習された。一方、GMM で定義されたタッチモデルは、端末は 4 節と同様のものを用い予め収集した、今回の被験者と異なる 7 名 (23~36 歳, 男性 5 名, 女性 2 名) により入力された合計 840 フレーズの英文入力データから最尤法により学習された。提案手法での GMM の分散はタッチ座標 (x_t, y_t) と注視点座標 (x_g, y_g) の入力に対し、注視点-タッチモデルから推定される分散に置き換えた。3 節で述べた入力に対する確率上位 N 個の数を定めるための閾値 ϵ は 0.5, 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、英文 20 フレーズの入力を 1 セットとし、通常の QWERTY キーボード (B1)、従来の確率的キーボード (B2) と提案手法の 3 つのキーボードにより各々 6 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 Total-ER (Error Rate) を、入力速度は WPM (Words Per Minute) [7] を用いた。キーボードの提示順序は全 18 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入力し、誤入力文字の訂正および候補単語の選択により提示されたフレーズと等しくなるよう入力した。その際、端末は (x_t, y_t) と (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 6 セットの入力を対象とした。被験者は 23~40 歳までの日本人男性 7 名, 女性 4 名の計 11 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。

評価の際、被験者により候補単語利用率に差があり、利用率に応じて効果が異なると考えられるので、全被験者と、利用率が上位 50% の被験者の二群に対し評価を行った。

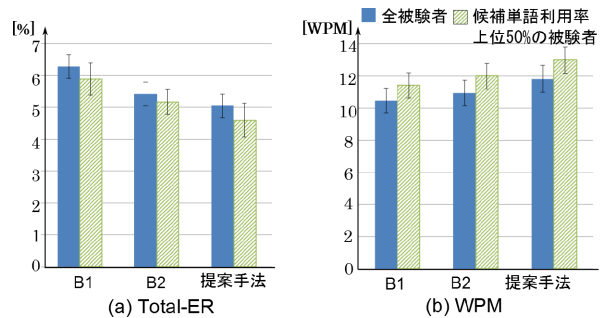


図5 入力方法毎の Total-ER と WPM の平均 (標準誤差)

5.3 実験結果

図 5 (a) (b) に、入力方法と被験者群毎の Total-ER と WPM の平均を示す。全被験者の値に注目すると、提案手法は B1, B2 と比較して Total-ER を 19.7%, 7.0% 削減し、WPM を 12.7%, 8.0% 向上させた。しかし、ANOVA 検定 ($\alpha=0.05$) で 3 者のいずれかに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。一方、候補単語利用率上位 50% の被験者に注目すると、提案手法は B1, B2 と比較して Total-ER を 21.9%, 11.2% 削減し、WPM を 13.9%, 9.2% 向上させた。これらの値は全被験者の示す値より大きな改善を示し、ANOVA 検定、Tukey の多重検定ともに有意差を示した。提案手法では、被験者が不正確な入力をした場合にも、候補単語を選択したことで、訂正にかかるキー入力数と入力誤りが低減していることから、提案手法は、候補単語を選択する割合が多い被験者により有効であることが確認された。

6. おわりに

本報告では、注視点情報を確率モデルによるソフトウェアキーボードに用いるために、まず、タッチ目標点と注視点間の距離と、タッチ目標点に対する実際のタッチ位置の分散との関係をモデル化し、次に、同モデルから算出されるタッチの分散を確率モデルを用いたキーボード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、通常のキーボード、注視点を使わない確率的キーボードと比較して、入力誤りは 19.7%, 7.0% 削減し、速度は 12.7%, 8.0% 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] Dmitry Rudchenko et al "Text Text Revolution: A Game That Improves Text Entry on Mobile Touchscreen Keyboards", Pervasive 2011.
- [2] Tyler Baldwin et al, "Towards Online Adaptation and Personalization of Key-Target Resizing for Mobile Devices", IUI 2012.
- [3] Shiri Azenkot et al, "Touch Behavior with Different Posture on Soft Smartphone Keyboard", MobileHCI 2012.
- [4] Toshiyuki Hagiya et al, "Probabilistic Keyboard Adaptable to User and Operating Style Based on Syllable HMMs", ICPR 2012.
- [5] Zhai, S., Morimoto, C., and Ihde, S. Manual and gaze input cascaded (MAGIC) pointing. In Proc. CHI 1999, 246-253.
- [6] Sophie, S., and Raimund, D. Still Looking: Investigating Seamless Gaze-supported Selection, Positioning, and Manipulation of Distant Targets. In Proc. CHI 2013, 285-294.
- [7] Kucera, H., and Francis, W. Computational Analysis of Present-Day American English.
- [8] Ahmed Sabbir Arif et al, "Analysis of Text Entry Performance Metrics", TIC-STH 2009.