

システム特性と人的要因を考慮した ポインティングタスクのパフォーマンスモデル

任 向実[†], 姜 興起^{††}

[†]高知工科大学, ^{††}旭川大学

ポインティングデバイスのパフォーマンス評価は HCI(Human Computer Interaction)の分野における重要な課題であり、このために Fitts の法則が有力なツールとしてしばしば使用されてきた。Fitts のモデルは、一方向のポインティングタスクにおけるパフォーマンス指標の基礎となるものとして ISO でも採用され、広く使われてきたが、その妥当性に対する疑問も指摘されており、HCI におけるタスク設計の立場から見ても Fitts のモデルに対する改良の必要性が迫っているといえる。そこで本論文はポインティングタスクの効率性を評価するための新しいモデルを提案し、このモデルの妥当性と有効性を示すために、赤池情報量規準 AIC を評価基準として用いて伝統的 Fitts モデルとの比較を行った。その結果、予測力において本論文で提案したモデル(SH モデル)が伝統的モデルより優れていることが検証された。

A Systematic Human-factor Model for Pointing Task Evaluation

Xiangshi REN[†], Xing-Qi JIANG^{††}

[†]Kochi University of Technology, ^{††}Asahikawa University

One challenging job for those who are doing research on Fitts' law, an evaluation tool used in human computer interaction, is to resolve the problem of the input hits' distribution. Although one method has been proposed and used for many years as the ISO standard for one-direction pointing task, doubts now exist regarding its validity. Design tasks in HCI demand a revision or a replacement of the model offered in Fitts' law. In response to this need, we developed a new model for evaluating input efficiency of pointing tasks. The prediction effect of the new and the traditional models are compared using AIC (Akaike Information Criterion), a criterion for statistical model selection. The results of AIC evaluation show that the new model, developed by us and described in this paper, is better than the traditional ones in performance prediction and evaluation.

1. はじめに

コンピュータの普及とともにポインティングデバイスが多く使用されるようになった。しかしながら、現在使用されている入力装置の多くが十分な快適性をもっているとはいえず、これらのデバイスの性能と効率性の評価・予測のためのモデル開発が重要な課題である。入力装置およびインタフェースの選択と設計に先んじてパフォーマンスの評価と予測を行

う必要があり、そのためのモデルが欠かせない。最初のモデルは有名な Fitts の法則として 1954 年に Fitts によって提案されている[5]。近年、Fitts モデルの改良に関する試みが盛んに行われている[1,10]。例えば、Accot らは、汎用性のあるマウスはある種の特異なインタフェースに対しては必ずしも最適ではないと指摘している[1]。

Fitts の法則についての議論はモデルの定式化に集中しているといえる。Fitts のモデルはパフォーマンス

ス指標(IP :index of performance)と呼ばれる概念で導入され、その基礎は Shannon の通信容量(channel capacity)の概念である[14]。パフォーマンス指標 IP は次式で定義されている[9]。

$$IP = ID / MT \quad (1)$$

ここで、 MT はポインティングデバイスを使う操作時間で、 ID は困難度(index of difficulty)と呼ばれる指標である。Fitts のモデルでは困難度が

$$ID = \log_2(2A/W) \quad (2)$$

で与えられている。ただし、 W はターゲット幅、 A は二つのターゲット間の距離を表す。Fitts モデルの一般形は(1)式より次のように表されている。

$$MT = a + bID \quad (3)$$

(3)式では、 a と b はパラメータで、 MT は操作時間の理論値(または予測値)を表す。

しかし、(3)式に次のような欠点がある。つまり、 A に比べて W がより大きい場合、 ID の値が負になってしまう。これは論理的に考えれば明らかに不合理なことである。そのために、Fitts の法則に対する修正が行われた。その一つは Shannon の定理 17¹との類似性から次のように提案されたものである[8]。

$$MT = a + b \log_2\left(\frac{A}{W} + 1\right) \quad (4)$$

ここで困難度は

$$ID = \log_2\left(\frac{A}{W} + 1\right) \quad (5)$$

となる。(4)式と(5)式において、ターゲット間の距離 A は「信号」に対応し、ターゲット幅 W は「雑音」に対応している。これについて疑問点がまだ残されている。Shannon の理論は、信号は加法性の白色正規雑音(AWGN: Additive White Gaussian Noise)によって攪乱されているという仮説に基づいて厳密な数学的演繹推理の上に成り立つものである[14]。したがって、この理論をポインティングデバイスへ拡張するとき、ポイントの座標分布が正規分布に従うという条件(空間的制約)が必要である。文献[9,10]によれば、被験者に「できるだけ迅速でかつできるだけ正確にターゲットをポイントするように」という実験条件で行い、96%の成功率という空間的制約が達成できると考えられる。しかし、このような実験条件にしたがっても96%の成功率でターゲットをポイントすること、つまりエラー率を4%にコントロールすること

は確実ではない。そこで、多くの研究は如何に4%エラー率実験の条件を維持しながら Fitts の法則を改良できるかという点に集中している。その試みの一つとして有効ターゲット幅 We (effective target width)²の指標が提案されており、しかも ISO 規格 9241-9 に採用されている[3]。しかし、エラー率4%に修正してもポイントの座標分布は必ずしも正規分布に近づくとは限らない。また、ISO 規格 9241-9 の提案に対しても異論が見られる[16]。

要約すれば、Fitts の法則を活かす方策として、エラー率を、タスクに対する制約条件(add hoc)または実験後のデータ処理(post hoc)により、4%にコントロールする方法を取っているのが、現状である。これに対してもう一つの考え方は空間的制約に制限されない新しいモデルを構築し、Fitts のモデルに代わって用いることである。

2. 新しいモデルの提案

まず、本論文の基本的な考え方は情報理論と現代情報量統計学によるもので、従来の Fitts のモデルの基礎である Shannon の通信容量理論と異なる。

図 1 に示すように、もしポインティングを範囲 $A+W$ 以内(グレイの領域)に限定すればシステムの構造によるヒットの確率 P_s は次式で与えられる。

$$P_s = \frac{W}{A+W} \quad (6)$$

ただし、 A は二つターゲット間の距離を、 W はターゲット幅を表す。また、破線で示すような操作領域になると、(6)式を次のように書き換える。

$$P_s = \frac{W}{A+(I+1)W} \quad (7)$$

ここで、実際の入力領域は理論の入力領域より IW だけ広くなると仮定している。そこでシステムの構造に関する自己情報量(self-information)は次式で与えられる。

$$ID_s = \log_2 \frac{1}{P_s} = \log_2 \left(\frac{A}{W} + I + 1 \right) \quad (8)$$

¹ Shannon の定理 17 では通信路容量 C は $C = B \log_2(S/N+1)$ で与えられる。ここで、 B は雑音の帯域幅で、 S/N は信号・雑音比である[14]。

² 有効ターゲット幅 We で定義するモデルは $MT = a + b \log_2\left(\frac{A}{We} + 1\right)$ で与えられる。

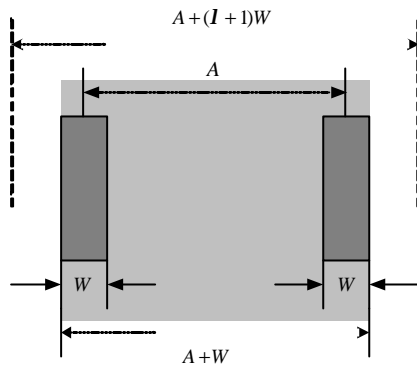


図1 . 入力タスクの範囲

(8)式の ID_s が情報量であり、(5)式で定義した困難度 ID とは意味が異なる。本論文では、 I をパラメータとして扱い、その値は後述の最小 AIC 法で決める。実際にはシステムの構造のほかに操作主体の能力や努力などの人的要因も操作時間に影響を及ぼす。したがって、操作者の人的要因も考える必要がある。

ここで、 P_h でターゲットに当たる確率を表し、それを成功率と呼ぶ。同時に当たらない確率をエラー率と定義すれば、 $P_h + \text{エラー率} = 1$ という関係になる。成功率で反映される人的要因の影響を数量化するために、次式のように人的要因に関する自己情報量 (self-information) ID_h を導入する。

$$ID_h = \log_2\left(\frac{1}{P_h}\right) \quad (9)$$

本論文では各ターゲットの幅とターゲット間の距離の組合せで P_h を計算する。そこで次のようなモデルができる。

$$MT = a + bID_s + cID_h \quad (10)$$

ここで、 MT は操作時間の理論値 (予測値) で、 a 、 b と c はモデルの係数である。

しかし、経験的に操作時間は正規分布に従うことは稀である。操作時間の分布が正規分布でない場合、そのまま(10)式についてパラメータの推定値を求めても偏りを持つ推定結果しか得られない[15]。そこで操作時間のデータをできるだけ正確に正規分布で表現できる形に変換する必要がある。したがって、操作時間について対数変換を施し、対数変換したデータについて線形回帰モデルを構築すれば信頼できる推定結果を得ることができる[4,7,12]。本論文での分布

の概念は操作時間 (時間的分布) に関するもので、従来の座標分布³とは異なる。

本論文で提案するモデルは次式で与える。

$$\ln(MT) = a' + b' \ln(ID_s) + c' \ln(ID_h) \quad (11)$$

ただし、 a' 、 b' および c' は未知係数であるが、前述モデルにおける a 、 b および c とは異なるものである。(11)式は、 $\ln(ID_s)$ と $\ln(ID_h)$ を独立変数とする操作時間の対数 $\ln(MT)$ の回帰式である。実際の操作時間データ (MT_{real}) についてモデルを書くとき、次式のように右辺に誤差項 (k) を付け加えなければならない。

$$\ln(MT_{real}) = a' + b' \ln(ID_s) + c' \ln(ID_h) + k \quad (12)$$

このモデルでは、 ID_s はシステムの構造的要因、例えばターゲット間距離やターゲット幅などの影響を反映するもので、 ID_h は操作主体に関する人的要因の影響を反映するものである。したがって、(11)式のモデル並びに (12)式のモデルはシステムの構造的要因と人的要因を包含している。ここで、このモデルを SH モデル (SH-Model: S と H はそれぞれ Systematic と Human の略) と呼ぶことにする。

ID_s はシステム構造の設定によって決まる (I 値の決定については後述する)。 ID_h は P_h の値によって決まる。ここで、 P_h は二項分布のパラメータとして扱い、ポイントがターゲットに当たるか否かの情報を持つものである。最尤推定 (maximum likelihood estimation) によれば、 P_h の推定値は次のように求められる。

$$P_h = \frac{n}{m} \quad (13)$$

ただし、 n はポイントのヒットした回数で、 m はポイントした総数を表す。

(13)式で P_h を推定する際にして二つの極端なケースを考える必要がある。一つは、すべてのポイントがヒットした場合、つまり、 $P_h = 1$ のときである。もう一つはその逆、つまり $P_h = 0$ の場合である。(11)式はこの二つのケースに適應できない。そこで、 P_h に一様分布の事前分布 (uniform prior distribution) を導入し、ベイズの方法で P_h を求める。その事後分布の平均値は次式で求められる[11]。

$$P_h = \frac{n+1}{m+2} \quad (14)$$

本論文ではこれを P_h の推定値として使用する。

(11)式のもう一つの形式は次式で表す。

³空間的分布とも呼ぶ。画面上にポイントした座標値の集合を指す。

表 1 . 実験 1 の各モデルの推定結果と AIC

モデル	定式化	AIC
Shannon 型モデル	$MT = 197.39 + 75.31 \log_2 \left(\frac{A}{W} + 1 \right)$	38927
We モデル	$MT = -5.05 + 165.31 \log_2 \left(\frac{A}{We} + 1 \right)$	39078
SH モデル	$MT = e^{5.27134} \log_2 \left(\frac{A}{W} + 1 \right)^{0.63651} \log_2 \left(\frac{1}{P_h} \right)^{-0.02755}$	37696

$$MT = e^{a'} ID_s^{b'} ID_h^{c'} \quad (15)$$

3 . モデル評価のための最小 AIC 法

モデルの評価法には 2 つのアプローチがある。一つは従来の誤差の標準偏差や決定係数(R^2)などでモデルを評価する方法である。もう一つは情報量規準によるモデル比較の方法である。前者を用いるとモデルにパラメータを多く導入すれば決定係数が(よって相関係数も)大きくなるが、モデルの予測能力は必ずしも改善されない。つまり、モデルを複雑化すれば見かけ上モデルのデータへの当てはまりがよくなるが、モデルの予測能力が逆に悪くなる可能性もある。これに対して、後者はモデルのデータへの当てはまりとモデルの複雑さを総合してモデルのよさを評価する方法である。これまでいくつかの異なる情報量規準が提案されているが、よく利用されているものに AIC⁴(Akaike Information Criterion) [2,7,12] や BIC[13]などがある。本論文では多くの成功例があることと使いやすさを考慮した結果、AIC を利用することにする。AIC は統計モデル選択基準の一つとして提案されている。「より小さい AIC の値を導くモデルが(より強い予測力をもつという意味で)よりよいモデル」というモデル選択のルールが最小 AIC 法として提案されている[2]。

線形回帰モデルの場合、応用性に優れている最小二乗法が最尤推定法に一致するため、パラメータの

$$^4 AIC \text{ は } AIC = -2\ell(\hat{q}) + 2(\text{parameters' number}) \\ = N(\log 2p\hat{s}^2 + 1) + 2(m+1)$$

で定義される。ただし、 $\ell(\hat{q})$ は最大対数尤度で、回帰モデルの場合 $\ell(\hat{q}) = -\frac{N}{2} \log 2p\hat{s}^2 - \frac{N}{2}$ で計算される。 N

はデータ数、 m はモデルにおける説明変数の数(定数項を含む)、 \hat{s}^2 は誤差分散の推定値であり、

$$\hat{s}^2 = \frac{1}{N} \sum_{i=1}^N e_i^2 \text{ で計算する}(e_i \text{ は誤差を表す})。$$

推定値を求める過程で AIC の結果も同時に計算される。これも最小 AIC 法のメリットの一つである。

HCI の分野ではオリジナル観測データの代わり各条件ごとの平均値(例えば、各ターゲット幅およびターゲット間の距離ごとの平均値)でモデルを推定する手法がよく使われている。線形回帰モデルについて、両方でも同じ推定値を得ることができるが、非線形のモデルではそうならない。統計学の理論によると、標本のサイズが大きいほど推定結果の信頼性が高い。したがって、平均値よりオリジナル観測データを利用したほうが望ましい。また、AIC の計算は大数の法則に基づくものであるから、AIC 計算の立場で平均値よりオリジナル観測データを利用するメリットが見える。したがって、本論文では、モデルの推定に操作時間のオリジナルデータを利用する。

4 . 実験 1: PDA 上の実験

4.1 実験方法

本論文で提案した SH モデルの妥当性と有効性を確認するために PDA 上の実験データを利用した。利用したタスクは Fitts の法則に基づき国際規格 ISO 9241-9 で定められた、一方向相互ターゲットポインティングタスクであった[3]。

実験に参加した被験者は大学生 12 名(男・女各 6 名、年齢は 20 歳から 22 歳)であった。

実験には OS に Windows CE が搭載された PDA を使用した。サイズは 157 mm (横幅) x 79 mm (高さ) x 18 (厚さ)、重さは約 200 g で、表示部分の広さは 480 x 160 pixels (1 pixel = 0.24 mm)であった。ポインティングデバイスとしてスタイラスペンを使用し、Java で開発したソフトを使用した。

被験者は椅子に座った状態で PDA を持ち、スタイラスペンを用いてターゲットをポイントした。タスク中、本体を机の上や膝の上に置かないように指示した。PDA の画面上には白と黒の長方形が交互に提

表 2 . 実験 2 の各モデルの推定結果と AIC

モデル	定式化	AIC
Shannon 型モデル	$MT = 136.46 + 119.99 \log_2 \left(\frac{A}{W} + 1 \right)$	47465
We モデル	$MT = 53.52 + 153.05 \log_2 \left(\frac{A}{We} + 1 \right)$	47859
SH モデル	$MT = e^{5.40134} \log_2 \left(\frac{A}{W} + 1 \right)^{0.70715} \log_2 \left(\frac{1}{P_h} \right)^{-0.00012}$	46077

示された。被験者が白い長方形がポイントするとその長方形は黒く表示され、同時にもう一方の黒い長方形は白く表示された。

また、実験の前に被験者に提示されたターゲットをできるだけ迅速かつできるだけ正確にポイントするよう指示した。各被験者に 10 回の予備練習をさせた。

実験計画は 3 x 3 の被験者内 2 要因計画であり、各要因および水準を次のように設定した。

- ターゲット幅: 10, 20, 40 pixels (2.4, 4.8, 9.6 mm)
- ターゲット間距離: 100, 200, 300 pixels (24, 48, 72 mm)

各被験者は 9 種類のターゲット幅とターゲット間距離の組合せの各条件につき 30 試行を繰り返した。ターゲットの高さは常に 90 pixels で固定されていた。各被験者につき異なる順序でターゲット条件を設定した。

30 回試行の中で最初の試行だけ操作時間は測定できないため、データ総数は 3(ターゲット間距離) x 3(ターゲット幅) x 29(試行) x 12(被験者) = 3132 であった。

実験の中で被験者が時折ターゲットでない方の長方形をポイントすることがあった。これは被験者の一時の勘違いや見込み間違いによるものと考えられる。これらのデータを除去したため、有効データ数は 3132(全部データ数)-118(異常データ数)=3014 であった。

4.2 モデルの推定結果と AIC

ここで前述した実験データを用いて提案の SH モデル((15)式)、Shannon 型モデル((4)式) および有効ターゲット幅 We を用いたモデルを推定し、AIC の値で各モデルの評価を行う。最小 AIC 法によると最小の AIC 値を得たモデルが最もよいモデルと判断できる[2]。

まず、伝統的 Fitts モデル(式(4))の推定結果は

$$MT = 197.39 + 75.31 \log_2 \left(\frac{A}{W} + 1 \right) \quad (16)$$

となり、AIC の値は 38927 である。図 2 はモデルの回帰直線とデータの散布図を示す。

次に、有効ターゲット幅を用いたモデルについてその AIC の値は 39078 であり、伝統的 Fitts モデルの AIC 値よりも大きい。その回帰直線と散布図は図 3 に示す通りである。

SH モデルについて、前述した各モデルと比較可能にするために $I = 0$ と設定した。モデルの推定結果は

$$MT = e^{5.27134} \log_2 \left(\frac{A}{W} + 1 \right)^{0.63651} \log_2 \left(\frac{1}{P_h} \right)^{-0.02755} \quad (17)$$

となっており、AIC の値⁵は 37696 である。

モデルは 3 次元空間の曲面であるが、図 4 はその $\ln(MT) - \ln(ID_s)$ 平面への射影だけを示す。こうすることによって視覚的にモデルのデータへの当てはまりを確認できるようになる。モデルにおける係数 c' の推定値がマイナス(-0.02755)となることは、より正確にターゲットにポイントして成功率 P_h を高めようとする時間がよりかかることを意味する。これは直感的にも容易に理解できることである。

各モデルの AIC の計算結果をまとめて表 1 に示す。この表に示すように、各モデルのなかで SH モデルが最小の AIC 値 (37696)を得ている。つまり、SH モデルが伝統的 Fitts モデルより予測の能力という点で優れている。

また、パラメータ I の値を観測データで適切に決め、あいまいな入力範囲を客観的に確定する必要がある。そこで、 I の値を 0 から 3 までの整数値で変

⁵他のモデルの AIC 値と比較可能なため、ここでの AIC 値は、 $(\ln(MT))$ のモデルの AIC) + 2(対数変換したデータの合計)で計算されている[7]。

動させた各モデルの AIC を計算した。その結果は表 3 に示す通りである。表からわかるように、実験 1 では $I = 1$ のモデルが最小 AIC を得ている。したがって、最適な I 値は 1 に近いことがわかる。

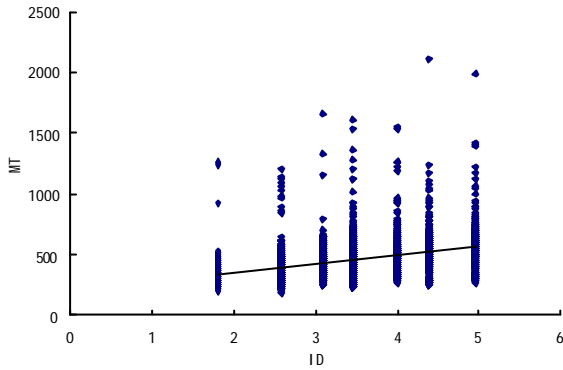


図 2 . Shannon 型モデルの回帰直線と散布図

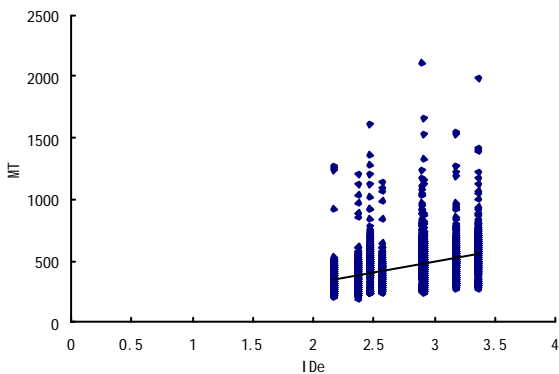


図 3 . We を用いたモデルの回帰直線と散布図

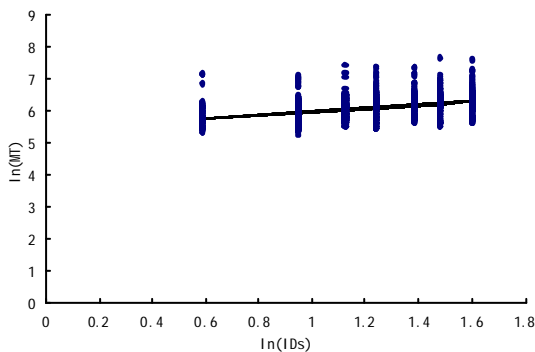


図 4 . SH モデルの予測曲線と散布図 ($I = 0$)

5 . 実験 2: TABLET PC 上の実験

5.1 実験方法

本論文で提案した SH モデルが PDA のデータに限らずほかのデータにも適用できるか、つまり、このモデルは一般性をもつものであるかを明らかにするために、別の条件で実験を行った。Fitts の法則に関する研究に広く利用される規範性のあるデータ作成も本実験の目的であった。

被験者は全員右利きの 12 名 (男性 9 名と女性 3 名で、年齢は 21 歳から 38 歳) であった。

実験プログラムは Windows XP 搭載の富士通 FMV STYLISTIC tablet PC 上で Java によって開発したものである。スクリーン画面のサイズは 21 cm x 15.6 cm (0.2055 mm/pixel) である。タスクの設定とポインティングの仕方は PDA の場合と同様であった。また、実験の前に被験者の全員に同程度の練習をさせた。

実験計画は 3 x 3 の被験者内 2 要因計画であり、各要因および水準を次のように設定した。

- ターゲット幅: 12, 36, 72 pixels
- ターゲット間距離 120, 360, 840 pixels

ターゲットの高さは常に 760 pixels で固定されていた。各被験者は 9 種類組合せの各条件につき 12 試行繰り返した。各被験者は “できる限り正確”、 “できる限り迅速”、 “できる限り正確かつ迅速” という 3 つの別々の教示を受け、実験を繰り返した。この目的は、広い入力範囲 (ポイントした領域) を取ることであった。同様の考え方は Fitts らによっても考案されている [6]。なお、各被験者につき異なった説明順序およびターゲット条件を設定した。

12 回試行の中で最初の試行だけ操作時間は測定できないため、データ総数は 3(回の繰り返し) x 3(ターゲット間の距離) x 3(ターゲット幅) x 11(試行) x 12(被験者) = 3564 であった。

実験 1 と同様に 28 個の外れ値を除いたため、最終的な有効データ数は 3564 - 28 = 3536 であった。

5.2 モデルの推定結果と AIC

各モデルの推定結果と AIC の値は、異なった教示で取ったデータを混合したもので計算した。表 2 はその結果を示す。

これらの結果から、実験 2 のデータについても SH モデルは最小の AIC 値(46077)を得ていることがわかる。つまり、本論文で提案した新しいモデルの有効性が実験 2 のデータでも検証されている。

表 3 . 各 I の値に対応する SH モデルの AIC

実験	$I = 0$	$I = 1$	$I = 2$	$I = 3$
実験 1	37696	37689	37691	37694
実験 2	46077	46037	46032	46039

表 4 . 各モデルの正・負誤差の比率

	SHモデル		W_e モデル		Shannon 型モデル	
	負エラー (%)	正エラー (%)	負エラー (%)	正エラー (%)	負エラー (%)	正エラー (%)
実験 1	47.8	52.2	54.9	45.1	55.1	44.9
実験 2	52.8	47.2	58.9	41.1	59.7	40.3

さらに、実験 2 のデータについての各 I 値に対応する SH モデルの AIC 値が表 3 の下欄に示す。 $I = 2$ のモデルが最小 AIC を得ている。

6 . 考察と結論

表 1 は、実験 1 のデータについて、SH モデルと 2 種類の伝統的なモデルを AIC を用いて比較した結果である。この結果から本論文で提案したモデルは伝統的なモデルと比べて説明力が高いことがわかる。強調すべきことは有効ターゲット幅 (W_e) で定義されるモデルは Shannon 型モデルにも及ばないことである。それは実験 1 が PDA 上で実現されたことに関係していると考えられる。PDA では被験者は手の置き場がなく手を固定することができなかつたため、正確にポイントすることがより難しかったと考えられる。したがって、操作時間のデータの分散はより大きく、各モデル間における AIC の差も大きい。これに対して、実験 2 のデータは相対的に分散が小さくもち、各モデル間における AIC の差も相対的に小さい。にもかかわらず、各モデル間の差は有意に大きなものであった。また、実験 2 のデータにおいて各モデルの AIC 値の順位は実験 1 と同様であることも意味深い。さらに、エラー率を計算した結果、実験 1 と実験 2 ではそれぞれ 26.1% と 10.9% であった。PDA と tablet PC のいずれにおいても SH モデルが最小の AIC 値を得ており、新しく提案されたモデルの有効性がいずれの実験によっても検証されたといえる。

本論文で提案した SH モデルでは、有効入力範囲を決めるパラメータ I が重要な役割を果たしている。各種 I 値について AIC を計算した結果 (表 3)、実験 1 では $I = 1$ 、実験 2 では $I = 2$ のとき AIC が最小と

なっている。つまり、図 5 に示すように実験 1 と実験 2 での有効入力範囲はそれぞれ $(A+2W)$ と $(A+3W)$ となる。最適な I 値が異なった理由は次のように解釈できる。実験のとき、被験者は主にターゲットの現れる位置に集中するので、全体の入力範囲を正確に把握することが困難である。また、相対的に PDA の画面が狭いためその有効入力範囲も狭い。それに対して Tablet PC の画面が広いいためその有効入力範囲は広い。このように、データから有効入力範囲を推定することができるのも最小 AIC 法を使用するメリットの一つである。

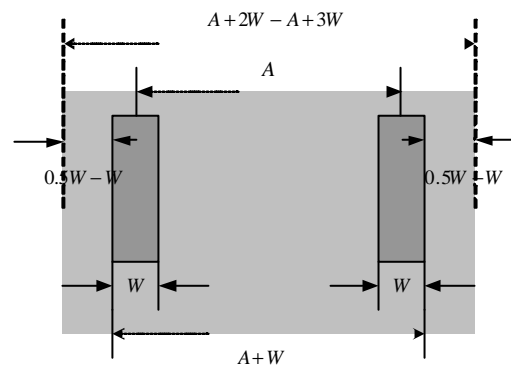


図 5 . $I = 1$ と 2 のときの有効入力範囲

伝統的なモデルの AIC は新しく提案されたモデルよりも大きな値となっている。これは伝統的なモデルに改良の余地があることを示唆していると考えられる。また、パラメータ I を通して有効入力範囲を最小 AIC 法で同定する方法も本研究ではじめて提案したものである。 I をパラメータとしてモデルに導入し、操作時間の観測データでその最適値を決めることに

よってモデルの予測力を一段と向上させることができる。

図 2 と図 3 からわかるように、伝統的モデルの誤差分布に明らかな非対称性が見られる。これに対して図 4 から、新しいモデルの場合データがより均一に回帰曲線の両側に分布し、誤差分布における非対称性の問題が明らかに緩和されたことがわかる(新しいモデルは、 $Ln(MT)$ に関する線形回帰モデルであるため、 MT の代わりに $Ln(MT)$ について作図した。)。表 4 は各種モデルについての正、負誤差の百分比を示してある。この表からもわかるように新しく提案した SH モデルは正・負の誤差の数がほぼ同等であり、ほかのモデルは正・負の誤差数が明らかに異なっている。誤差分布に明らかな非対称性があると、パラメータの不偏推定が得にくくなり、モデルの予測能力も著しく低下してしまう。この困難を緩和するために、本論文では対数変換したデータについて線形回帰モデルを構築する方針を採った。

実験計画の際にエラー率のことが考慮されない場合、または実験の中でエラー率の制御が実現できない場合、実験データの分布が正規分布に従わない可能性が高い。このような場合には Shannon 型モデルの予測力が下がる。しかし、SH モデルは空間的制約を必要としないため、このような困難が生じない。このモデルによって実験者は 4%のエラー率の制限から解放される。つまり、本論文の提案した新しいモデルはエラー率のコントロールやターゲット幅の修正などのことは一切必要としないうえに、高い予測力をもつのである。

本研究のもう一つ大きな貢献は、提案した新モデルにシステムの構造的要因だけでなく、操作主体の人的要因も考慮していることである。これは HCI の分野でははじめての試みである。

新しいモデルの提案とともに、最小 AIC 法によるパフォーマンスモデルの評価方法を HCI 分野に導入することも本論文のもう一つの貢献である。これによって、モデルの評価と選択を AIC 値から自動的に行うことが可能となり、データの特徴に即した新しいモデルの提案がより容易に実現できるようになった。

本論文で取り扱った 2 つの実験では、スタイラスペンを使用した。今後、異なるポインティングデバイスの効率性などの評価実験を行う予定である。

謝辞 本研究に貴重なコメントを頂いた Shumin Zhai 氏 (IBM Almaden Research Center) に感謝する。本研究に対する孔京氏、小笠原将文氏、任研究室のメンバーおよび被験者の皆様の協力に感謝する。

参考文献

1. Accot, J., Zhai, S., More than dotting the I's – Foundations for crossing-based interfaces. *Proc. CHI 2002*, CHI Letters 4(1), 73-80, 2002.
2. Akaike, H. A new look at the statistical model identification, *IEEE Trans. Auto. Control*, AC-19, 716-723, 1974.
3. ISO9241-9: Ergonomic design for office work with visual display terminals (VDTs) – Part 9: Requirements for non-keyboard input devices. 2000, *International Standardization Organization*.
4. Everitt, B.S. *The Cambridge Dictionary of Statistics*, Cambridge University Press, London, 1998.
5. Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381-391, 1954.
6. Fitts, P.M. and Radford, B.K. Information capacity of discrete motor responses under different cognitive sets. *Journal of Experimental Psychology*, 71 (4), 475-482, 1966.
7. Kitagawa, G. and Gersch, W. *Smoothness Priors Analysis of Time Series*, Springer-Verlag, New York, 1996.
8. MacKenzie, I. S., A note on the information-theoretic basis for Fitts' law. *Journal of Motor Behavior*, 21, 323-330, 1989.
9. Mackenzie, I. S. Fitts' Law as a performance model in human-computer interaction. *Doctoral dissertation*. University of Toronto: Toronto, Ontario, Canada, 1991.
10. Mackenzie, I. S. Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, 91-139, 1992.
11. Press, S. J. *Bayesian Statistics. Principles, Models, and Applications*, John Wiley & Sons, New York, 1989.
12. Sakamoto, T., Ishiguro, M. and Kitagawa, G. Akaike Information Criterion Statistics, *D. Reidel, Dordrecht*, 1986.
13. Schwarz, G. Estimating the dimension of a model, *Annals of Statistics*, 6(2), 461-464, 1978.
14. Shannon, C.E. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27, 397-423, 623-656, 1948.
15. Vinod, H.D. and Ullah, A. *Recent Advances in Regression Methods*, Marcel Dekker, New York, 1981.
16. Zhai, S. On the Validity of Throughput as a Characteristic of Computer Input. *IBM Research Report, RJ 10253(A0208-026)* August 21, 2002.