

クリック座標履歴に基づくユーザビリティ評価 (2)適用事例

谷本 諒 河井 悠毅 岡田 英彦

京都産業大学 工学部 情報通信工学科

1. はじめに

本研究では、汎用的かつ容易に記録可能なクリック座標の履歴を用いて期待手順とユーザ手順の差分を抽出する方法を提案し、当該手法を用いた評価ツールの開発および適用事例による有効性検証を行う。提案手法については河井らの論文[1]にて記載し、本論文では適用事例について報告する。

2. 適用事例

2.1 実験方法

Web上のサイトから10種類を評価対象として選択し、各サイトにおけるテストタスクをそれぞれ1種類設計した。これらのタスクを被験者に行ってもらい、本研究の提案手法に必要な操作履歴を記録するために開発したツールを用いて、被験者がPC上で行った操作をバックグラウンドで記録した。実験に協力を得た被験者は20歳代前半の大学生5名であり、PCやWebの利用経験は十分にあるが、本実験のテストタスクの正解操作手順はいずれも未知であった。各タスクの正解操作手順における必要クリック数は平均3.9回である。疲労の影響を抑えるため被験者1人当たりの実験時間が最長でも1時間程度となるよう、タスクの規模に応じて、1タスク当たり5分もしくは10分の制限時間を設けた。得られた5名*10タスク分のユーザ操作履歴、および、各タスクを評価者自身が正解手順に従って行うことで記録した正解操作履歴を用い、タスクごとに、本研究の提案手法を実装した解析ツールを用いてユーザビリティ上の問題点発見を試みた。

2.2 重みづけおよび閾値の設定

本評価手法における操作比較（ベクトルマッチング）では、水平・垂直方向の誤差に対する重み、ベクトル始点間距離および差ベクトルの大きさの加重平方和における重み、2ベクトルを同一と判

定する閾値の5つを決定する必要がある。今回の適用事例では、あらかじめ上記の被験者5名とは異なる被験者1名に対し予備実験および解析を行い、マッチング精度が良好となる重みおよび閾値を調査した。この調査によって得られた値を本実験から得たデータの解析に用いている。

なお、本研究の提案手法では、2操作（2ベクトル）間の距離をベクトル始点間距離および差ベクトルの大きさの加重平方和にて定義している[1]。この加重平方和の重みの一方を0に設定することで、2操作（2ベクトル）間の距離をベクトル始点間距離のみ、もしくは差ベクトルの大きさのみによって求めることができる。前者の場合には、操作の同一性をクリック位置のみに基づいて判定する（つまり、近い位置をクリックしている操作を同じ操作と判定する）ことになる。一方、後者の場合には、操作の同一性をベクトルの向きと大きさの類似性のみに基づいて判定する（つまり、クリックが行われた絶対位置は問わない）ことになる。これらのバリエーションの評価も行うこととした。各方法における5パラメータの設定値を表1に示す。閾値の単位はピクセルである。手法B/Cがそれぞれ上述の前者/後者のバリエーションであり、手法Aがベクトル始点間距離および差ベクトルの大きさを両方用いる手法である。4種類の重み $w_x \sim w_v$ については河井らの論文[1]を参照されたい。

2.4 動画解析との比較

提案手法を用いた評価の問題点発見上の網羅性と効率を評価するため、提案手法を用いて発見された問題点の数を、評価者が手作業で評価を行った場合に発見された問題点の数と比較する。後者の手作業評価は具体的には以下のように行った。まず、この評価に用いる目的で、被験者がテスト

表1: 重みづけおよび閾値の設定

手法	w_x	w_y	w_p	w_v	閾値
A	0.4	1.0	0.5	1.0	100
B	0.4	1.0	0.0	1.0	67
C	0.4	1.0	1.0	0.0	34

タスクを行う際に PC の画面を動画形式で記録しておいた。被験者実験後、評価者はこの動画データを再生して被験者が行った操作を詳細に分析し、期待される操作とは異なる操作が行われた箇所などから、問題点を発見した。また、この手作業評価は、提案手法を用いた評価を行う評価者とは別の評価者が担当し、互いに、一方の手法による問題点発見がもう一方の手法による問題点発見上のバイアスとならないようにした。

各手法を用いた評価によって発見された問題点の数を表 2 に示す。表 2 の値は 10 タスク分の合計値である。手法 A~C で発見された問題点のうち、動画解析で発見された問題点と共通する問題点は 3 手法とも 11 個 (61%) であった。動画解析では発見されず提案手法では発見された問題点の多くは、リンクではない部分をリンクであると勘違いしてクリックした操作から発見された問題点であった。この理由は当該の操作が PC 画面を記録した動画 (映像) のみではわからなかったためであり、タスク実行中のユーザ行動を観察して評価した場合には検出可能であったものと考えられる。

動画解析では各被験者の動画を再生して分析しなければならず、被験者数が多くなるほど所要時間も大きくなるが、提案手法では複数被験者の操作履歴を一括して分析し、共通的な問題点を発見できるため、動画解析に比べ所要時間が大幅に短くなる。この点と表 2 より提案手法は、問題点発見の網羅性では有利とは言えないが、所要時間に対する問題点発見の効率の観点では有利であることが示唆された。

2.5 問題点発見率

提案手法では、検出される余分操作および不足操作のすべてが何らかのユーザビリティ上の問題点を示唆しているとは限らない。したがって、余分操作もしくは不足操作のうち、問題点発見に貢献した操作の割合 (以下、発見率と表す) が大きいほど、より短時間でより効率よく問題点を発見できることになり、望ましいと言える。そこで、上述の手法 A~C について本適用事例における発見率を調査したところ、表 3 の通りとなった。なお、「余分操作 (先頭のみ)」の意味については河井らの論文 [1] を参照されたい。表 3 より以下のことわかる。

表 2: 発見された問題点の数

評価手法	個数
動画解析	18
手法 A	15
手法 B	14
手法 C	13

表 3: 不足 / 余分操作数と発見率

手法	不足操作		余分操作 (すべて)		余分操作 (先頭のみ)	
	総数	発見率	総数	発見率	総数	発見率
A	4	25.0%	375	7.5%	51	49.0%
B	8	37.5%	422	5.7%	58	39.7%
C	1	0.0%	299	9.4%	72	37.5%

- 手法 A~C のいずれにおいても、「余分操作 (すべて)」より「余分操作 (先頭のみ)」のほうが発見率が高い。つまり、検出された余分操作をすべて分析するより、連続する余分操作のうちの先頭のみを分析したほうが効率の観点では望ましいことを支持する結果となった。
- 手法 B (クリック位置のみに基づく操作比較) や手法 C (ベクトルの向きと大きさのみに基づく操作比較) より、手法 A (両方を加重付きで用いた操作比較) のほうが、「不足操作」および「余分操作 (先頭のみ)」において、発見率が概ね上回る。例外は、不足操作において手法 B の発見率が手法 A を上回る点であった。これらより、評価効率の観点では手法 A が概ね有利であること、不足操作の分析に関しては手法 B の能力も高いことが示唆された。

3. まとめ

クリック座標履歴に基づくユーザビリティ評価のための提案手法の有効性を評価するため、Web サイトを対象に提案手法を適用し、他手法との比較や手法のバリエーション間での比較を行った。この結果、評価効率における提案手法の優位性を示唆する結果などが得られた。今後の課題としては、他の評価対象への適用事例や被験者数を増加させた適用事例に基づく評価が挙げられる。

参考文献

- [1] 河井他: クリック座標履歴に基づくユーザビリティ評価 (1) 評価手法; 情報処理学会第 69 回全国大会 (2007)