5Z-01

視線情報と音声情報を組み合わせた Webブラウジング用インタフェースの提案と実装

中川 千拡　　松原 俊一　　Martin J. Dürst

青山学院大学理工学部情報テクノロジー学科

1　はじめに

視線や音声を単独で用いる入力の研究は現在盛んに行われており，新しい入力方法として大いに期待されている [1]．しかし，目は感覚器官として情報を収集する段階と，視線を制御する段階が分離されていないため操作が困難という問題点が知られている [2]．

本研究では，視線情報と音声情報を組み合わせる Web ブラウザ用のタッチレスな操作手法を提案する．具体的な機能として，音声入力による「戻る」・「進む」・「更新」機能，スクロール機能，視線情報と音声入力によるページ遷移機能やテキストボックスへの文字入力機能を実装した．また，実験用の Web ページにおいてマウスとキーボードによる入力と提案手法による入力の操作速度を比較した．

2　提案機能

提案機能の原則として，主に視線情報はマウスカーソルの様にポインティング，音声情報はマウスのボタンやキーボードの様にイベントや入力に用いられる．注視座標に目のアイコンを表示することにより，利用者にカーソル位置を明示する．

2.1　選択の視覚化

注視座標が Web ページのハイパーテキストリンクの表示領域に入った際に注視したと見做し，図 1 のようにリンクを赤枠で装飾する．テキストボックスを注視した際も同様に，図 2 のようにテキストボックスを青枠で装飾する．Web ページ上の装飾は視線情報を取得するたびに更新される．

図 1: 注視されたリンクの装飾の例

図 2: 注視されたテキストボックスの装飾の例

2.2　「戻る」・「進む」・「更新」機能

音声入力によって Web ブラウザの履歴の遷移とページの内容の再読み込みを行う機能である．「もどる（戻る）」・「すすむ（進む）」・「こうしん（更新）」のいずれかの発声によって実行される．

2.3　スクロール機能

音声入力によって Web ページのスクロールを行う機能である．「うえ（上）」・「した（下）」・「みぎ（右）」・「ひだり（左）」のいずれかの発声によって，その方向に滑らかなスクロールを実行する．スクロール量はブラウザの表示領域の幅や高さそれぞれの 75%としている．

2.4　ページ遷移機能

音声入力と視線入力によって Web ページ内のリンクから他の Web ページに遷移する機能である．「くりっく（クリック）」と発声することによって，最後に注視したリンク先のページに遷移する．

2.5　文字入力機能

音声入力と視線入力によって Web ページのテキストボックスに文字を入力する機能である．「にゅうりょく（入力）」と発声することによって，最後に注視したテキストボックスに文字入力が可能になる．「にゅうりょくかいじょ（入力解除）」及び「そうしん（送信）」と発声するまでテキストボックスに音声による文字入力が行われ，「つぎ（次）」・「まえ（前）」と発声することにより文字入力を行うテキストボックスの変更ができる．また，「そうしん（送信）」と発声することによりフォームを送信する．

3　実装

Web ブラウザは Firefox を使用し，利用者の視線情報と音声情報は Firefox の拡張機能である Greasemonkey[1] により処理を行う．Greasemonkey は JavaScript

[1]Greasemonkey, http://www.greasespot.net

Proposal and Implementation of a Web Browsing Interface Combining Eye Tracking and Voice Recognition

Chihiro Nakagawa, Shunichi Matsubara and Martin J. Dürst
Department of Integrated Information Technology, College of Science and Engineering, Aoyama Gakuin University
5–10–1 Fuchinobe, Chuo-ku, Sagamihara, Kanagawa 252-5258, Japan
duerst@sw.it.aoyama.ac.jp

Copyright ©2016 Information Processing Society of Japan.
All Rights Reserved.

で記述されたスクリプトファイルを読み込むことにより，任意の Web ページにスクリプトを追加できる．システム全体の構造は，図 3 に示す．

3.1 視線情報の処理

視線情報は 1 秒間に 60 回視線追跡装置で取得され，Node.js[2] により作成したローカルサーバを経由し，WebSocket プロトコルで Web ページに送信される．

3.2 音声情報の処理

音声情報はオープンソースの音声認識ソフトウェア Julius[3] より文字列として出力される．出力された文字列は，視線情報と同様に Web ページに送信される．また音声認識の精度向上のため，Julius プロセスを二つ起動し，一方は「クリック」や「入力」などの命令を記した辞書，もう一方は文字入力用の辞書を搭載した．

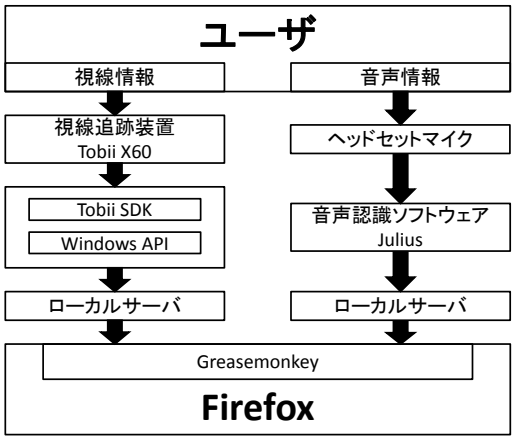

図 3: システム全体の構造図

4 評価

マウスとキーボードを使用する従来手法と，視線入力と音声入力のみで操作を行う提案手法の比較により評価を行う．被験者の視線情報の取得は Tobii X60 で行い，マイクはサンワサプライ MMZ-HSUSB10W を使用した．被験者数は 18 名で，20 代男性が 17 名，30 代男性が 1 名であった．

4.1 ページ遷移機能の評価

実験 Web ページには都道府県名のリンクが 15 種類配置されている．14 個は青色で 1 個は赤色で表示され，赤色のリンクを選択することによって次のページに遷移する．10 回遷移するまでにかかった時間を計測し，従来手法と提案手法を比較した．1 ページ当たりの平均ページ遷移時間は従来手法は 1.214 秒，提案手法は 3.100 秒であった．

4.2 文字入力機能の評価

実験 Web ページには都道府県名とその県庁所在地名，テキストボックス二つ，送信ボタンが表示されている．二つのテキストボックスに指定された文字列を入力し，フォームを送信することで次のページに遷移する．四つのページに入力し，送信するまでにかかった時間を計測し，従来手法と提案手法を比較した．1 ページあたりの平均ページ遷移時間は従来手法は 7.482 秒，提案手法は 11.716 秒であった．

4.3 考察

ページ遷移機能は約 2.6 倍提案手法の作業時間がかかったのに比べ，文字入力機能は約 1.6 倍提案手法の方が遅かった．ページ遷移機能に関して，従来手法ではクリック動作はボタンを押下するのみでよいが，提案手法では「くりっく（クリック）」と発声して音声データを分析するまでの時間がかかる．文字入力機能においては，従来手法では入力や変換のためにキーボードをより多く押下しなければならないが，提案手法では「にゅうりょく（入力）」と入力したい単語を発声するのみで実現する．したがって，提案機能は文字入力のような煩雑な処理により向いていると考えられる．被験者の視線による操作や音声認識に適した発声などの習熟により操作速度やユーザビリティなどの向上が期待できる．また，視線追跡装置や音声認識ソフトウェアの精度の向上もユーザビリティに影響を与えると考えられる．

5 まとめと今後の課題

本研究では，利用者の視線情報と音声情報のみで Web ページのタッチレスな操作を実現した．また，マウスとキーボードを使う従来手法との比較を行った．音声認識の精度向上や利用者の習熟，機能の追加などを行うことによって，さらなるユーザビリティの向上が見込まれる．

参考文献

[1] 岡本龍太郎, 古堂淳也, 松原俊一, Martin Dürst. 視線情報を用いた Web ブラウジング機能の提案と実装. 第 77 回全国大会講演論文集, No. 1, pp. 217–219, 2015.

[2] 大野健彦. 視線を用いたインタフェース. 情報処理, Vol. 44, No. 7, pp. 726–732, Jul 2003.

[3] Julius development team. 汎用大語彙連続音声認識エンジン Julius. http://julius.osdn.jp/, 2016 年 1 月 7 日閲覧.

[2]Node.js, https://nodejs.org/en/

Copyright ©2016 Information Processing Society of Japan.
All Rights Reserved.