

HTML5 対応音声ブラウザの開発

- Chrome 拡張機能の適用 -

Development of Speech Browser using Chrome extension

吉田 龍馬†

大山 実†

† 東京電機大学 情報環境学部

1. はじめに

近年、音声入力インターフェースが急速に注目を集め各種のシステムが開発されている[1][2]. それと同時に HTML5 の登場により動画や音楽等が容易に扱えるようになり、Web も大きく進化を遂げた. 本稿では、このような HTML5 の多様な機能にも対応できる音声ブラウザの開発結果を述べる.

2. 開発環境

OS : OS X Yosemite 10.10.1

ブラウザ : Google Chrome 39.0.2171(64-bit)

開発言語 : Javascript, jQuery 1.11.1, PHP 5.5.3

3. 音声入力インターフェースの実現

3.1 音声認識の手法

音声ブラウザの核となる音声認識機能は Google が提供する Web Speech API を用いる.

3.2 音声認識インターフェースの実装

Web Speech API を用いた音声認識機能, 認識結果に対応するブラウザの動作を Google Chrome 拡張機能に実装することにより, 音声入力インターフェースを実現する.

3.3 ブラウザ操作に対応する単語の検討

ブラウザの操作を行う際は, ユーザーが操作に対応した単語を発声し, それに応じて処理を行うが, 認識結果に誤りがあると意図しない操作が実行される場合がある. 従って, 認識結果の信頼度が低い極端に短い単語は避け, 信頼度の数値に応じて長い文字列を設定する.

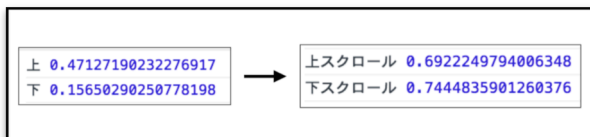


図1 文字列の長さによる信頼度の違い

4. 音声ブラウザの機能

ブラウザ操作の機能として以下の機能を実装した.

- (1) Web ページのリロード, 進む, 戻る, 閉じる
- (2) Web ページの上下左右スクロール
- (3) Web ページの拡大, 縮小
- (4) Twitter による共有機能

- (5) 任意の単語を Google 検索

- (6) 動画ファイルの再生, 停止, 音量調整

- (7) 目的のリンクへのアクセス

閲覧中の Web ページ内のリンクに付加された通し番号の中から, 任意のリンクに対応する番号を発声することにより, 操作が実行される. この機能は, Web ページを読み込む際にリンクを示す a タグ要素すべてに通し番号と, その番号を示すカスタムデータ属性を付加させることによって実現している.

- (8) 現在地情報を用いた天気情報の通知

- (9) 音声フィードバックによる対話型の実現

5. まとめ

本研究において, 音声ブラウザに実装した機能は以下のとおりである.

表1 音声ブラウザに実装した全機能

ブラウザ操作	操作に対応する単語
Webページの更新	更新
進む	進む
戻る	戻る
Webページを閉じる	閉じる
上下左右スクロール	上/下/左/右スクロール
Webページの拡大・縮小	ページ拡大/縮小
Twitterによる共有	ツイッター共有
Google検索	検索*** (任意の単語を付加)
動画の再生・停止	動画再生/停止
音量調整	音量上げる/下げる
任意のリンクへのアクセス	(リンクに付加された番号)
天気情報の通知	天気

6. 今後の課題

従来の音声ブラウザに加え, HTML5 特有の機能である動画・音量制御等を実現した音声ブラウザを開発した. 今後は, 開発したブラウザの詳細な評価を行うとともに, 音声ブラウザならではの音声を活用したサービスを実装していきたい.

参考文献

- [1]宮崎英一: “音声合成を用いた視覚障害者のためのブラウザの開発” 香川大教育実践総合研究 8, 125-133, 2004-03
- [2]Masahide YUASA, Minoru OHYAMA: “Support Tool Using Positions of Contents for Voice/Gaze-controlled Browsers” IEEE Tencon 2005 1D-02.2