

1 はじめに

近年、繁華街のいたる所や電車やバスに置いてあるディスプレイなどに、商品の宣伝や天気情報などを伝える多くのデジタルサイネージが見られる。しかし、街などにあるデジタルサイネージは単にユーザへ情報を伝えるのみで、情報が一方通行のため、必ずしもユーザがほしい情報が手に入るとは限らない。そこで本研究では、Kinect を利用したモーション入力とマイクを利用した音声入力の2つのインタラクティブな要素を加えて、自分が欲しい情報が簡単に得られるような、新しいデジタルサイネージのシステムを提案する。

2 新しいデジタルサイネージの提案

提案するデジタルサイネージでは、動画を再生している時を通常のデジタルサイネージとし、画面の前に人が立ち止まると、身体の入力と Google 社の WebSpeechAPI[1] を使ったコンテンツの移動やズーム、音声認識を使ってコンテンツの検索を行う。コンテンツの検索は一問一答方式を採用する。また、音声処理、身体処理、全体処理は別々のプログラムで動作している。図1に今回提案するシステム概要図を示す。



図1 提案するインタラクティブなデジタルサイネージの概要図

3 WebSpeechAPIの有効性の検証

今回提案するデジタルサイネージシステムでは WebSpeechAPI を利用しているが実装実験の結果として WebSpeechAPI の問題点が浮き彫りになった。WebSpeechAPI の主な問題点の1つとして被験者の滑舌の問題がある。WebSpeechAPI には音声認識結果の候補が出力されるので、その候補機能を使って滑舌の問題を解決できるか調べる。また、候補機能の結果を踏まえて認識率から WebSpeechAPI がデジタルサイネージで有効であるかを調べる。

3.1 仮説

音声認識の結果をデータベースで検索するとき、必ずしも滑舌が良いわけではないので、実際に認識されない場合がある。この問題が起こると、データベースでの検索に引っかからない現象が起きる。そこで、WebSpeechAPI の出力結果の候補機能を使って滑舌の補完ができるのではないかと考えたので検証する。

3.2 実験

あらかじめデータベースには検索されるキーワードとコンテンツ情報が入っている。ここでは、学園祭に関する情報が入っている。実験方法としては、あらか

じめデータベースに入っている単語を録音し、それを WebSpeechAPI に音声認識をさせてデータベースに適合するかを調べる。単語数は272単語とし、被験者は10人とする。認識率はデータベースにヒットしたものを数えるとする。

3.3 結果

実験結果を表1に示す。表1から、候補を含む場合は認識率が上がっていることがわかる。そこで、なぜ認識率が向上したかについて考察する。

表1 音声認識の認識結果(被験者10人の平均)

	候補を含まない	候補を含む
認識できた数	170.90	214.80
認識率	62.83%	78.97%
単語数	272	

3.4 考察

ここでは、候補を含まないで認識させ認識不可だったものと、候補を含むと検索に成功したものを抽出して考える。候補ありで検索した結果で一番多かったのは、滑舌の修正である。例えば、「5CS」が最初の認識では、「OCS」となっていた。これは WebSpeechAPI の結果が検索したいワードに近い音を最初に出力してしまったために、このような現象が起きていると考えられる。

次に空白の修正がある。例として「物理情報研究室」と音声認識させると、「物理情報」と「研究室」の間で半角のスペースが空いてしまう。元々 WebSpeechAPI は Google 社の Android や Chrome ウェブブラウザの検索用に作られているので複数のワード検索と誤動作してこのような動作をするのではないかと考えられる。しかし認識結果から、候補を使った認識率は全体の16.14%しか改善していないので、あまり認識率が向上してないことがわかる。

3.5 結論

候補ありで検索した結果、微量であるが音声認識が向上し、限りなく同じような発音であるという条件での滑舌の補完が確認できた。また、認識率から WebSpeechAPI がデジタルサイネージで有効であることもわかった。

4 まとめ

本稿では、Kinect を利用したモーション入力とマイクを利用した音声入力で、自分が欲しい情報が簡単に得られるような新しいデジタルサイネージを提案し、WebSpeechAPI の候補機能を使って滑舌の問題が解決できるか調べ、WebSpeechAPI がデジタルサイネージで有効であるかを調べた。今後は SQL でのパターン検索を使用し複数の結果を画面上に表示させ、ユーザビリティ向上をはかっていきたい。

参考文献

[1] Web Speech API Specification

<https://dvcs.w3.org/hg/speech-api/raw-file/tip/speechapi.html>