

大型タッチディスプレイでの情報検索のための マルチモーダルインタフェースの検討

及川 隼平 川本 真一
群馬工業高等専門学校 電子情報工学科

1. はじめに

近年、音声インタフェースが様々な機器に搭載され、デジタルサイネージ(電子看板)への応用などが試みられている [1]. 本研究では、デジタルサイネージ(電子看板)のような端末における複数人での情報検索を想定した音声を含むマルチモーダルインタフェースの実現様式について検討する。

2. 大型タッチディスプレイでの情報検索

デジタルサイネージのような端末に入力を行うとき、画面上のキーワードを選択するような場面では、タッチディスプレイのようなポインティングデバイスが素早く正確な操作が可能である。一方自由なキーワードを入力するといった場面では、日常的に利用する音声というモダリティの利用は入力速度の面からも有利と考えられる。しかし、音声入力には認識誤りがあるため、何らかの誤り訂正機能を導入が必要となる。この誤り訂正機能として、入力キーワードの「削除」と「修正」を検討した。音声入力した内容が意図したものでない場合、その内容を削除すれば良い。この「削除」については、音声認識されたキーワードの表示を、タッチ入力ですワイプすることが考えられる。正しく音声入力できたとしても、検索エンジンにとって適切なキーワードであるとは限らない。この「修正」については、入力されたキーワードに適したキーワード候補を提示し、選択することを考えた。このキーワード候補については、音声認識におけるN-best候補を利用することが考えられる。しかし、話した音声に関する類似度であり、検索エンジンにとって適しているとは限らない。その代替として、入力キーワードに対する類語検索結果を候補とするのが望ましいと考えた。

3. 実験

類語検索結果を利用することによる入力修正の有用性について、「レストラン検索」を対象として検討した。グループで飲食店を探す行為は、グループでの相談を基に情報を検索/提示する典型的なタスクの1つであるからである。

入力キーワードが飲食店サイトの検索タグと微妙に異なる場合、十分な数の情報が表示されない可能性がある。そこで、類語検索を行い、最適なキーワードに変換を行うことでどのくらい検索精度が向上するか調べた。検索精度の評価には、再現率、適合率、F値、平均適合率を用いた。F値は再現率と適合率の調和平均であり、この値が高いほど性能が良いことを表す。また、平均適合率は数値が高いほ

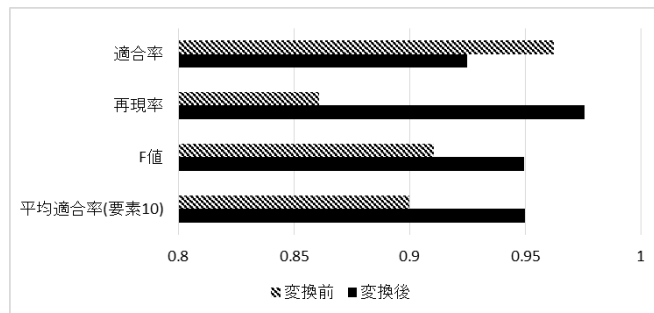


図1: 変換前後の数値評価

ど早い段階でヒットしていることになる。

レストラン検索にはぐるなびのレストラン検索APIを用いた[2]. 類語検索にはWeblio類語辞典を用いた[3].

「群馬県前橋市、喫茶店」を「群馬県前橋市、カフェ」に変換した場合、「埼玉県本庄市、イタリアン」を「埼玉県本庄市、イタリア料理」に変換した場合の検索結果を目視で比較し、数値評価を行った。キーワードの変換前後の数値評価の平均を図1に示す。結果より、類語によるキーワードの置き換えにより、F値・平均適合率共に数値が高くなる場合が確認された。これは、類語検索による入力キーワードの修正が有用であることを示唆する結果であるといえる。

4. まとめと今後の課題

レストラン検索を題材として、大型タッチディスプレイでの利用を想定したマルチモーダルインタフェースについて検討した。今後の課題として、類語の選定基準や提示数など、インタフェース全体を通して複数人が扱いやすいものになっているかという点について、検証を進めることが挙げられる。

謝辞

本研究はJSPS科研費JP25240026, JP15K21024の助成を受けたものです。

参考文献

- [1] 大浦 他, 人工知能学会誌, Vol.28, No.1, pp.60-67 (2013)
- [2] 株式会社ぐるなび, “レストラン検索API,” <<http://api.gnavi.co.jp/api/manual/restsearch/>> (最終アクセス: 2016年10月5日)
- [3] ウェブリオ株式会社, “Weblio類語辞典,” <<http://thesaurus.weblio.jp/>> (最終アクセス: 2017年1月4日)