

理工系

不特定多数のユーザによる誤認識の訂正協力に基づいて
性能が向上する音声情報検索システムを開発

産業技術総合研究所情報技術研究部門主任研究員 後藤 真孝

【研究の背景】

インターネット上でテキストデータが増え続けた結果、テキストの全文検索サービスは不可欠になりました。次に必要なのは、キーワードをタイプ入力するだけで、それを含む音声データを全文検索できる音声情報検索サービスです。

音声データは、それ自体を索引として使えないので、全文検索のためには音声認識によるテキスト化(書き起こし)が必要となります。しかし、従来の音声認識技術では誤認識が多く、新しい言葉にも対応できないため、音声情報検索は実現困難でした。

【研究の成果】

私は、研究分担者の緒方淳らと共に、不特定多数のユーザの協力によって、運用中に性能が向上していく音声情報検索システムを開発しました。そして、日本語のポッドキャスト(インターネット上の代表的な音声データ)を対象とした全文検索サービス「PodCastle」(<http://podcastle.jp>)として公開し、実証実験を開始しました(図1)。

この研究では、従来の音声認識では対応できないような新しい言葉(時事用語、芸能人名など)を、インターネット上のニュース記事や辞書などから自動学習する技術を開発しました。その際、音声認識辞書に言葉を追加するだけでなく、その周辺のつながり(文脈)も学習するようにしたため、日々増え続ける新しい言葉を、的確に認識することが可能となりました。

さらに、音声認識において不可避な誤認識にユーザが気づいたときに、ウェブブラウザ上で効率的に訂正ができる新たなインタフェースを実現しました(図2)。多数のユーザが訂正すればするほど、索引付けが正しくなって検索性能が向上でき、さらに、訂正結果を学習する仕組みを導入することで、音声認識性能も向上できます。

【交付した科研費】

平成19-20年度 基盤研究(B)「不特定多数のユーザが音声認識誤りを訂正した結果を活用した音声情報検索に関する研究」

【今後の展望】

今後は、この「新しい言葉に強く日々成長する音声認識システム」を、様々な応用に展開する予定です。また、今回の研究で提案した「ユーザによる誤り訂正や、インターネット上の情報で性能が向上していくパターン認識技術」という概念自体が学術的に新しいため、その有効性を音声認識分野やそれ以外の分野で検証していきたいと考えています。



図1 ポッドキャストに対する全文検索の画面例:
日々自動収集しているポッドキャストの音声データを音声認識でテキスト化して索引付けし、検索キーワードと照合して全文検索結果を表示します。個々の検索結果では、キーワード周辺の音声認識結果が表示され、実際にものサイト上から音声を聴いて確認できます。

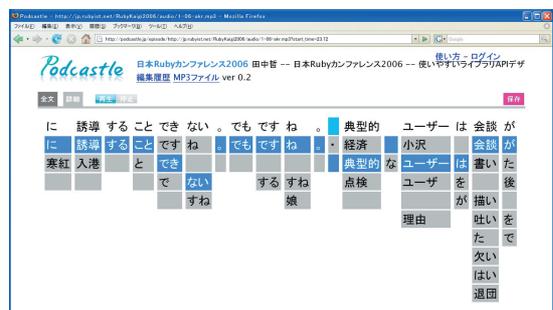


図2 誤認識をユーザが簡単に訂正できるインタフェースの画面例:
音声認識結果に対する複数の候補を求めて下に列挙することで、ユーザは正しい候補を選択するだけで訂正ができます(候補がない場合にはタイプ入力します)。