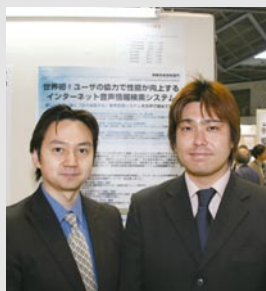


ユーザーの協力で性能が向上する「PodCastle」

新しい言葉に強く、日々成長する音声認識システムを公開



緒方 淳 おがた じゅん(右)
jun.ogata@aist.go.jp
情報技術研究部門
メディアインタラクショングループ
研究員 (つくばセンター)

機械による自動音声認識・理解に関する研究に取り組んでいます。産総研入所以降は、特に実用化を意識し、実環境向けの音声認識技術やインタフェース、さらにインターネットサービスの開発を行っています。今回のPodCastleの開発と運用を通じて、「ユーザーとともに歩む研究」の楽しさを知りました。

後藤 真孝 ごとう まさたか(左)
m.goto@aist.go.jp
情報技術研究部門
メディアインタラクショングループ
主任研究員 (つくばセンター)

計算機によって実世界の音楽・音声コンテンツを自在に扱える技術の確立を目指し、音楽・音声の音響信号の自動理解と、それに基づくユーザーインタフェースの研究を中心に、さまざまな研究課題に取り組んでいます。PodCastleのような、研究として面白くかつ世の中への貢献が大きい技術開発をこれからも全力で頑張ります。

関連情報：

● 参考文献

[1] M.Goto et al.: A Web 2.0 Approach to Speech Recognition Research, Proc. Interspeech (2007).

[2] J.Ogata et al.: Automatic Transcription for a Web 2.0 Service to Search Podcasts, Proc. Interspeech(2007).

[3] 緒方 淳, 後藤 真孝: 情報処理学会論文誌, 48:(1), 375-385(2007).

● プレス発表

2008年6月12日「ユーザーが協力すると性能が向上する音声情報検索システムを実現」

音声情報検索の現状

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が不可欠となりました。次に必要なのは、キーワードを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。しかし、音声データはそれ自体を索引として使えないため、音声認識によるテキスト化が必要となります。従来の音声認識技術では誤認識が多く、新しい言葉に対応できないため、実用的な精度の音声情報検索は実現困難でした。

日々成長する音声情報検索システム

私たちは、不特定多数のユーザーの協力によって性能が向上する音声情報検索システムを開発し、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で、ブログの音声版に相当）を対象とした全文検索サービス「PodCastle（ポッドキャスル）」<http://podcastle.jp>として公開し、実証実験を行っています。PodCastleは、以下の3つの特長をもっています。

(1) インターネット上の日本語ポッドキャストの全文検索が可能

PodCastleでは、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を付けておき、ユーザーが入力した検索キーワードと照合して全文検索結果を表示します(図1)。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、実際にもののサイトから音声を聴いて確認することもできます。



図1 ポッドキャストに対する全文検索の画面例



図2 音声認識の誤認識をユーザーが簡単に訂正できるインターフェースの画面例

(2) 次々と出現する新しい言葉を自動学習

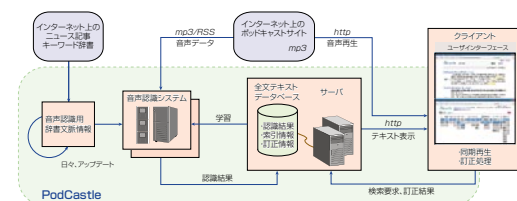
従来の音声認識では対応できないような新しい言葉（新語、時事用語、芸能人名など）を、インターネット上のニュース記事や辞書などから自動学習する技術を開発しました。その際、音声認識辞書に言葉を追加するだけでなく、その周辺のつながり（文脈）も学習するため、よりの確な認識が可能となりました。

(3) ユーザーが協力すればするほど音声の検索・認識の性能が向上

音声認識において不可避な認識誤りにユーザーが気づいたときに、効率的に訂正ができる新たなインターフェースを開発しました。具体的には、音声認識結果に対する複数の候補も求め、図2の画面例のように提示することで、ユーザー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）。これにより索引付けが正しくなるので、検索性能が向上します。さらに、訂正結果を自動学習することで、多数のユーザーが訂正すればするほど音声認識の性能を向上させることができます。

今後の展開

今後は、新しい言葉に強く日々成長する音声認識システムを、インターネットサービスだけでなく、ロボット、コールセンター、会議議事録作成などのさまざまな応用に展開していく予定です。また、ここで提案した「ユーザーによる誤り訂正や、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、この概念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ます。



音声情報検索システム概要