

平成21年6月15日現在

研究種目：基盤研究(B)
 研究期間：2007～2008
 課題番号：19300065
 研究課題名(和文) 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
 研究課題名(英文) A study of speech information retrieval that utilizes results of correcting speech recognition errors by a number of anonymous users
 研究代表者
 後藤 真孝
 独立行政法人産業技術総合研究所・情報技術研究部門・主任研究員
 研究者番号：20357007

研究成果の概要： ユーザ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た。そして、日本語のポッドキャスト（インターネット上の音声データ）を対象とした全文検索サービスをインターネット上で公開して実証実験をした。音声認識誤りにユーザが気づいたときに、効率的に訂正ができる新たなインターフェースを実現した結果、ユーザが訂正すればするほど検索精度が向上でき、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た。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	8,400,000	2,520,000	10,920,000
2008年度	6,000,000	1,800,000	7,800,000
年度			
年度			
年度			
総計	14,400,000	4,320,000	18,720,000

研究分野：総合領域

科研費の分科・細目：情報学、知覚情報処理・知能ロボティクス

キーワード：音声情報処理、情報検索

1. 研究開始当初の背景

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った。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ある。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）が必要となる。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あった。

2. 研究の目的

本研究は、エンドユーザに音声認識誤りを訂正する協力をしてもらうことで、音声情報検索性能、音声認識性能をどこまで高くできるかを探求することを学術的目的とする。また、日本語ポッドキャスト検索のための日本初のWebサイトを公開してエンドユーザの役に立つという社会的目的も持っている。

3. 研究の方法

近年、音声版のブログ(Weblog)ともいえる「ポッドキャスト」が普及し、Web上の音声データ(MP3ファイル)として企業や個人によって多数公開されるようになったため、そうした音声に対する全文検索の重要性が増してきた。そこで本研究では、Web上の日本語のポッドキャストを音声認識によって自動的にテキスト化することで、それらをユーザが検索できるだけでなく、詳細な閲覧、認識結果の訂正も可能なシステムを実現する。

本システムでは、検索したポッドキャストの全文をテキスト表示することで、音を聞く前にポッドキャストを読んで内容を把握することも可能にする。しかし、ポッドキャストの音声は話題が多岐に渡るため、正しく書き起こせるような既存の音声認識器は存在しない。本研究では、これを解決するために、システムが持つすべての情報を積極的にユーザに開示し、ユーザに認識誤りを訂正する協力をしてもらうことで、音声認識率をシステムの運用中に向上させる枠組みを新たに提案する。これにより、音声認識器が徐々に強力になって認識誤りが減ることで、本システムの有用性も高まっていく。

具体的には、課題(1)「Web上のポッドキャスト等のテキスト全文検索技術」を実現すると共に、課題(2)「ユーザが音声認識による自動書き起こしテキストを閲覧し、認識誤りを訂正することを促す方法」を検討して実装し、課題(3)「ユーザが訂正を繰り返した正解情報の自動学習等に基づく音声認識性能の向上技術」を実現してユーザの利便性を高めることに取り組んだ。

4. 研究成果

不特定多数のユーザの協力によって、使用に伴い性能が向上する音声情報検索システムを開発した。そして、日本語のポッドキャスト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った。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精度が向上でき、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た。また、一部のポッドキャストに対しては人手による書き起こしをおこなって、性能評価のための新たなデータベースを構築した。

より詳細には、三つの課題に対する成果を下記に述べる。

- (1) 課題(1)に関しては、Webクライアント用インタフェース、音声認識器に加え、音声認識状態管理部、データベース管理部、検索サーバを実装し、検作用Webサイトを構築した。音声認識状態管理部では、複数の音声認識器の負荷をモニタリングすることで、データベース管理部と連携してポッドキャストを次々と認識可能にした。検索サーバでは、ユーザによるテキスト全文検索とインタフェースの画面遷移を可能にした。また、単純な全文検索だけでなく、より高精度で、効率のよい検索を目指して、ポッドキャストのエピソード間の類似検索の検討もおこなった。
- (2) 課題(2)に関しては、ユーザがポッドキャストの音声認識結果を閲覧し、音声認識誤りを訂正できる機能をWebクライアント用インタフェースとして実現した。効率的に閲覧(視聴)するために、音声の再生に同期してテキスト中のカーソル(ハイライト)を動作させる機能や、音声の再生速度を自在に速くしたり遅くしたりできる機能を構築した。音声認識結果の訂正に関しては、ユーザが見通し良く効率的に訂正できるように、音声認識の中間表現(単語グラフ)を要約したconfusion network(信頼度付き競合候補)を訂正インタフェースとして利用した。また、音響イベント検出において音楽の有無の自動判定を改良し、笑い声等の音響イベントの検出に関する検討もおこなった。
- (3) 課題(3)に関しては、PodCastleを通じて得られる集合知(不特定多数のユーザからの訂正結果)に基づく音響モデル、言語モデルの学習手法を実現した。ポッドキャストは、ニュース、講演、インタビューなど、言語(内容)的にも音響的にも多様な特性の音声データを含むといった問題がある。そこでポッドキャストのメタ情報(RSS)の仕組みを利用して、ポッドキャストの各番組に依存した音響モデル、言語モデルを構築した。これにより、ユーザからの訂正結果をより生かすことができ、実験の結果、ポッドキャストに対する音声認識性能を大幅に向

上させることができた。ユーザからの訂正結果に含まれる未知語については、語彙制約なし音声認識(音素タイプライタ)を用いた発音系列の自動推定をおこない、音声認識用の単語辞書に自動登録する機能(未知語自動獲得機能)を構築した。

実際に、訂正が比較的多くなされたポッドキャストにおいて、上記手法により学習が進むにつれて、新しく配信される音声データの認識結果ではあきらかに認識誤りが減り、ユーザの訂正の負担や訂正時間が減少していたことが確認された。

本研究の成果は、学術的に高い評価を得ているだけでなく、産総研が2008年6月12日(木)にプレス・リリース「ユーザ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：インターネット上の音声情報を認識・検索するサービスを開始」(http://www.aist.go.jp/aist_j/press_release/pr2008/pr20080612/pr20080612.html)をおこなった結果、社会的にも高い関心を集めた。具体的には、テレビの二つの番組(日本放送協会(NHK)「首都圏ネットワーク」及び日本放送協会(NHK)「いばらき わいわいスタジオ」)と新聞9紙(日本経済新聞、毎日新聞、読売新聞、日経産業新聞、日刊工業新聞等)で報道された。さらに、科研費 NEWSレター 2008年11月発行 p.8でも「最近のユニークな研究成果の例」として本研究成果「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」が取り上げられ、文部科学省からもユニークで優れた研究成果であることが認められていることがわかる。

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ある。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る。



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する。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、実際にもとのサイト上から音声を確認できる。



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る(候補がない場合にはタイプ入力が可能)。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計1件)

- ① 緒方 淳, 後藤 真孝, 伊藤 克亘: “有声・無声休止区間の自動検出を考慮したデコーディングによる自由発話音声認識の性能改善”, 電子情報通信学会論文誌 D-II, Vol. J92-D-II, No. 2, pp. 226-235, February 2009. (査読あり)

[学会発表] (計7件)

- ① 緒方 淳, 後藤 真孝: “PodCastle: 集合知を活用した音響モデル学習による音声認識の性能向上”, 日本音響学会 2009年 春季研究発表会, 2009/03/18, 東京, 日

- 本
- ② 緒方 淳, 後藤 真孝: "PodCastle: ポッドキャスト音声認識のための集合知を活用した音響モデル学習", 第3回音声ドキュメント処理ワークショップ, 2009/02/28, 愛知, 日本
 - ③ Junta Mizuno, Jun Ogata, and Masataka Goto: "A Similar Content Retrieval Method for Podcast Episodes", 2008 IEEE Workshop on Spoken Language Technology (SLT 2008), 2008/12/18, Goa, India
 - ④ 緒方 淳, 松原 勇介, 後藤 真孝: "PodCastle: 集合知に基づく Web キーワードを考慮した言語モデリング", 日本音響学会 2008 年秋季研究発表会, 2008/09/12, 福岡, 日本
 - ⑤ 緒方 淳, 後藤 真孝: "confusion network と語彙制約なし音声認識を用いた動的発音モデリング", 情報処理学会第 72 回音声言語情報処理研究会, 2008/07/18, 盛岡, 日本
 - ⑥ 水野 淳太, 緒方 淳, 後藤 真孝: "ポッドキャストを対象とした類似エピソード検索手法", 情報処理学会第 71 回音声言語情報処理研究会, 2008/05/23, 神奈川, 日本
 - ⑦ 松原 勇介, 緒方 淳, 後藤 真孝: "ポッドキャスト音声認識の性能向上手法: 集合知によって更新される Web キーワードを活用した言語モデリング", 情報処理学会第 71 回音声言語情報処理研究会, 2008/05/23, 神奈川, 日本

[その他]

○報道等

- ① 日本放送協会 (NHK) 「首都圏ネットワーク」 2008 年 6 月 12 日 (木) 18:10~19:00
- ② 日本放送協会 (NHK) 「いばらき わいわいスタジオ」 2008 年 6 月 12 日 (木) 18:10~19:00
- ③ 日本経済新聞 2008 年 6 月 13 日 (金) 13 面
- ④ 日経産業新聞 2008 年 6 月 13 日 (金) 3 面
- ⑤ 日刊工業新聞 2008 年 6 月 13 日 (金) 26 面
- ⑥ 毎日新聞 2008 年 6 月 13 日 (金) 27 面
- ⑦ 読売新聞 2008 年 6 月 13 日 (金) 33 面
- ⑧ 東京新聞 2008 年 6 月 13 日 (金) 3 面
- ⑨ 茨城新聞 2008 年 6 月 13 日 (金) 21 面
- ⑩ 常陽新聞 2008 年 6 月 23 日 (月) 3 面
- ⑪ 日本情報産業新聞 2008 年 6 月 30 日 (月) 2 面
- ⑫ 文部科学省 科学技術動向 2008 年 7 月号 2008 年 7 月 26 日 (土) p. 5
- ⑬ 文部科学省 科研費 NEWS レター 2008 年

- 11 月発行 Vol. 2, 2008, p. 8
- ⑭ 日経 BP 社 日経エレクトロニクス Tech-On!: 「みなさんで育ててください」, 産総研が Podcast 向け音声情報検索システムを開発 (2008/06/12)
<http://techon.nikkeibp.co.jp/article/NEWS/20080612/153228/>
- ⑮ 毎日 jp: 音声検索: 音声データを文字変換・検索するシステム開発 (2008/06/12)
<http://mainichi.jp/select/science/news/20080613k0000m040078000c.html>
- ⑯ YOMIURI ONLINE: ネット上の音声データを全文検索、国内初のサービス公開 (2008/06/12)
<http://www.yomiuri.co.jp/science/news/20080612-OYTIT00825.htm>
- ⑰ インプレス INTERNET Watch: 日本語ポッドキャストの全文検索サービス、産総研が公開 (2008/06/13)
<http://internet.watch.impress.co.jp/cda/news/2008/06/13/19932.html>
- ⑱ RBB TODAY: 産総研、インターネット上の音声情報を認識・検索するサービスを開始～ユーザが協力で性能向上 (2008/06/13)
<http://www.rbbtoday.com/news/20080613/51954.html>
- ⑲ ASCII.jp: ポッドキャストの音声データを検索できる! (2008/06/30)
<http://ascii.jp/ele/000/000/145/145386/>

○広報

- ① 産総研プレスリリース (2008/06/12): 「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: インターネット上の音声情報を認識・検索するサービスを開始」
http://www.aist.go.jp/aist_j/press_release/pr2008/pr20080612/pr20080612.html
- ② 産総研 TODAY Vol. 8, No. 8 (2008/08/01): 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
http://www.aist.go.jp/aist_j/aistinfo/aist_today/vol08_08/p17.html
- ③ AIST TODAY 2008 No. 30 Autumn (2009/01/13): PodCastle: A speech information retrieval system in which its performance can be improved by users' contribution
http://www.aist.go.jp/aist_e/aist_today/2008_30/hot_line/hot_line_14.html

○ホームページ情報

- ① PodCastle (ポッドキャストル)
<http://podcastle.jp/>
- ② PodCastle とは
<http://podcastle.jp/info/>
- ③ 産総研プレスリリース
http://www.aist.go.jp/aist_j/press_release/pr2008/pr20080612/pr20080612.html

6. 研究組織

(1) 研究代表者

後藤 真孝 (GOTO MASATAKA)
独立行政法人産業技術総合研究所・情報技術
研究部門・主任研究員
研究者番号: 20357007

(2) 研究分担者

緒方 淳 (OGATA JUN)
独立行政法人産業技術総合研究所・情報技術
研究部門・研究員
研究者番号: 10392599

江渡 浩一郎 (ETO KOUICHIROU)
独立行政法人産業技術総合研究所・サービス
工学研究センター・研究員
研究者番号: 20311277