

2000時間の録音の中から3秒でキーワードを含む音声を検索

高精度かつ高速な大規模音声検索技術を開発

株式会社日立製作所(執行役社長:古川 一夫/以下、日立)は、このたび、大規模な録音データの中から、任意のキーワードを含む音声を高い精度で高速に検索する技術を開発しました。本技術は、録音データを、音素(音声の基本単位)記号処理を用いて高速に検索した上で、音声特徴量^{*1}を用いたより精緻な再検索を行い、段階的に精度を高めながら音声を検索する技術です。これにより、2000時間の録音の中から3秒でキーワードを含む音声を検索することが可能になります。本技術は、今後、音声付き映像コンテンツの検索をはじめ、コールセンターにおいて顧客の生の声を迅速に検索し、商品やサービスの向上に役立てるなど、幅広い応用が期待されます。

近年、ストレージの大容量・低価格化やブロードバンドの普及にともない、音声付き映像コンテンツを含む大規模な録音データが、データセンターやネットワーク上、個人のPCなどに蓄積されています。これらの録音データを有効活用するために、膨大な録音データの中から必要なキーワードをすばやく探し出す検索技術が求められています。従来の音声検索技術には、録音データを音素の記号列に変換しておき、高速にキーワードの照合を行う方式と、録音データの音声特徴量を解析し、高い精度で照合を行う方式がありましたが、前者には検索の精度に、後者には、照合処理に膨大な時間がかかるという問題がありました。

このような背景から、日立では、音素の記号列と音声特徴量による検索方式を組み合わせ、任意のキーワードに対して段階的に照合を行うことによって、2000時間という大規模な音声データに対しても、高い精度を保ったまま、3秒という高速で検索する音声検索技術を開発しました。

今回開発した技術は、2008年10月8日から10月10日まで、オーストラリアのケアンズで開催される国際学会「IEEE Signal Processing Society 2008 International Workshop on Multimedia Signal Processing」で発表します。

なお、本研究では、音声特徴量の学習およびシステム性能評価に際し、独立行政法人国立国語研究所、同じく情報通信研究機構、国立大学法人東京工業大学が共同開発した『日本語話し言葉コーパス(CSJ)』を利用しました。

■開発技術の詳細

(1) 音素記号インデックス(索引)を利用した高速検索技術

録音データを音素の記号列に変換し、音素がどのタイミングで出現するかをインデックス(索引)として保存する方式を用いた検索を行います。今回、このインデックス処理を最適化し、不要なインデックスを削減することで、任意のキーワードを含む音声を高速に検索することができます。

(2) 音声特徴量を利用した高精度な検索技術

音素による検索処理の後で、音声特徴量を用いたより精度の高い再検索を行います。今回、音声特徴量を用いた処理を2段階にわけ、段階的に精度を高めながら合計3回の検索処理を行うことで、検索速度を損なうことなく、高精度なキーワード検索を実現しました。

*1 「あ」らしさ、「い」らしさなど、いわゆる音韻を特徴づける物理量。例えばスペクトル(周波数成分)など。

■照会先

株式会社日立製作所 中央研究所 企画室 [担当:木下]
〒185-8601 東京都国分寺市東恋ヶ窪一丁目 280 番地
電話 042-327-7777(直通)

以上

このニュースリリース記載の情報(製品価格、製品仕様、サービスの内容、発売日、お問い合わせ先、URL 等)は、発表日現在の情報です。予告なしに変更され、検索日と情報が異なる可能性もありますので、あらかじめご了承ください。
