

音声認識技術

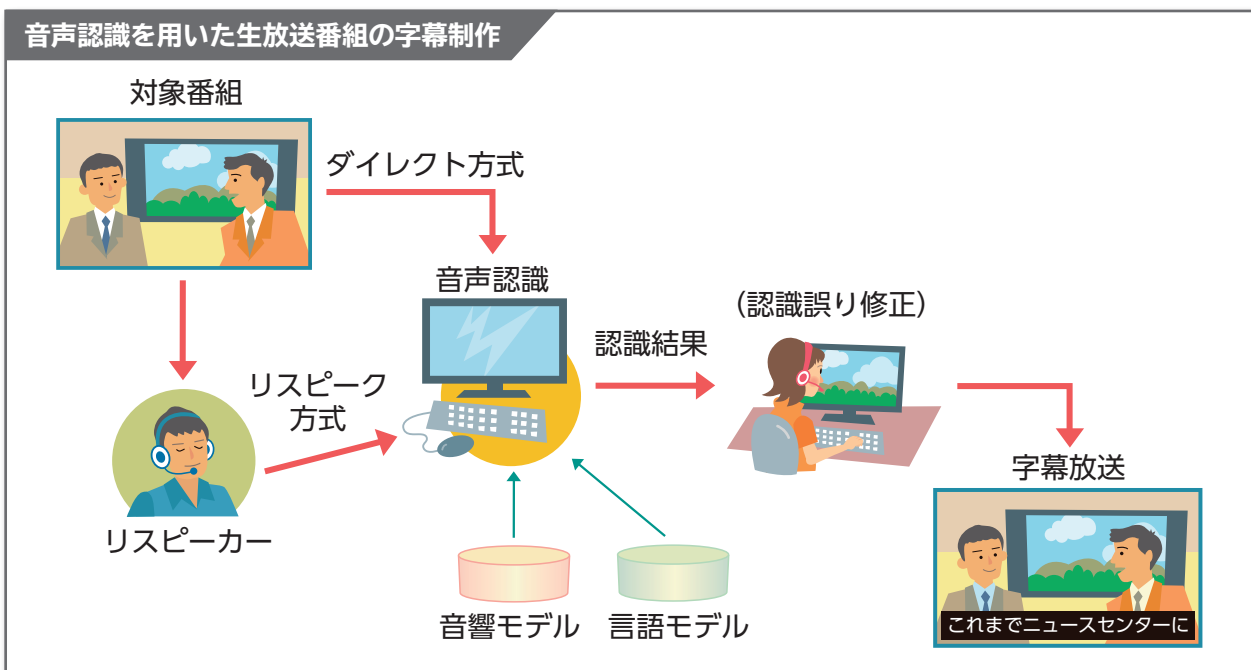
ニュースや情報番組、スポーツ中継などの生放送番組の字幕制作で利用している音声認識技術です。

利用分野

- 字幕制作システム
- 放送番組の書き起こし作成
- 音声コンテンツのアーカイブ、検索システムの構築

特長

- 1 アナウンサーの原稿読み上げ、現場からレポートする記者の音声などを直接認識することができ、95%以上の認識精度が得られます。
- 2 スポーツ中継や対談が含まれる情報番組の音声を、別の話者が言い直した音声を認識することで、字幕制作に十分な認識精度が得られます。
- 3 発話の終わりを待つことなく、リアルタイムに認識結果が得られ、字幕の表示遅れを短縮することができます。
- 4 認識対象番組に合わせた話題に適応化して認識精度を向上することができます。



キーワード ▶ 字幕制作 / 音声書き起こし / 音声認識

1 音声認識技術

● **音響モデルと言語モデル**：音声認識には、入力された音声かどの単語かを算出する音響モデルと、単語と単語のつながりやすさから認識結果を決定する言語モデルが用いられています。これらのモデルを話者や話し方のスタイル、番組で取り上げられる話題に適応化することにより認識精度を向上することができます。日々のニュースを認識するための言語モデルの適応化をはじめ、スポーツ番組の競技に合わせた言語モデルの適応化や、情報番組の広範な話題を認識するための言語モデルの適応化技術が字幕制作に用いられています。

● **音声認識アルゴリズム**：音声認識アルゴリズムは、発話の終わりを待つことなく、認識結果を逐次出力できるアルゴリズムにより、生放送番組の字幕の表示遅れを短縮することができます。

2 字幕制作システムへの応用

● **2つの音声認識方式**：生放送番組における音声認識技術を用いた字幕制作では、番組の音声を直接認識するダイレクト方式と、番組とは別の話者が番組の内容を言い直した音声を認識するリスピーク方式があり、番組音声の認識の難易度に応じて使い分けています。特にリスピーク方式は、背景雑音や感情的な話し方がある番組に有効で、複数話者の対談を整理して入力することもできます。ダイレクト方式とリスピーク方式の長所を組み合わせたハイブリッド方式によって、さまざまな番組で高精度な音声認識が実現可能です。

3 放送番組書き起こしシステムへの応用

ニュースなどの放送番組をすべて収録し、音声認識を用いて書き起こしを付与するシステムです。キーワードを用いて過去のニュース番組などを検索し、キーワードが発声された位置から番組を再生することができます。

提供可能な技術

- リアルタイム音声認識技術
- 音声認識適応化技術
- ダイレクト・リスピーク方式の音声認識システム構築技術

関連特許

特許第 4791857 号	発話区間検出装置及び発話区間検出プログラム
特許第 4981076 号	発音辞書修正装置、音声認識装置、およびコンピュータプログラム
特許第 5149107 号	音響処理装置およびプログラム
特許第 5184467 号	適応化音響モデル生成装置及びプログラム
特許第 5296598 号	音声情報抽出装置