

音声認識技術を用いた自動車室内音環境の評価に関する研究

- 音声認識を用いた評価システムの構築 -

照明音響科 竹居 翼 木野直樹

Study on evaluating the sound environment inside vehicles using automatic speech recognition technology

- Building an evaluation system using automatic speech recognition -

TAKEI Tasuku and KINO Naoki

Keywords: automotive interior, automatic speech recognition, sound environment

自動車室内の音環境の調整には自動車内装が重要であり、当所では自動車内装の開発に必要な技術支援を実施している。本研究では、自動車内装の技術支援に活用するため、音声認識技術を利用した車室内音環境の数値化を試みる。この実現のため、音声認識ツールキットを利用して評価システムを構築し、既存の音声コーパスを利用して雑音の無い音声の評価した。その結果、単語誤り率2.5%が得られ、雑音の無い音声を概ね正しく認識できることを確認した。

キーワード：自動車内装、音声認識、音環境

1 はじめに

近年、EV化によるモーター騒音や車両接近通報装置の警報音等、自動車の走行音は大きく変化しており、自動車室内の静粛化には、自動車内装がこれまで以上に重要となっている。工業技術研究所では、自動車内装の開発に必要な吸音率等の音響試験を実施している。

本研究では、自動車内装の技術支援に活用する目的で、近年利用が盛んな音声認識技術を用いた車室内音環境の数値化を試みる。この実現のため、音声認識ツールキットを利用して評価システムを構築し、既存の音声コーパスを利用して雑音の無い音声の評価した。その結果、単語誤り率2.5%が得られ、雑音の無い音声を概ね正しく認識できることを確認した。

2 方法

音声認識技術を利用した車室内音環境の数値化の流れを図1に示す。

発話内容付きの音声データが収録された音声コーパスを使用した。

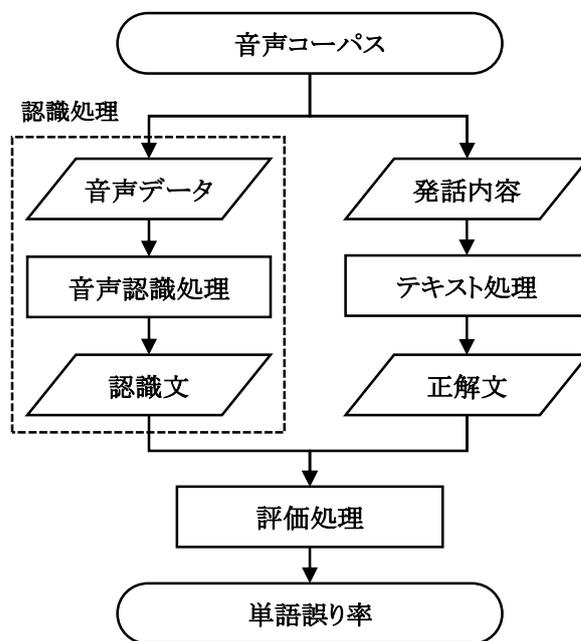


図1 音声認識を用いたシステムの概要

認識処理では、音声データを音声認識処理して得られた認識文を作成する。ここでは、クラウドサービスによる音声認識API等を利用しない音声認識ツールキットを利用して実装した。

評価処理では、音声データに付属する発話内容をテキスト処理して作成した正解文を使用して、認識文との差異を数値化する。ここでは、正解文を作成する処理、および認識文と正解文を比較して単語誤り率を計算する処理を実装した。

3 結果と考察

3.1 構築した音声認識システム

音声認識ツールキットは、Open AI, Inc. が提供しているオープンソース版 Whisper¹⁾ を利用した。このツールキットは、モデルパラメータ数が異なる5種類 (tiny、small、base、medium、large) の事前学習済み認識モデルを利用できる。今回は medium の認識モデルを利用して認識処理を実装した。単語誤り率を計算する評価処理は、Hugging Face, Inc. が提供している音声認識性能の評価用の Python モジュール²⁾ を利用して実装した。これら2つの処理は、ローカルコンピュータ上に構築した。

3.2 雑音の無い音声データの評価

英語音声コーパスの LibriSpeech ASR corpus³⁾ に収録されているデータセットのうち、雑音の無い音声データセット (test-clean セット、全 2,620 データ) を評価した。

評価したデータセット全体の単語誤り率は 2.5% であった。このときの各音声データの単語誤り率の分布を図2に示す。この分布より、2,000 ファイル以上の音声データの単語誤り率が 10% 未満であった。今回構築したシステムによって、雑音の無い英語音声を概ね正しく認識できることを確認した。

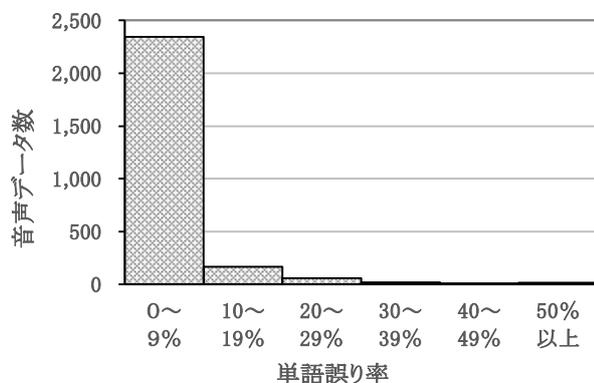


図2 英語音声の単語誤り率の分布

4 まとめ

音声認識結果を評価するシステムを構築し、雑音の無い音声データを評価した。その結果、雑音の無い音声を概ね正しく認識できることを確認した。今後、雑音等のある音声や実車室内を再現した音声データの認識評価を進め、自動車内装が車室内音環境へ及ぼす影響の指標化を目指す。将来的には、本研究の成果を活用して、自動車内装の技術支援の充実化を図る。

謝辞

本研究を進めるにあたり、音声情報処理に関することについてご指導いただきました中部大学工学部山本一公教授に深く感謝いたします。

参考

- 1) OpenAI, Inc.: GitHub openai-whisper, <https://github.com/openai/whisper> (2024.3.18 アクセス)
- 2) Hugging Face, Inc.: GitHub evaluate, <https://github.com/huggingface/evaluate> (2024.3.18 アクセス)
- 3) Open Speech and Language Resources: LibriSpeech ASR corpus, <https://www.openslr.org/12/> (2024.3.18 アクセス)