次の停車駅を文字表示させるシステム

東京都立多摩科学技術高等学校
門野玲也　北田淳之介　髙橋風純

研究背景

聴覚障碍者の方は電車の走行音や人の話し声が主な原因となって駅や電車のアナウンスを正確に聞き取ることができず、取得できる情報に限界がある。[1]

緊急時、電車の運行状況等に関する情報は電波によって特殊な機械に現在の情報を表示させる方法があるが、場所による制約がある。[2]

研究方針

車内アナウンスを直接文字に起こすことによって
場所を限定せずに聴覚情報を文字情報に変換する
アプリケーションを提案。

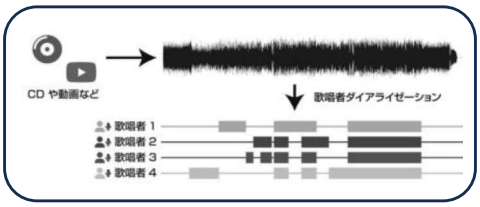
研究目的

聴覚障碍者の方が聞き取りにくい電車のアナウンスを文字に視覚化して現在地がわからないことに対する不安を解消する。

設計

使用言語　Python
実行環境　Google Colaboratory
使用ライブラリ
・pyannote.audio ・whisper ・pydub ・numpy ・webvtt ・ginza
・pyaudio ・pykakasi ・pandas ・spacy

1．電車のアナウンスを録音、話者ダイアライゼーション※1によって話者分離



2．whisper※2で文字起こし[3]

3．特定のワードを含む文を取り出す

まもなく立川、立川。お出口は右側です。立川の次は国分寺に停まります。国立、西国分寺へおいでのお客様と、青梅線、南武線、多摩モノレールはお乗り換えです。

↓ まもなく、つぎは

まもなく立川、立川。

4．文と電車のアナウンスの定型文とのコサイン類似度※3を計算し、最も値が高い定型文を画面に表示[4]

※1：どこの時間でどの話者がしゃべったのか、話者認識をせずに実施する技術。

※2：高い精度で入力した音声を文字起こしすることが可能な音声認識モデル。

※3：ベクトルの向きの類似度を測る指標で、２つのベクトルが「どのくらい似ているか」という類似性を表す尺度。自然言語処理においては文書の類似性を評価するもの。

実験

（１）駅のホームで録音した場合
YouTube上の、駅のホームから録音した電車のアナウンスの音声データを使用し、人の話し声がある状況で電車のアナウンスを正しく文字起こしできるか確認する。

（２）電車内で録音した場合
電車内でスマートフォン、イヤホン内臓のマイクを使用して電車のアナウンスを録音し、ノイズが多く含まれる状況で電車のアナウンスを正しく文字起こしできるか確認する。

結果

（１）ホーム
駅のホームからのアナウンスは読み取ることが出来た。
この時音声データには人の話し声などの音も入っていたが、実験の結果は問題無かった。

[00:00.000 --> 00:06.000] 今日も新幹線をご礼送らさえまして、ありがとうございます。
[00:06.000 --> 00:10.000] この電車は、謎にも、新大阪雪です。
[00:10.000 --> 00:18.000] 途中の停車駅は、しながは、しんよこはま、なもや、きょうとです。

（２）電車内
電車でのアナウンスはwhisperがアナウンスを聞き取ることができなかったため、正しく文字お起こしすることが出来なかった。
この時音声データには人の話し声などは入っておらず、電車自体のノイズなどは入っていた。

[00:00.000 --> 00:30.000] お手に入らして、お手に入れ、お手に入れ、お手に入れ、お手に入れ、お手に入れ、お手に入れ、お手に入れ、お手に入れ、お手に入れ、

考察

結果より、ホームと電車内での実験結果を比較した時、読み取りの結果は人の話し声などには影響を受けないと考えられる。従って、読み取りの結果は電車内でのジョイント音などのノイズに影響されていると推測できる。

今後の課題

・録音の精度をあげるために音声データのノイズ除去が必要。
・現在限られた駅での実験となっているので、中央線以外の私鉄会社での実験が必要。
・さらに、現在停車駅のアナウンス読み取りの場合に限定しての実験なので様々なアナウンスのパターンを読み取りできるプログラムにする必要がある。

参考文献

[1]東洋経済ONLINE「「聴覚障害者」にとって通勤電車はこんなに怖い」　2019/7/15　柳澤　美樹
[2]特許情報プラットフォーム　文字情報提供システム
1994/9/2　新橋　龍男　近藤　芳人
[3]WhisperとPyannoteを用いた話者分離と音声認識
2023/10/30　Hakky
[4]「コサイン類似度」で文書がどれだけ似ているかを調べてみた
2022/12/21　shirota