

# 大正時代の歴史的音源の音声認識精度の向上に関する検討

Study on Improving the Accuracy of Speech Recognition  
for Historical Sound Sources of Taisho Era

轟 禎政  
指導教員 三輪 賢一郎

サレジオ工業高等専門学校 機械電子工学科 情報コミュニケーション研究室

大正期の歴史的音源の内容を文字情報として後世に伝えていくことには大きな意義がある。しかし、人力で文字起こしを行うとなると、莫大なコスト、時間、労力を要する。本稿では音声認識 API による認識精度の向上のために雑音除去の前処理を施すことで、文字起こしのさらなる省力化を検討する。

キーワード：歴史的音源, 語りもの, 雑音除去処理, 音声認識, 文字起こし

## 1. 緒言

明治・大正期から昭和の初めにかけて、当時の演説や講談、落語などの「語りもの」をはじめとして、様々なジャンルの貴重な録音がレコード盤に遺されている[1]。これらの膨大な価値あるコンテンツを、文字情報として後世に伝えていくことは重要である。そこで、当時の音源を耳で聞きながら文字起こしを手作業で行うとなると、莫大なコスト・時間・労力を要することは明白である。

先行研究では、音声認識 API を用いた大正期の歴史的音源に対する文字起こしの自動化を検討したが、一定の可能性は認められたものの、省力化までは至らない結果となった[2]。そこで、本研究では音声信号に雑音除去の前処理を施すことで、歴史的音源に対する文字起こしの省力化の可能性を検討する。

## 2. 方法

本検討では使用する音声認識システムとして、OpenAI が公開している Whisper を使用した[3]。また雑音除去の手法として、歴史的音源の雑音除去に有効とされる「GHA Denoiser」[4]と、一般的な雑音除去手法である「スペクトルサブトラクション法(Spectral Subtraction Method)」の2種類で比較を行った。

認識結果を評価する指標としては、以下の算出式で示される文字認識率(%)を用いることとし、モーラ(拍)に着目する平仮名ベースでの文字認識率に加え、漢字ベースの文字認識率も算出して評価を行った。

文字認識率 =

$$\frac{\text{正解文字数} - \text{誤挿入文字数} - \text{誤削除文字数} - \text{誤置換文字数}}{\text{正解文字数}} \times 100$$

対象とした音源は「国立国会図書館デジタルコレクション「歴史的音源」」[1]内に所蔵されている、「政治の倫理化」(後藤 新平、日本コロムビア、大正 13 年頃)を使用した。

## 3. 結果

平仮名ベースでの結果を図 1、漢字ベースでの結果を図 2 に示す。GHA Denoiser、スペクトルサブトラクション法のいずれにおいても、音声除去処理無しの場合の文字認識率を大きく下回っており、Whisper での音声認識処理においては雑音除去の前処理が有効とは言えない結果となった。

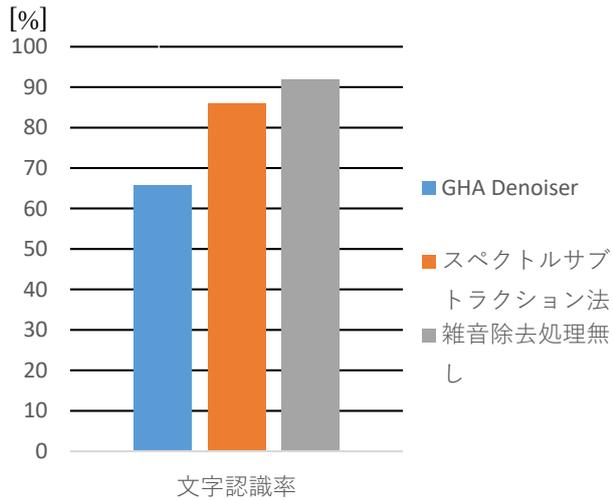


図1 文字認識率(平仮名ベース)の比較

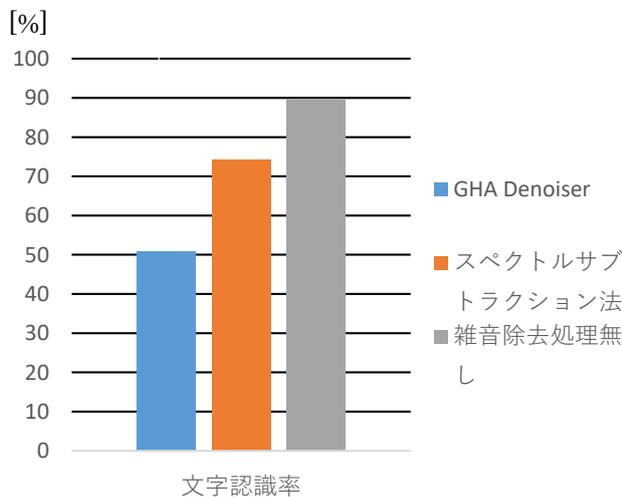


図2 文字認識率(漢字ベース)の比較

#### 4. 結言

音声信号に雑音除去処理を施すことで、大正期の歴史的音源の文字起こしの省力化の可能性を検討した。結果から、Whisperでの音声認識処理において、雑音除去の前処理が悪影響を及ぼすことが確認され、今回の手法が文字起こしの省力化には適さないことが分かった。

#### 今後の予定

今回の結果を踏まえて、引き続き歴史的音源に対する音声認識率の向上に有効な方策を検討していく予定である。

#### 謝辞

本研究では、国立国会図書館のご厚意により、「国会図書館デジタルコレクション歴史的音源」[1]に所蔵の音源を用いております。

また本検討の遂行にあたり、プログラム構築に協力していただいた、サレジオ工業高等専門学校・情報通信工学研究室所属、桶谷 晃誠氏にお礼申し上げます。

#### 文献

- [1] 国立国会図書館デジタルコレクション「歴史的音源」Web サイト  
(<https://rekion.dl.ndl.go.jp/>)
- [2] 三輪賢一郎, 山崎右京, ”音声認識 API 「Whisper」を用いた大正期の歴史的音源の文字起こし,” 電子情報通信学会 2024 年総合大会, D-14-01, March. 2024.
- [3] 音声認識システム Whisper 企業 OpenAI Web サイト  
(<https://openai.com/index/whisper/>)
- [4] GHA Denoiser 一般化調和解析(GHA)を用いた音声ノイズ除去ソフト Web サイト  
(<https://www.vector.co.jp/soft/winnt/art/se510574.html>)