

# ウェアラブル型自動翻訳システムの開発

小山工業高等専門学校 3年 印南諒祐 3年 萩原大貴 3年 関根龍充 2年 飯野雅翔

## 1. 背景・目的

グローバル化により、海外旅行や、日常生活などでも外国語の使用機会が増え、言語の壁が浮き彫りになっている。それに伴い様々な翻訳デバイス・アプリが使用されている。しかし、既存デバイスには

例1：原文の文字をデバイスに打ち込むタイプの翻訳アプリ

- ◆ 複雑な操作が必要
- ◆ 難解な文字等の区別が必要

→ **使える人が限られ  
利便性が悪い**

例2：Google社「Googleレンズ」や、ソースネクスト社「ポケットク」など

- ◆ デバイスを手で操作する
- ◆ 目視で翻訳結果を確認する

→ **「ながら行為」誘発の危険  
安全性が確保されていない**

といった問題がある。

私たちは、現在多くの製品が利用者の**安全性**か**利便性**のどちらかに偏っていると考えた。そこで、

- ①直感的で親しみやすい翻訳の合図
- ②ハンズフリーで、利便性と安全性を両立させる開放的なシステム
- ③デバイスを操作せず観光を楽しむことができ、「ながら行為」の防止へ

という3つの独創的なシステムで**安全性**と**利便性**を両立し、解決することを提案する。そして「言語の壁を無くしていく」ことを目標とし、実証するための新たな翻訳システム「**翻訳展開!!**」の開発を行った。

更には、このシステムを役立てることで、気軽に旅行に行きやすくなり、観光客の増加が期待でき、コロナ収束後の観光業界活性化などにも貢献できると考える。このようにして、私たちは翻訳デバイス業界の革新を起こしていきたい。

## 2. システム概要

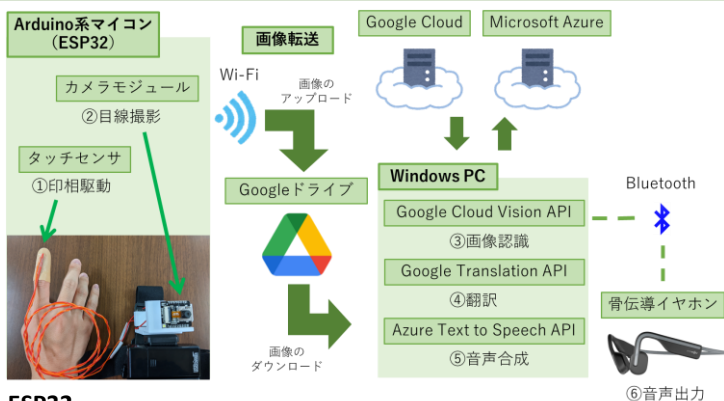


「**翻訳展開!!**」は、専用のヘッドバンドを着用することで、目の前にある翻訳したい文、例えば入力が難しい外国語の看板や案内板、更には本やパンフレットなど、文字列に**目線を合わせ**、**「印を結び」**指に付けたタッチセンサに触れるだけで、翻訳し利用者に音声で伝えることができる。デバイス本体を手に持ったり、カメラを向けたり、文字を入力したり、翻訳結果を確認する、などの従来の翻訳デバイスで必要だった動作が一切省かれ、**「見て」「印を結ぶ」**という動作のみで完結する。

## 3. 機能構成

- ①印相駆動
  - 印を結んだことを検知し、翻訳開始のトリガーとする。
- ②目線撮影
  - 使用者の目線カメラで翻訳対象物を撮影する。
- ③画像認識
  - 光学文字認識 (OCR) で画像から文字列を抽出する。
- ④翻訳機能
  - 抽出された文字列を翻訳する。
- ⑤音声合成
  - 翻訳された文字列を音声に変換する。
- ⑥音声出力
  - 使用者に音声で情報を伝える。

## 4. ソフトウェア・ハードウェア構成



### ESP32

- カメラモジュール、タッチセンサが取り付けられており、指に装着したタッチセンサが印を結ぶことで反応し、カメラモジュールで撮影する。
- 撮影した画像をWi-Fi経由でGoogleドライブへ転送する。
- ESP32はBluetoothを介してもPCへ接続されており、タッチセンサに触れたタイミングでキーボード入力を送信する。
- プログラムはC++で実装。

### Windows PC

- ESP32からのキーボード入力を検知すると、Googleドライブから画像をダウンロードする。
- 文字認識 (OCR)、翻訳、音声合成、音声出力を順次実行する。
- 文字認識はGoogle Cloud Vision API、翻訳はGoogle Translation API、音声合成はAzure Text to Speech APIを用いた。
- 入力言語は180言語に対応し、利用者の母国語に合わせて自由に設定できる。
- プログラムはPython3で実装。

### 骨伝導イヤホン

- 音声を使用者へ伝達する。
- 画面表示の場合視覚を占領してしまう、通常のイヤホンの場合聴覚を占領してしまう、スピーカーの場合周囲への迷惑になる、など既存デバイスで起こる問題を、骨伝導イヤホンを採用することで解決した。

## 5. 成果

### 安全性の確保

- ✓ **ハンズフリー**で、手でデバイスを握る動作を必要としない。
- ✓ 印を結ぶことで翻訳が可能のため、**五感のうち1つでも独占しない**。
- ✓ **「ながら行為」を防止**することができる。

### 利便性の向上

- ✓ 単純な動作で見た文字を自動翻訳でき、国籍、年齢によらず**誰もが簡単に使える**。
- ✓ **素早く翻訳**を行え、旅行などの場面では景色などを優先して楽しめる。
- ✓ **難解な文字等は、事前入力なし**で自動翻訳。

上記の項目を達成することができた。このシステムを実用化することができれば、「1.背景・目的」で述べたような問題解決・社会貢献の実現可能性を見出すことができるであろう。

## 6. 今後の展望

本プロジェクトでは、機能実現のためESP32、Windows PCなど複数のデバイスを使用したり、翻訳開始トリガーに物理センサを用いたりした。そのため、持ち運びや装着に手間が掛かるなどの問題がある。今後の展望では、一連のプロセスを一つのデバイスに完結できるように、計算処理端末のウェアラブル化や、物理センサの代わりに画像処理を用いて印を認識するなど、ユーザーインターフェースの改良をしていき、より実用化に近づけるように開発を進めたい。

また、現状は印を結ぶ動作をしてから翻訳が完了し音声が出力されるまで、約8秒ほど処理に時間が掛かる。実用化の上ではこの処理時間の短縮も目指したい。

※この研究は第32回全国高等専門学校プログラミングコンテストに提出した作品を元に、これを学術的観点にまとめたものである。