



翻訳展開!!  
-ハンズフリー音声翻訳システム-

自由部門  
登録番号：20033

対象者：観光客



# ～開発の背景～

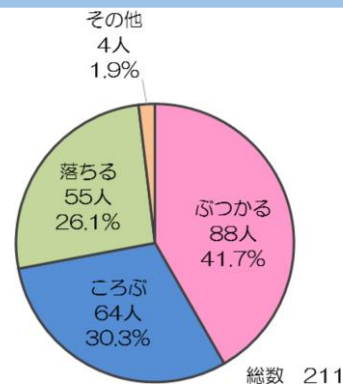
- 近年スマートフォンを使った、道案内や翻訳機能（例：Yahoo! MAP、Google 翻訳）など「仕事、旅行などにおいて欠かせない便利な機能」が広く使用されています
- それに伴い、歩きながらスマートフォンを見ている人（以下、**ながらスマホ**）が増加し、**ぶつかる、ころぶなどの問題が発生しており**（グラフ参照）、解決が難しい社会課題になっています



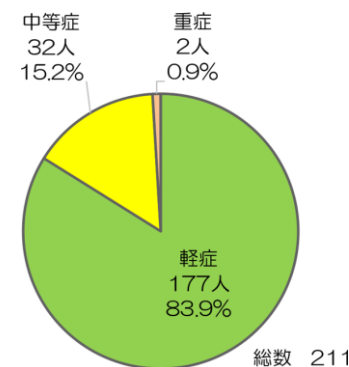
ながらスマホを防ぐには、**ディスプレイを見ないこと**が大切です!!

## ながらスマホによる被害

事故種別ごとの  
救急搬送人員

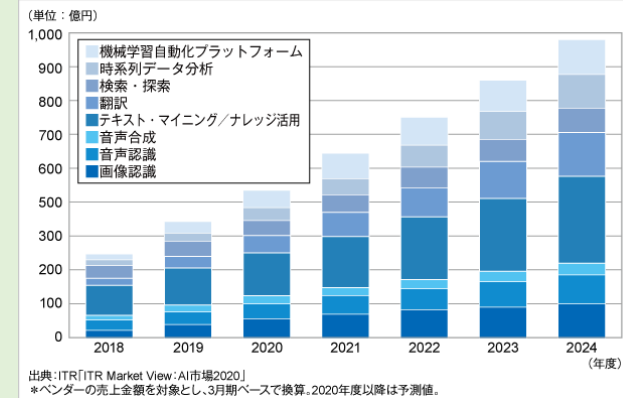


初診時程度別救  
急搬送人員



# 開発の背景・目的

クラウド翻訳アプリは、自分が知らない言語情報を気軽に調べることができ、インターネット環境があれば場所を選ばないことから、年々利用者が増加しています。



## 翻訳アプリ問題点

撮影時



自分で撮らなければいけない

対象者が知りたい情報は、カメラのフレームに収めて撮る必要がある

翻訳実行後



翻訳された文を読まなければならない

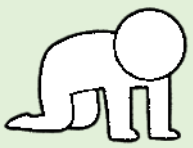
翻訳された情報は、ディスプレイを通し確認する必要がある



翻訳アプリは、**必ずディスプレイを見なければならない**

ながらスマホを引き起こしてしまう

観光客が翻訳の際、ながらスマホを解消し、安全な観光案内をサポートするシステムを創出することが私たちの目的です



# 既存のサービスと関連技術



## 既存のサービスの問題点と課題点

### 携帯型翻訳デバイス

(例：Google翻訳, ソースネクストポケトーク)

- 対象者が知りたい情報を、カメラのフレームに収めながら撮る  
→ **知りたい情報を自動で撮影する機能が必要**
- 文字を打ち込んで調べる場合、難しい字を打ち込むことは容易ではない  
→ **文字認識機能が必要**
- 翻訳結果は、ディスプレイを見て確認する  
→ **視覚ではなく聴覚で伝える**



### ヘッドマウントディスプレイ

(例：Google Glass)

- 目線の焦点が合わず、眼の負担が大きい  
→ **視覚ではなく聴覚で伝える**
- 余分な情報も入る  
→ **条件下でトリガーとなる仕組みが必要**



## 関連技術と問題点

### 道案内アプリ

(例：Yahoo! MAP, NAVITIME)

- 音声ガイドによりディスプレイを見る必要をなくす  
→ **スピーカーの場合、聞き逃してしまう** また、イヤホンの場合、周囲が聞こえない



# 提案

## 課題と解決方法

- 知りたいの情報を自動で撮影する機能が必要
- **対象者の目線に合わせたウェアラブルカメラと、自動で撮影するプログラムを作成**
- 文字認識が必要
- **画像解析機能を使用**
- 条件下でトリガーとなる仕組みが必要
- **ハンズフリーで、誰も簡単に使用できる、安全性と快適性を両立したトリガーを作成**
- 視覚ではなく聴覚で伝える
- スピーカーの場合、聞き逃してしまう また、イヤホンの場合、周囲が聞こえない
- **聞き取りやすく、周囲の音が聞こる、骨伝導イヤホンを使用**

私たちは、ハンズフリー音声翻訳システム  
「翻訳展開!!」で既存の問題を解決します!!

# サービス利用フロー



印を結ぶ



対象者の印が  
撮影開始のサイン。

信号送信

カメラで撮影



対象者と同じ目線の  
カメラが撮影

画像を送信

文字認識・翻訳

音声データ送信

音声読み上げ



骨伝導イヤホンから  
読み上げられる



Lemon

写真から文字列データを抽出

翻訳

音声データに変換



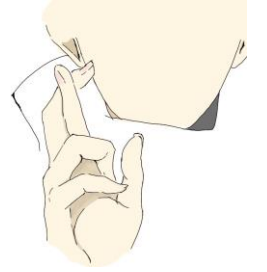
「翻訳展開!!」は、印を結ぶことにより、対象者が知りたい文字を翻訳を行うことで、**観光客の安全性と快適性を提供する**システムです。

# 本システムの特徴

## Mudra trigger system (印相駆動機能)

- 右の絵のような印を結ぶと、翻訳が開始される
- 利点として、音声などをトリガーにすると誤認が起こりやすいが、印を結ぶ場合、誤認することがなくなる
- ハンズフリーになることにより、安全性が格段に上がる

特定の動作なので誤認しにくい!!



## Viewpoint shooting system (目視点撮影機能)

- 対象者の目線に設置し、撮影する
- 利点として、カメラのフレームを気にせず撮影ができるので、誰でも簡単にできる
- ディスプレイを見ないで撮影できる



## Bone conduction system (骨伝導伝達機能)

- 骨伝導イヤホンで、対象者に翻訳結果を伝える
- 利点として、周囲の環境音に関係なく、聞き取りやすい
- 耳をふさがないので、安全性が格段に上がる

音声は視覚より五感割合が低い!!



# システム概要図

画像解析機能

Tesseract OCR

クラウドサービス

翻訳機能

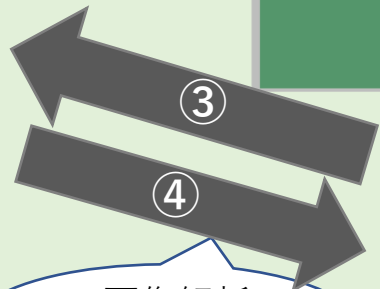


Google翻訳

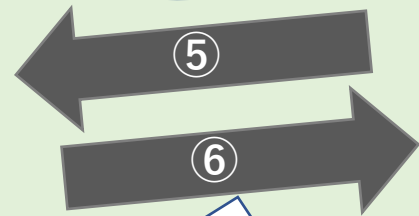
音声合成機能



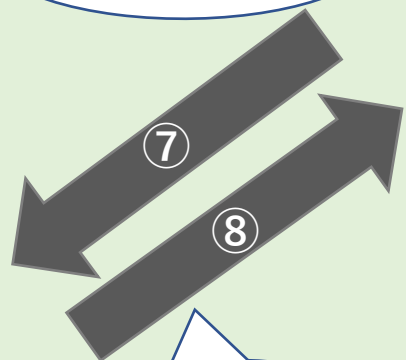
Google  
Text-to-Speech



画像解析  
プログラム (自作)

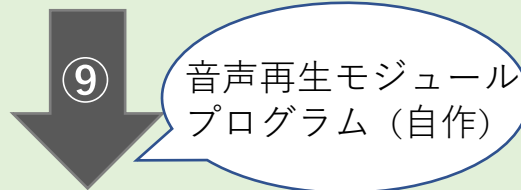


翻訳連携システム  
プログラム (自作)



音声合成変換  
システムプログラム  
(自作)

プログラム実行環境  
Windows10 PC



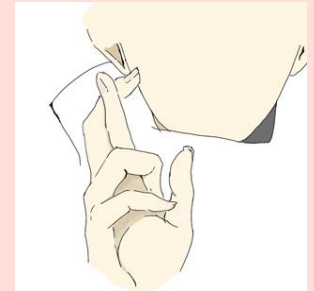
音声再生モジュール  
プログラム (自作)

骨伝導伝達機能

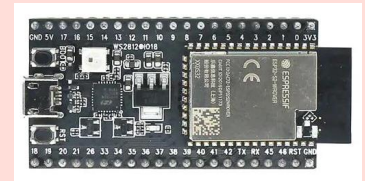


Cheergo製骨伝導イヤホン

印相駆動機能



指装着型接触検知デバイス (自作)



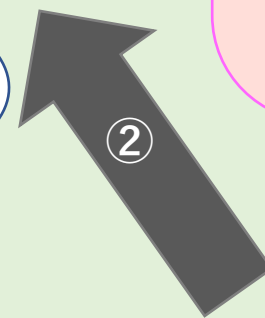
駆動検出プログラム (自作)  
ESP32使用



目視点撮影機能

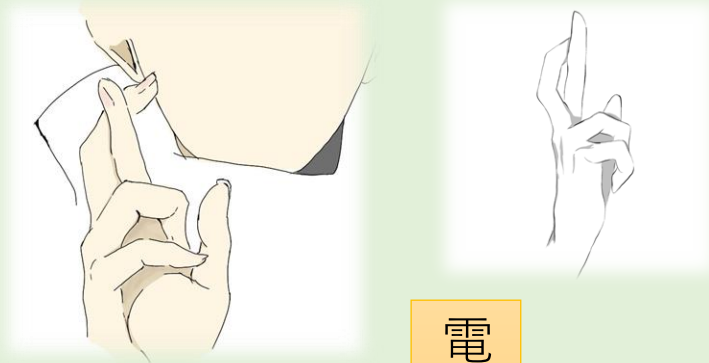


視点固定型カメラデバイス (自作)



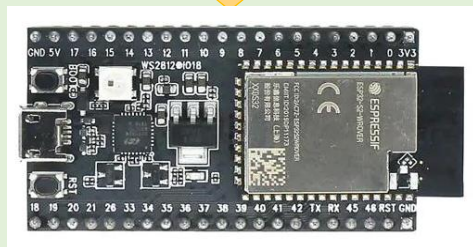


# 印相駆動機能



- 指に電極を付けておく。
- 印を結ぶことで指につけた電極が触れ電流が流れる。

電  
流



ESP32

テキストを入力  
Wi-Fi, Bluetooth

信号送信



## ESP32の利点

- 必要な電圧が他のマイコンより少ないため、バッテリー等の場所を取らない。
- Wi-Fi やBluetoothといった無線通信ができる。

- ESP32で電流を検知し、印が結ばれたか判断。
- 印が結ばれたらカメラに信号を送り、撮影開始。

- C言語で実装
- 開発環境  
Arduino IDE

# システム連携機能



画像ファイル

## 文字の抽出

文字認識 機械学習データをもとに画像を解析し  
文字列に変換  
(Tesseract-OCRを使用)

- Python3.7で実装
- 開発環境  
VScode、Windows10 PC
- 実行環境  
Windows10 PC

“案内”

原文の文字列

## 翻訳

高精度・多言語対応の機械翻訳で任意の言語へ変換  
(Google翻訳API (googletrans)を使用)

## 音声合成

音声合成 機械学習データをもとに音声に変換・出力  
(Google Text-to-Speech APIを使用)

高精度な  
機械学習を  
活用

“information”

翻訳後の文字列

# 既存の技術との比較

	携帯型 翻訳デバイス	ヘッドマウント ディスプレイ	翻訳展開!!
簡易性	×	◎	◎
ながら防止対策	×	○	◎
安全性	△	◎	◎
負担が軽い	△	×	◎
快適性	×	×	◎

## 従来の翻訳デバイス

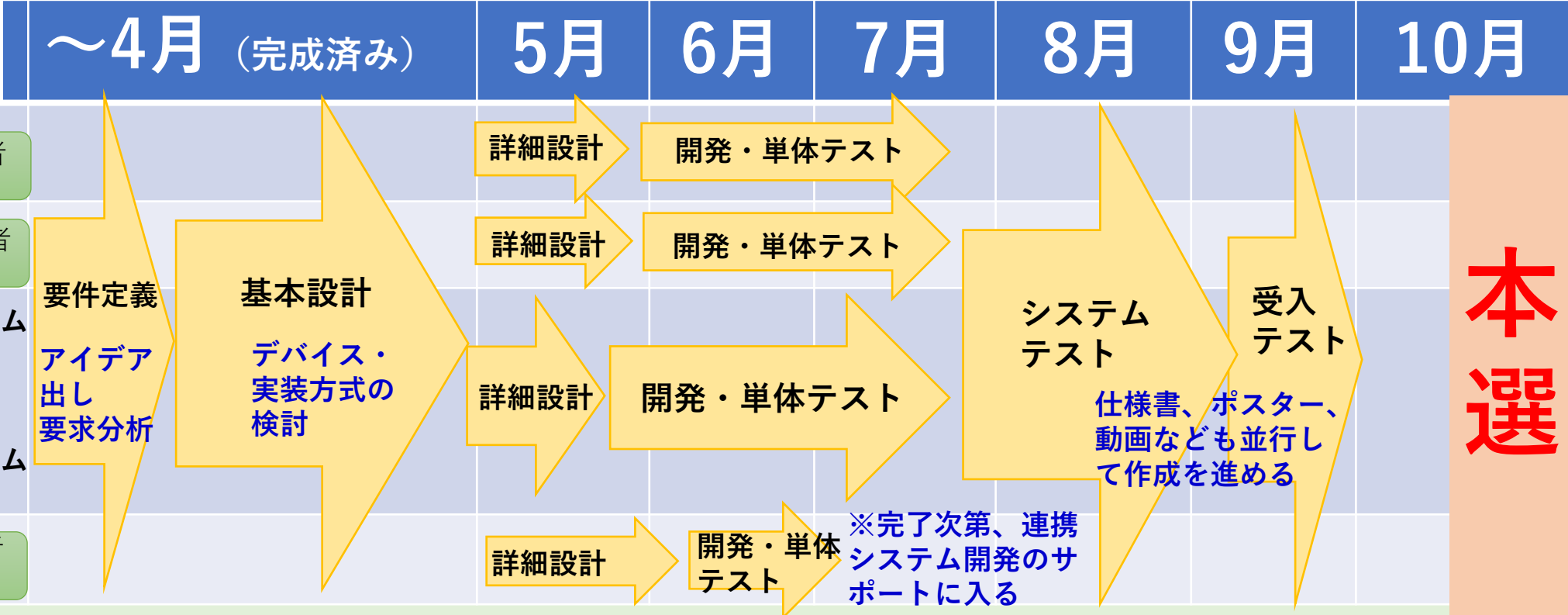
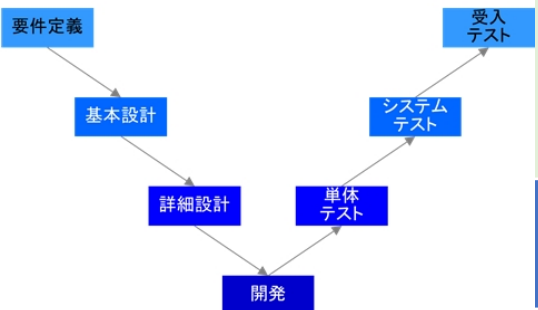
- 翻訳情報を見るときにディスプレイを見る必要があり、ながらになりやすい
- 必要な操作が多く、使用できる人と、できない人に分かれている
- スマホの光を見続けるので、眼の負担が大きい



## 翻訳展開!!

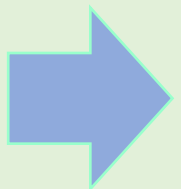
- 周囲の音も聞こえるため、ながら行為を防止できる
- 印を結ぶだけで翻訳されるので、誰でも簡単に使用できる
- 骨伝導なので、鼓膜の負担が軽い

# 開発スケジュール・まとめ



**本選**

- ・ **安心、安全**を第一に考えた飛躍的なシステム
- ・ **事故のない社会**を実現できるシステム
- ・ **不便と感じない**ように常に利用者に**配慮**するシステム



私たちは“**翻訳展開!!**”で  
**観光地の活性化ができる**と  
 考え提案します!!