

# スマートグラスにおける AR を用いた音声認識結果表示手法の提案

## A Proposal for Showing Speech Recognition Result with AR on Smart Glasses

阪田 大輔 †    森田 健太郎 ‡    長田 剛典 ‡    佐藤 健哉 ‡  
Daisuke Sakata   Kentaro Morita   Takenori Osada   Kenya Sato

### 1 はじめに

近年、音声認識の研究が盛んに行われている。現在、この音声認識は身近なものとなり、iOS における Siri や、Android における Google 音声検索で利用できるようになった。

音声認識はコミュニケーション支援技術にも用いられており、認識した音声をテキスト変換することで、聴障者が視覚的に情報を入手できるようになる。既存の支援技術におけるテキスト表示は、パソコンやスマートフォン、タブレットなどの携帯端末の画面に表示するものであるが、使用できる場所が屋内である必要性や手がふさがれてしまうことから、使用場所や使用状況に制限が生じる。

本提案手法では、音声認識結果をスマートグラスに表示したテキストを AR (Augmented Reality) として扱う。スマートグラスの動きの検知にジャイロセンサを用いることで、ハンズフリーで会話内容の確認を可能にし、聴障者が様々な環境で使用できるコミュニケーション支援技術を実現する。

### 2 関連研究

#### 2.1 LiveTalk

LiveTalk[1] は、聴障者参加型コミュニケーションツールであり、会議などの複数人が情報を共有する場で、発話者の発言を音声認識し、テキストに自動変換後、複数のパソコン画面に出力する。音声入力だけでなく、キーボード入力やスタンプ入力、定型文入力が可能である。

#### 2.2 こえとら

こえとら [2] は、音声認識技術や音声合成技術を用いたコミュニケーション支援アプリケーションである。アプリがインストールされたスマートフォンを相手に渡し、音声認識した発言を画面に表示することでコミュニケーションをとることができる。

#### 2.3 既存手法の問題点

上述した支援手法以外にもいくつか手法は存在する。それらの音声認識の表示方法は、パソコンや、スマートフォン・タブレットなどの携帯端末の画面に表示するものである。しかし、デスクトップ型のパソコンであれば屋内に設置する必要があり、使用場所における制限が生じる。また携帯端末の場合では、使用場所の制限を払拭できても、デバイスを持つことで手がふさがれてしまい、会話と並行して他の動作をすることが難しくなり、使用状況の制限が課されるという問題がある。

### 3 提案手法

#### 3.1 概要

図 1 に提案手法の概要を示す。本提案手法では、マイクから得た発話内容をスマートグラスの画面上に、テキストを AR として表示する。AR をスマートグラスの画面に表示することにより、使用場所、使用状況の制限を改善し、聴障者がより様々な環境下でコミュニケーションをとることができるようになる。

また、既存手法の場合、デバイスを用いてスクロールなどの操作をすることで、過去の発言内容を確認することができたが、提案手法ではハンズフリーを実現するために、スマートグラス内蔵のジャイロセンサを用いることで AR を操作し、過去の発言内容の確認を可能とする。

#### 3.2 構成

##### ● スマートグラス

ジャイロセンサが内蔵されており、無線通信可能なものを用いる。ジャイロセンサから、スマートグラスの回転を検知し、回転量・回転方向を計算する。今回の実装ではクラウド上で音声解析を行うため、無線通信が可能である前提とする。

##### ● マイク

スマートグラスに有線あるいは無線接続可能なものを用いる。

#### 3.3 動作手順

1. マイクに向かって発話した内容を、音声データとしてスマートグラスに送信する。
2. 受信した音声データを音声認識を用いて解析し、テキスト変換する。
3. テキストデータを保存し、画面上に表示する。

図 2 に、ジャイロセンサを用いて画面の移動量を取得する手順を示す。ジャイロセンサから取得した角速度  $\omega$  と時間  $t$  から回転角度  $\theta$  を求める。

手順 3：画面に AR としてテキスト表示

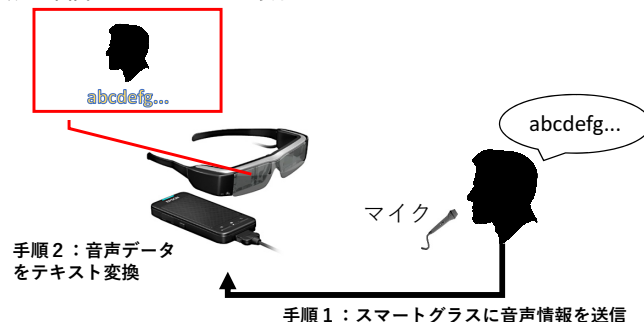


図 1 提案システム概要

† 同志社大学 理工学部 情報システムデザイン学科

‡ 同志社大学大学院 理工学研究科 情報工学専攻

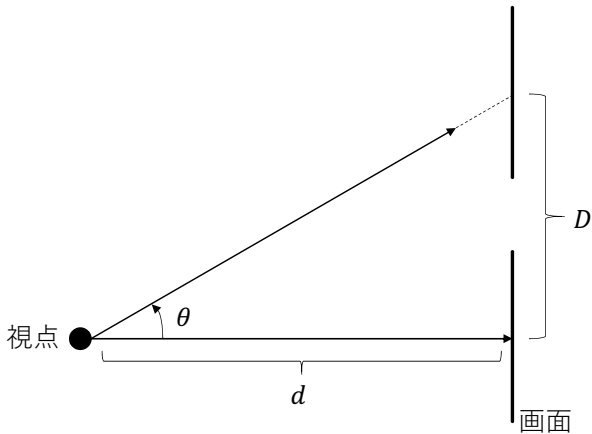


図 2 ジャイロセンサによる画面移動量算出

$$\theta = \omega t$$

視点からスクリーンまでの距離  $r$  として回転したときのスクリーンの画面の移動距離  $D$  は以下の式で表される。

$$D = d \sin \theta$$

算出された  $D$  をもとに、表示すべきテキスト情報を画面に表示することで、会話内容の確認を可能とする。

図 3 に、ジャイロセンサを利用して取得したテキスト情報をスマートグラスに表示する一例を示す。図中の X 軸、Y 軸は、ジャイロセンサの回転軸を示す。首を左右に振った場合、ジャイロセンサが Y 軸中心の回転を感知し、回転量を算出して表示すべきテキスト情報を取得する。また首を上下に振った場合、X 軸中心の回転を感知し、同様の処理を行う。

#### 4 評価環境

使用するスマートグラスは、EPSON 社の MOVERIO BT-200 を用いた。Android Studio で音声認識したテキストを AR 表示するアプリケーション開発を行い、MOVERIO にインストールする。音声認識は Android に実装されているものを用いて、ネットワーク経由でサーバ上で音声解析を行った。使用するマイクは、スマートグラスに無線接続可能な Plantronics 製のモバイル Bluetooth ヘッドセット M70 を用いた。評価は、日常生活の中で頻繁に会話が行われると想定される環境で、聴障者がコミュニケーションをとることが可能であるかを検証した。表 1 に既存手法と提案手法の比較結果を示す。既存手法には、2 章で取り上げた LiveTalk、こえとらを用いた。

表 1 既存手法との比較

	LiveTalk	こえとら	提案手法
使用場所	屋内	屋内外	屋内外
ハンズフリー性	低	低	高
使用中の行動制限	有	有	無

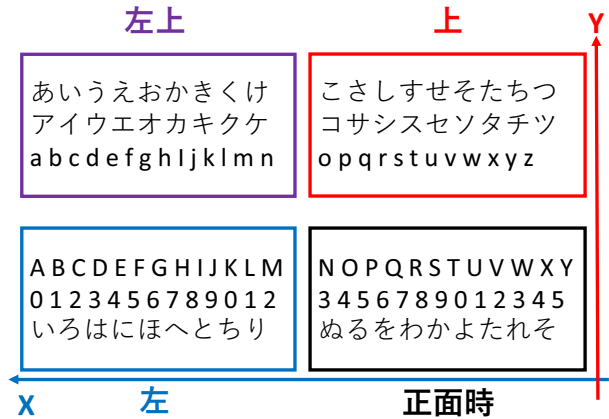


図 3 ジャイロセンサを用いた表示文字取得

#### 5 考察

音声認識した発言内容をスマートグラスの画面に表示することで、聴障者がより様々な場所でのコミュニケーションをとることが可能となる。過去の発言内容を確認する場合、スマートグラスの操作は、コントローラを用いたり、グラス本体を叩いたりなど、従来のように手による操作が必要となる。よってスマートグラスの動きで過去の発言内容を確認できるようにすることで、ハンズフリーを実現でき、他の動作、作業と並行して会話ができるようになる。ウェアラブルデバイスを用い、更にハンズフリーでの操作を可能とすることで、聴障者のコミュニケーションがよりバリアフリーなものとなる。しかし、本提案手法ではスマートグラスをかけている聴障者しか認識結果を視認することができず、認識結果が正確かどうかを発話者本人は知ることができないという問題がある。

#### 6 おわりに

本提案手法では、聴障者向けコミュニケーション支援技術の使用環境をよりバリアフリーにすることを目的として、スマートグラスに発言内容を AR としてテキスト表示する手法を提案した。既存手法では、屋内での使用に限られたり、会話に従事せざるを得ないという制限などがあったが、本提案手法では、ウェアラブルデバイスを用いて、ハンズフリーで操作することで聴障者が限られた場所、場面だけでなく、より自由にコミュニケーションをとることができるようになる。

今後の課題として、音声認識に誤認識があった場合の発話者の確認、訂正方法の実装があげられる。発話者もスマートグラスを装着し、手や指のジェスチャーでスマートグラスを操作することで誤認識の訂正を行うことができると考えられる。

#### 参考文献

- [1] FUJITSU, Software LiveTalk <http://www.fujitsu.com/jp/group/ssl/products/software/applications/ud/livetalk/>, (参照 2016-5-4)。
- [2] FEAT, こえとらアプリサポートページ, <http://www.koetra.jp/>, (参照 2016-5-6)。