

聴覚障害児の授業支援のための HMD による 音声認識結果呈示システムの設計

徳田 浩一 駒谷 和範 尾形 哲也 奥乃 博
京都大学大学院 情報学研究科 知能情報学専攻

1. はじめに

一般校で健聴児とともに学ぶ聴覚障害児（生徒・学生を含む。以下、利用者）の授業参加を促進させるために、要約筆記やノートテイクによる授業参加支援が行われている[1]。最近ではパソコン上でタイピングを行う要約筆記、さらには音声認識による要約筆記の入力の自動化も試みられている[2]。

パソコン要約筆記で作成された字幕は、ノートパソコンのディスプレイか前方スクリーンに表示されることが多い。しかし、この字幕呈示手法では教師や黒板やテレビといった字幕以外の注視対象が字幕の呈示位置と離れてしまうので、視線を頻繁に切り替える必要が生じ、授業への積極的な参加が難しくなるという問題が残っている。

これまで、音声認識を活用した要約筆記の入力の自動化についてはいくつかの研究があるものの、音声認識結果の字幕の出力方法について、利用者の視線の移動負荷を軽減する観点から検討した研究はない。本稿では、光学透過型ヘッドマウントディスプレイ（HMD）の有効性に対する予備実験について報告する。

2. 目標とするシステムの構成

本研究での最終目標は、音声認識技術を用いて字幕を作成し、手で修正したのち、視線移動の少ない装置で字幕を呈示するシステムである。目標システムは図 1 に示したように、音声認識部、認識結果通信部、認識結果修正部、認識結果呈示部から構成される。

2-1. 音声認識部

音声認識は大語彙連続音声認識システム Julius を用いる。音響モデル、言語モデル、単語辞書は認識対象に適合合わせる。Julius はモジュールモードで起動し、音声認識サーバとして利用する（ただし、本稿の予備実験では、字幕呈示機器の差異のみを評価するため、音声認識過程は使用しない）。

2-2. 認識結果通信部

音声認識部で得られた認識結果を認識結果修正部へ送信する。通信プロトコルとして UDP を用いる。ここでは Julius モジュールモードのクライアントツールである jcontrol を利用する。

2-3. 認識結果修正部

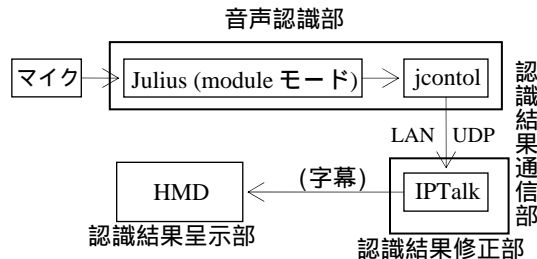


図 1 目標とするシステムの構成図

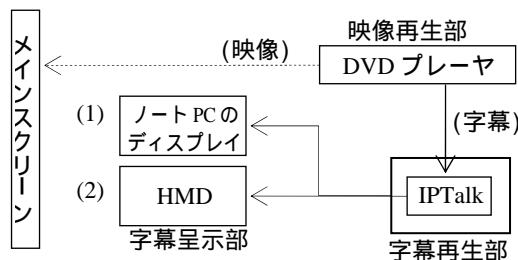


図 2 予備実験のシステム構成図

認識結果を手入により修正・整形する。栗田茂明氏作成のパソコン要約筆記用ソフトウェア IPTalk [3]を活用して音声認識結果に含まれる誤認識を修正し、利用者に呈示する字幕を作成する。IPTalk は LAN で接続されたパソコンを用いて、話者の発言内容を複数の支援者が交代で入力し、協調的に修正した上で利用者に呈示する字幕を作成するためのフリーソフトウェアである。これらの修正作業の支援や自動化も今後の検討を要する。

2-4. 認識結果呈示部

認識結果である字幕を利用者に呈示する。通常のディスプレイを用いた場合、字幕と字幕以外の注視対象の位置が離れてしまい、視線移動が頻繁になるという問題が生じる。視線移動を軽減するために、透過型 HMD を用いて両方の視野を重ねる方式を検討する。予備実験では、光学透過型単眼 HMD の島津製作所製の Data Glass 2/A を用いた[4]。

3. 予備実験

3-1. 実験内容

透過型 HMD の有効性を評価するために予備実験を行った。実験システムは目標システムのサブセットである。その構成を図 2 に示す。

被験者の視線移動の負荷を検証するため、字幕を(1)ノートパソコンのディスプレイと(2)HMD に呈示する。

A Design for Speech Recognition Output System with HMD for Class Support for Hearing Impaired Children
Koichi Tokuda, Kazunori Komatani, Tetsuya Ogata and Hiroshi G. Okuno (Kyoto Univ.)

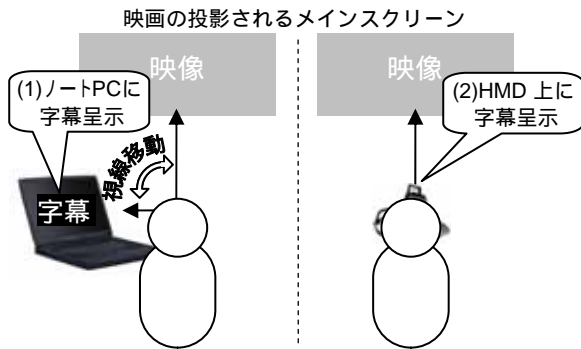


図3 予備実験で対比した2つの状況

(1)(2)各評価機器上の字幕呈示部に意識を向けさせるため、次の(a) (b)の方針で実験を設計した。

- (a) 音声認識の誤認識、遅延の影響の排除：映画を使用し、各評価機器に呈示する字幕は事前に作成しておく。
- (b) 字幕の切り替わりタイミングの視覚による判断を強制：被験者に映画の音声を聞かせない。メインスクリーンには字幕を表示しない。これにより、利用者の使用時の状況を再現する。

3-2. 方法

DVD 映画を音声と字幕なしで実験室前方のメインスクリーンに投影し、字幕のみを次の方法で呈示した。

- (1) 机の上に置いたノート PC のディスプレイ
- (2) HMD

対比した 2 つの状況を図 3 に示す。被験者に前半(1)で、後半(2)で、それぞれ 3 分程度見てもらった。その後感想（選択式と自由記述式を含む）を集めた。

(1)と(2)で呈示する字幕はあらかじめ IPTalk slide editor で作成しておき、出演者の台詞と同期して表示されるように設定した。

被験者は学部学生、大学院生からなる計 8 人で、HMD の使用経験については、全くない（5 人）か、試用経験がわずかにある（3 人）という構成である。

3-3. 結果

それぞれの字幕呈示機器を利用したときの感想として、表 1 のような集計結果を得た。数字は各質問に「はい」と回答した人数である。また、自由記述からは下記のような意見を得た。

- (1) ノートパソコンのディスプレイの場合
 - ・場面の切り替わるテンポが早い場合は見逃すことが多いと思う。
 - ・誰がしゃべっているのか分からなくなることがある。
 - ・視線を移動させるのが疲れる。
 - ・複数の人が会話するとき、発話されたかどうか口に注意しているせいか、かなり字幕を見落とした。
- (2) HMD の場合
 - ・ほとんど通常字幕に近い感覚で見られた。
 - ・画面が明るいと文字が見づらくなる。
 - ・目は疲れるが、映画自体は見やすい。
 - ・字幕を画面のどこに合わせれば見やすいかを試行錯誤

表 1 アンケート結果

	(1)ノート PC	(2)HMD
見逃した字幕がある	3	0
見逃したシーンがある	3	0
目が疲れる	2	3
ピントが合わない	0	4
視線を移すのが大変	4	0
どちらかを見逃す	3	0

しながら見る必要がある。

- ・ピントを合わせるので目が疲れる。
- ・字幕と映像の位置が違うので、ピントを合わせるのが大変。
- ・字が大きいためか、思ったより見やすかった。
- ・HMD にすぐに慣れて見ることができた（予想していたほど違和感はなかった）。
- ・HMD では、装着時の角度によっては画面最上部が切れることがある。

3-4. 考察

HMD の有利な点として、視線の移動が少なくなることが確認できた。実際、表 1 では HMD を利用した際には見逃した字幕やシーンの報告はない。

一方、HMD の欠点としては、目が疲れる、特にピントが合わないという点が明らかになった。欠点については、事前使用なしでの実験の結果であるため、慣れた場合についての評価を長期間にわたって実施する必要がある。

4. まとめ

本稿では、字幕を透過型 HMD に呈示し、字幕と字幕以外の注視対象の視野を重ねることの有効性を評価した。HMD を使用した音声認識結果呈示は、視野移動の負荷軽減の可能性があるものの、目に対する疲労、特に長時間装着の弊害を考慮したシステム運用が必要という知見を得た。

聴覚障害児の授業参加支援システムについては、音声認識に基づいた要約筆記の自動化技術を開発するとともに、利用者の視線移動の負荷を軽減する呈示手法の観点から、さらに検討をすることが必要である。

【謝辞】島津製作所には HMD のデモ機を貸していただきました。心よりお礼を申し上げます。

参考文献

- [1] 白澤麻弓, 徳田克己, 斉藤佐和(監): 聴覚障害学生サポートガイドブック, 日本医療企画 (2002).
- [2] 立入哉, 井上かおり, 宮武由佳: 音声認識を利用した聴覚障害学生学習保障システムについて, 信学技報, ET2003-8, pp.43-48 (2003).
- [3] 栗田茂明: <http://iptalk.hp.infoseek.co.jp/>
- [4] 島津製作所: <http://www.shimadzu.co.jp/hmd/>