

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 18 日現在

機関番号：12103

研究種目：基盤研究(C)

研究期間：2011～2014

課題番号：23501133

研究課題名(和文) タブレット端末を用いた聴覚障害者のための字幕情報保障システム

研究課題名(英文) Captioning support system for hearing impaired using tablet devices

研究代表者

皆川 洋喜 (MINAGAWA, Hiroki)

筑波技術大学・産業技術学部・教授

研究者番号：00273285

交付決定額(研究期間全体)：(直接経費) 4,000,000円

研究成果の概要(和文)：タブレット端末を用いた手書きイメージによる聴覚障害者のための情報保障システムを開発した。会議等で健聴者の発話を音声認識により字幕化しても、多くの誤変換が含まれ、そのまま聴覚障害者が理解することは困難である。そこで、誤変換を含む字幕に対し、他の健聴者がタブレット端末から手書き入力により字幕を修正したり、聴覚障害者が意味が分からない部分を指摘し、これを全参加者で共有することにより、特別な情報保障者なしでも参加者相互が協力しあい、情報保障が可能なシステムを実現した。評価実験により、適切な意思疎通が可能になることが明らかになった。

研究成果の概要(英文)：We developed a captioning support system for hearing impaired using tablet devices. It is difficult for hearing impaired to recognize the caption using voice recognition software, because of its conversion error. Our system can correct the error by hearing people with rewriting image using a tablet device. Hearing impaired can also point out the error which he cannot understand the meaning. The participants can attend the meeting by their collaboration.

研究分野：視聴覚情報保障

キーワード：聴覚障害 情報保障 字幕 要約筆記 音声認識

1. 研究開始当初の背景

筑波技術大学産業技術学部は、聴覚に障害がある学生のみを対象とした大学部局である。本学では、聴覚障害のある学生集団と直接向き合い、手話や板書、視覚教材などの視覚的コミュニケーション手段を用いて直接学生たちと接し、学生たちの学習状態を確認しながら授業を進めている。しかし、大学教育における幅広い知識を身に付けるためには、学外の非常勤講師による授業も必要となる。この場合、非常勤講師は直接聴覚障害学生とコミュニケーションをとることができないため、何らかの情報保障が必要になる。

大学の教育、特に産業技術学部で行っている技術系の教育においては、手話表現が未定義の専門用語が頻出するため、情報保障の方法としてはパソコン要約筆記等の文字によるものが有効である。しかし、大学レベルの高度で専門的な授業内容を要約するためには、要約者に相応の知識が求められることになる。

研究代表者ら筑波技術大学の研究グループでは、このような問題を解決するための1つの方法として、「遠隔情報保障」の研究を進めてきた。これにより、高度で専門的な知識を有する要約筆記者と、情報保障を必要とする聴覚障害者が出席する講義室とをインターネット回線で結び、講義室の講師の音声と映像を遠隔にいる要約筆記者に提示し、パソコン要約筆記により生成された文字を講義室にいる聴覚障害者に提示することが可能になった。このシステムは、筑波技術大学における非常勤講師の字幕授業で利用されている他、筑波大学や日本福祉大学等の一般の大学で学ぶ聴覚障害学生にも利用されている。

このようなシステムはまだ実験段階であり、機材や人件費等にコストがかかるため、現状のまま全国の聾学校や一般大学などに普及させることは困難である。このようなコスト削減と、更なる精度の向上のための研究は、本研究の研究分担者である三好らによって研究が推し進められている。

ところで、字幕は「文字」による情報保障のため、数学や物理学などの授業で頻出する数式や図のイメージを表現することはできない。

数式を用いた講義での情報保障方法については、名古屋大学と連携し、音声認識と画像認識を用いた指示領域抽出の研究を行ってきた。「この」などの指示語を音声認識し、カメラの映像からその時に指し示した領域にある図形や数式などの領域を画像として抽出するもので、要約筆記の文字と合わせて聴覚障害者のための情報保障に利用することができる。

2. 研究の目的

本研究では、健聴者と聴覚障害者が同席する会議場面を想定し、健聴者により発話され

た誤変換を含む字幕に対し、他の健聴者がタブレット端末から手書き入力により字幕を修正したり、聴覚障害者が意味が分からない部分を指摘し、これを全参加者で共有することにより、特別な情報保障者なしでも参加者相互が協力しあい、情報保障が可能なシステムの実現を目指す。

そのため、まず手書きイメージによる情報保障システムのアイデアを実現し、実際に聴覚障害者に評価してもらい、その有効性、可能性、課題について検討する。

システムは複数のタブレット端末により構成される。会議に参加する複数人の聴覚障害者と健聴者に、それぞれ専用のソフトウェアを組み込んだタブレット端末を配布し、これを介して相互に通信をしながら会議を進めていく。

健聴者による発言は、市販の音声認識ソフトウェアを用いてテキスト化され、字幕データとして各タブレット端末に配信される。この字幕データには多くの誤変換が含まれている。

この誤変換が含まれる字幕データに対し、会議参加者である他の健聴者が、適宜手書き入力によりタブレット画面上で必要な修正をイメージデータとして上書きする。上書きされたイメージデータは、ネットワークを介して同期し、他の参加者全てのタブレット端末の画面に表示される。聴覚障害者は元の字幕データに加え、修正されたイメージデータを見ることで、より正確な情報を得ることができる。

また、聴覚障害のある会議参加者が字幕および修正イメージを読んでも意味が理解できない部分は、自分のタブレット端末にその都度書き込むことで、今何が分からないかを発言者はもとより会議参加者全員に知らせることができる。

聴覚障害のある会議参加者からの発言は、キーボード入力により行うが、タブレット端末からの手書き入力により行うことも可能である。

本研究では、このような新しい情報保障システムを実現し、実際に聴覚障害者を対象とした評価を行うことを目的とする。

3. 研究の方法

本研究では、次のような字幕情報保障システムを開発する。

(1) 字幕生成・配信PC

市販の音声認識ソフトウェアにより生成されたテキストデータを、字幕データとして参加者全員のタブレット端末に配信する。音声認識ソフトウェアの設定は不特定多数者とし、個別の学習も行わない。これにより認識率は悪くなるが、システムを簡易化することができる。音声認識の悪さは、会議参加者による手書き入力修正でカバーすることが本研究の主旨である。

(2) 会議参加者用タブレット端末

字幕生成・配信PCより受信した字幕データを、全ての会議参加者（聴覚障害者、健聴者）のタブレット端末で受診する。画面上には、生の字幕データと共に、(3)の専用サーバを介して得られる他者が手書き入力したイメージデータも見ることができる。

発話者自身、および発話者以外の健聴の会議参加者が、誤変換を含む生の字幕に対し、手書きイメージにより適宜修正を加えることができる。

また、聴覚障害の参加者にとって理解不能な部分を、手書きイメージにより指摘することができる。

システムとしては聴覚障害者も健聴者も区別はなく、各会議参加者が入力した手書き入力データは、(3)の専用サーバを介して他の参加者と共有することができる。

データの履歴は端末に蓄積されるので、必要に応じて指先でタブレットを操作し、過去のデータを参照することもできる。

(3) 専用サーバ

字幕データと、各会議参加者により入力された手書きイメージを同期させ、全ての会議参加者に配信する。

システムの評価方法としては、実際に聴覚に障害のある被験者にシステムを使用してもらい、評価を得る。

アンケート等による主観評価の他、機器を操作する様子や講義に参加する様子などの行動もビデオに記録し、分析する。

聴覚障害者たちの手話による発言、動作、視線、他の聴覚障害者との協調などの様子を記録し、分析する。

評価実験の結果に基づいて、システムの改良を加えていく。システムに要求される機能やユーザインタフェースは、会議内容や会議形態、ユーザの好み等により多種多様になることが想定される。個々の条件やニーズに合わせてカスタマイズ可能な自由度のある設計が求められる。

4. 研究成果

(1) プロトタイプシステムの開発

前述のアイデアをプロトタイプシステムとして実現した。Androidをベースに、専用サーバソフトウェアおよび各タブレット端末用のクライアントソフトウェアを開発し、これと市販の音声認識ソフトウェアと組み合わせることでシステムを構成した。

本プロトタイプシステムにおいては、本研究で課題としている基本的な動作を実現することを念頭に置き、操作のしやすさや画面の見易さ等については最低限の機能を実現した。

(2) 国際会議での報告・情報収集

聴覚障害者が同席する講義および会議の場面を想定した字幕情報保障システムについて、オーストリアのリンツにおいて開催された国際会議 ICCHP(International Conference on Computers Helping People with Special Needs)で研究発表を行なった。ここでは、研究の背景、システムの作成及び作成者らによる試用と問題点について報告した。また、Universal Learning Designの発表セッションにおいて、各国の聴覚障害者らが多数参加する会場で、聴覚障害者自身の立場からの意見や、各国事情の違いなど、様々な有用な意見交換をすることができた。

さらに、チェコ共和国ブルノにおいて開催された Universal Learning Design International Conference Brno 2013に参加し、情報収集を行なった。この国際会議には各国の聴覚障害者が多数参加し、ここでの情報保障には Polygraf という字幕配信システムが採用されており、その実際の動作を体験し、直接説明を受けることができた。Polygraf は言語間通訳や文字・手話通訳など、非常に多くの人的パワーを必要とする点で我々の提案するより手軽なシステムとは思いが大きく異なるが、誰でも自由に使える点、クライアントはフリーで配布している点、プレゼンテーションファイルと字幕を同時に配信している点など、弱視の聴講者にとっても、また外国人の参加者にとっても有用であり、今後のシステム改良の方向性として大いに参考になった。

(3) システムの改良

(1)のプロトタイプシステムを元に、(2)等で得られた知見を加えてシステムを改良した。主に画面構成等のインタフェースを改良し、手書き入力のしにくさを改良した。

また、今後の評価実験をふまえ、必要なプロトコルデータの記録が行えるように改良を加えた。

(4) 聴覚障害者による評価実験

(3)で改良を加えたシステムを用いて、実際に聴覚に障害のある被験者によりシステムの評価を行った。その結果、以下のような点が明らかとなった。

情報保障システムとしての機能

音声認識によるテキストデータを字幕として全会議参加者に配信し、各タブレット端末画面に表示させることができた。また、各タブレット端末から手書きイメージを上書き入力し、他の会議参加者と共有するという、最低限の機能が実現できた。

聴覚障害者自身が情報保障の役割の一部を担うという、これまでになかった発想の情報保障システムに対し、聴覚障害者らは、戸惑いながらもシステムを操作することができ、健聴者と相互に協力しながら会議を進めることができた。

音声認識の精度

本システムは、音声認識の精度が低くてもそれを参加者全員で補うという発想に基づくものである。しかし、誤認識が多すぎる場合は、健聴者による修正入力が間に合わなかったり、聴覚障害者が指摘するにも手がかりが少なすぎて指摘すらできない事態となる。

本システムを用いる場合、どの程度の精度の音声認識であれば対応可能であるか、様々な条件で詳しく調べる必要がある。

文字サイズと行間

本システムでは、手書き入力のために字幕の行間をダブルスペースで空けて提示し、ここに手書きイメージを書き込むようにしてある。また、手書き修正を加えるために、文字サイズも(設定可能であるが)大きめに表示させている。

ところが、普段の字幕表示に慣れている聴覚障害者にとっては、一画面に表示される文字数が少なくなり、字幕として情報を得ることが難しくなってしまう。

これについては、画面構成の見直しや、機器の大画面化等の検討が必要である。

字幕履歴と画面スクロール

本システムには、自分の端末で自由に字幕の履歴をさかのぼって見ることができる機能がある。これ自体は便利であるが、履歴を見ている時に画面修正が入ると、画面が最新の字幕に戻ってしまう。

また、修正している間に新しい字幕データが入力されると、修正イメージが履歴をさかのぼらないとみることができない位置に挿入され、他の会議参加者に気づいてもらえないという状況も起きる。

自分のアクションを他者にいかに伝えるか等については、さらに検討が必要である。

リアルタイム性

現状のシステム仕様では、手書き入力データを書き終えてから、確定ボタンにより他者に配信するようになっている。従って、評価実験の際には何人かによる手書きイメージが重なって画面に出力されてしまうということが起きた。

システムの仕様を見直し、リアルタイムで「誰かが何かを書いている」ことが全員で共有できれば、重ね書きや修正作業の重複を避け、より円滑なコミュニケーションが可能なシステムとなると思われる。

専用入力スタンプ等

「何？」など、よく使う言葉を記号化した専用入力スタンプや、作業内容や入力者毎の色分けなど、実際の使用状況に特化したメニュー構成についての要求があった。

このことについてはさらに詳しく調べる必要がある。

聴覚障害者からの発言

現状では、聴覚障害者からの発言は、キーボード入力と手書きイメージ入力を想定している。しかし、聴覚障害者にも音声認識を使用したいというニーズを持っている者がおり、様々なニーズに対応したインタフェースを検討する必要がある。

本研究により、以上のような成果と、今後の発展につながるいくつかの課題を見出すことができた。

5. 主な発表論文等

〔雑誌論文〕(計1件)

Makoto Kobayashi, Hiroki Minagawa, Tomoyuki Nishioka, Shigeki Miyoshi, Meeting Support System for the Person with Hearing Impairment Using Tablet Devices and Speech Recognition, Computers Helping People with Special Needs(2), Springer, LNKS7383, 2012, pp.221-224 (査読有)

〔学会発表〕(計1件)

Makoto Kobayashi, Hiroki Minagawa, Tomoyuki Nishioka, Shigeki Miyoshi, Meeting Support System for the Person with Hearing Impairment Using Tablet Devices and Speech Recognition, 14th International Conference on Computers Helping People with Special Needs, 2012年07月12日, Linz, Austria

6. 研究組織

(1)研究代表者

皆川 洋喜(MINAGAWA, Hiroki)
筑波技術大学・産業技術学部・教授
研究者番号: 00273285

(2)研究分担者

西岡 知之(NISHIOKA, Tomoyuki)
筑波技術大学・産業技術学部・教授
研究者番号: 70310191

三好 茂樹(MIYOSHI, Shigeki)
筑波技術大学・障害者高等教育研究支援センター・准教授
研究者番号: 80310192

小林 真(KOBAYASHI, Makoto)
筑波技術大学・保健科学部・准教授
研究者番号: 60291853