

視覚障害者用学術情報インタフェースに関する研究

Research on man-machine interface for academic information to be used by the severely visually handicapped

- | | | |
|-------|--|------|
| 代表研究者 | 東京理科大学基礎工学部教授
Prof., Faculty of Industrial Science and Tech., Science Univ. of Tokyo
Hiroya FUJISAKI | 藤崎博也 |
| 協同研究者 | 東京大学工学部助教授
Assoc. Prof., Faculty of Engineering., The Univ. of Tokyo
Keikichi HIROSE | 広瀬啓吉 |
| | 東京大学医学部教授
Prof., Faculty of Medicine, The Univ. of Tokyo
Shigeru KIRITANI | 桐谷滋 |
| | 早稲田大学理工学部教授
Prof., School of Science and Engineering., Waseda Univ.
Katsuhiko SHIRAI | 白井克彦 |
| | 早稲田大学人間科学部教授
Prof., School of Human Sciences, Waseda Univ.
Shizuo HIKI | 比企静雄 |
| | 千葉工業大学工学部教授
Prof., Faculty of Engineering, Chiba Inst. of Tech.
Ken-iti KIDO | 城戸健一 |
| | 中京大学情報科学部教授
Prof., School of Computer and Cognitive Sciences, Chukyo Univ.
Minoru SHIGENAGA | 重永實 |
| | 聖心女子大学文学部教授
Prof., Faculty of Liberal Arts, Univ. of the Sacred Heart
Shuukou TORII | 鳥居修晃 |
| | 特殊教育総合研究所部長
Director, National Inst. for Special Education
Yasuhiro KIZUKA | 木塚泰弘 |
| | 特殊教育総合研究所所員
Researcher, National Inst. for Special Education
Kouichi ODA | 小田浩一 |
| | ATR 翻訳通信研究所室長
Chief, ATR Interpreting Telecommunications Res. Lab.
Norio HIGUCHI | 樋口宜男 |

Based on the analysis of the problems and the difficulties encountered by the severely visually handicapped people in accessing as well as in generating academic information, the present study aims at adapting current technologies of speech and natural language processing, which are generally developed with the tacit assumption that the user is not visually impaired, in order to provide better means of academic activities for the blind. Both the source of information and the intended output of the blind person are assumed to be academic documents readable by ordinary people, with speech and Braille symbols as the direct channels. The use of ordinary keyboard is also considered. Thus the following five tasks are examined assuming a blind person as the user.

1. Conversion of academic text into speech.
2. Conversion of academic text into Braille symbols.
3. Conversion of Braille symbols into academic text.
4. Conversion of speech into academic text.
5. Use of ordinary keyboard to generate academic text.

Whereas all these input and output modes may be necessary for the blind depending on the situations, the emphasis in the current study is laid on the use of speech both for input and for output, since it provides the easiest and the fastest means of communication for the blind. Several new techniques have been developed in the present study both in speech synthesis from text and in speech recognition, that will considerably facilitate information transfer when used by a blind person.

研究目的

重度の視覚障害者（盲人）は、知能が優れていても、その障害のゆえに情報の入出力が困難である。特に現行の情報の表現形式および情報入出力のための補助機器は、初歩的・基本的な情報の入出力を目的としたものであり、高度の内容をもつ学術情報の入出力に適した表現形式は確立されておらず、また補助機器も不備である。このことは重度視覚障害者、特に盲人が大学を受験する際の問題の提示および答案の評価に大きな支障となっており、現在は晴眼者による点訳および墨訳に頼っているが、精度・速度の点で問題がある。また、重度視覚障害者が大学・大学院レベルで高度の学術情報を検索・利用したり表現したりするには、さらに大きな支障がある。本研究は、このような重度視覚障害者にとっての学術情報の入出力における問題点を徹底的に検討することによって、現行の学術情報の表現形式、および現存する補助機器の不備な点を明らかにするとともに、前者に関しては統一的な見地から改良を加え、また後者に関しては晴眼者の利用を前提として研究され、開発されている視聴覚情報処理技術を拡張・発展させることによって、補助機器とその利用技術の改善を図ることを目的として行ったものであ

る。

研究経過

本研究でいう学術情報とは、通常晴眼者が読むことのできる文書を指す。一方、視覚障害者への情報の提示手段としては音声または点字、視覚障害者からの情報の表現手段としては通常のキーボード、点字キーボード、または音声とする。したがって視覚障害者を中心とする情報の流れは、図1に示すように表現することができる。

研究の経過としては、(A) 視覚障害者への情報の提示手段、(B) 視覚障害者からの情報の表現手段、の両面から検討を行った。以下にその詳細を示す。

A 視覚障害者への情報の提示手段

A-1 漢字・かな・その他の記号による表記の音声への変換

まず、通常の漢字・かな混じりのテキストから音声合成するシステムを開発し、次にローマ字・記号・数式などを含む漢字かな混じりテキストを音声に変換する際に必要な特殊処理の手法を開発する。ここで特殊処理とは、漢字・かな以外の記号などを発音するための処理である。

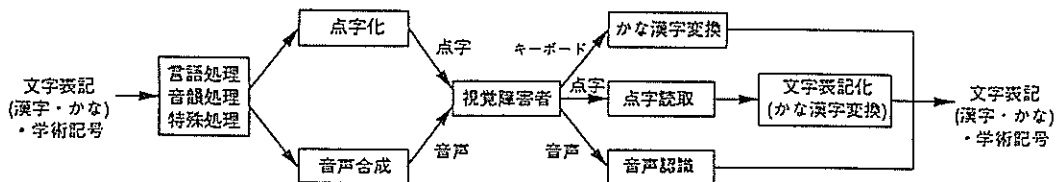


図1. 重度視覚障害者にとっての学術情報入出力の主要な手段

A-2 漢字・かな・その他の記号による表記の点字への変換

ローマ字・記号・数式などを含む漢字・かな混じりテキストを点字に変換する際に必要な特殊処理の手法を開発する。これは A-1 の場合と共通する点も少なくないが、点字ではアクセント・イントネーションなどを付与できないため、同音異義語をはじめとする各種の曖昧性の解消に留意する必要があること、また特殊記号の点字化に関しては従来試みられているが、学術情報の表記として統一がとれていないことが問題である。

B 視覚障害者からの情報の表現手段

B-1 点字表記から漢字・かな・その他の記号による表記への変換

視覚障害者が点字器を用いて作成した点字文書を、晴眼者が容易に理解できる通常の漢字かな混じり文書（ローマ字・記号・数式などを含む）に変換する際に必要な言語処理の手法を開発する。

B-2 視覚障害者のキーボード入力による通常の学術文書作成支援

視覚障害者が直接に通常のキーボードを介して晴眼者向けの学術文書を作成するのを支援するための手法を開発する。この場合には、通常のワードプロセッサと同様、若干の試行錯誤を許容するが、視覚障害者にとってできるだけ負担が少なく、しかも文書作成能率の良い手法の開発を目的とする。

B-3 視覚障害者の音声入力による通常の学術文書作成支援

多くの場合、視覚障害者は音声には全く障害がないので、音声を用いて晴眼者向けの学術文書を作成するのを支援する手法を開発する。全く誤りのない音声自動認識は現在の技術水準では達成されていないので、この場合にも上記 B-2 と同様、

若干の試行錯誤を許容するが、システムを特定の話者に適応させることによって認識率を高め、使用者の負担を軽減することを主眼とする。

研究成果

A 視覚障害者への情報の提示手段

A-1 漢字・かな・その他の記号による表記の音声への変換

まず、通常の漢字かな混じりのテキストから、音声を合成するシステムを開発した。このシステムは、(a) 言語処理、(b) 音韻処理、(c) パラメータ生成、(d) 音声波形生成、の4段階からなる。それぞれの内容は以下のとおりである。

(a) 言語処理：まず、形態素解析を行ってテキストを単語単位に区分する。形態素解析は、辞書の照合によるテキストの区分、形態素間の接続関係の妥当性チェックによる区分の修正、の2段階からなる。次に、統語・意味・談話解析によって統語構造および形態素の重要度などに関する情報を抽出する。

(b) 音韻処理：まず、隣接する形態素の読みに関する連濁、促音化などの規則を用いてテキストを構成する単語を音素記号列に変換し、次に異音化規則を用いて音素記号列を実際の発音と直接に対応する単音記号列に変換する。一方、アクセント結合規則を用いて複合語のアクセント型を決定し、また、隣接する単語間にもアクセント結合規則を適用してテキストを韻律語の系列に変換する。さらに、統語境界の深さと形態素の重要度の指標から、韻律記号生成規則を用いて休止記号・フレーズ記号・アクセント記号からなる韻律記号を生成し、上記の単音記号列中の該当する位置に挿入する。

(c) パラメータ生成：単音記号に対応した音節蓄積パターンを呼び出し、接続規則を適用して接

続し、声道伝達特性および音源強度・有声/無声パラメータの時間パターンを生成するとともに、休止記号から決まる休止時間を挿入し、また、フレーズ記号・アクセント記号をそれぞれフレーズ指令・アクセント指令の生起時点と振幅に変換して基本周波数パターン生成過程のモデルを駆動し、基本周波数の時間パターンを生成する。

- (d) 音声波形生成：基本周波数パターンおよび音源強度・有声/無声パラメータから音源波形を生成して、声道伝達特性パラメータにより制御したフィルタ回路を駆動し、音声波形を合成する。

次に、このシステムを漢字・かな以外のローマ字・記号・数式などを含む学術文書に適用するための処理として、(e) 特殊処理の機能を設けた。以下にその概要を述べる。

(e) 特殊処理

- 1) 漢字かな混じり文中に存在する他の記号の処理

- a. 欧文の語・句・固有名詞など：原語の発音に準ずる
- b. ローマ字・ギリシャ文字などの記号：アルファベットとして発音するか、もとの単語として発音するかを文脈から判断する(例：mを“エム”と読むか“メートル”と読むか)。この場合、厳密な正確さを求めるならば高度の文脈処理が必要となるが、直前および直後の単語の品詞を参照することによって、かなりの程度まで正確な判断が可能である。
- c. 上記以外の記号：慣用に従って発音する

- 2) 漢字かな混じり文外に存在する他の記号の処理

- a. 数式・化学式など：それぞれの分野の慣用に従って発音する
- b. その他(例えば\$や@、数式の後の式番号など、かならずしも読み方が決まっていないもの)：分かりやすい読みを付ける

- 3) 未登録記号列の処理

上記 1), 2) に挙げた記号などが閉じた集合を

なす場合にはリストとして登録すれば十分であるが、言語は閉じた集合ではないので常に未登録、つまりシステムにとって未知のものが出現する可能性がある(ただしここでは個々の記号は既知であるとする)。したがってこの場合のシステムの対応として、以下の3通りのモードを設けた。

- a. 個々の記号ごとに発音する
- b. 規則によって推定して発音する
- c. ユーザ(オペレータ)の介入を求める

これらのモードの中で、当然ながら発音の正確さはa, b, cの順に高くなるが処理時間もこの順に長くなる。したがってaは正確さをある程度犠牲にしても実時間の動作を必要とする場合に、bは大量の学術文書をバッチ処理で音声に変換する場合に、またcは正しい知識を持った人間と同じ正確さの発音を必要とする場合に適している。

このうち、bでは未登録記号列の発音の推定が必要となる。これは特殊な記号の場合だけではなく、漢字かな混じり表記の未登録語の場合にも必要な処理であり、帰納的推論に基づく推定法を検討した。

A-2 漢字・かな・その他の記号による表記の点字への変換

点字はかなと対応しているため、この問題はA-1の音声への変換と共通する点も少なくないが、点字ではアクセントやイントネーションなどの韻律的特徴を与えられないので、同音異義語や構文的多義文などによる各種の曖昧性の除去ないし軽減に、特に留意する必要がある。ことに漢字の熟語には同音異義語が多いため、点字の場合には読み替え、言い替えによって区別する必要がある。しかしながら、大部分の曖昧性は文脈中では解消するため、特定の文脈の中で誤解を生ずるおそれのあるものだけを区別するのが最も合理的である。

この見地から、話題(専門分野)別に同音異義語を調査し、それらを区別するための読み替え、言い替えを検討した。この場合にも、対象を例えば高校教育の範囲に限れば網羅的な対処が可能で

あるが、より程度の高い学術文書を対象とする場合には網羅的な方法で対処することはできず、高度の文脈処理が必要になる。

B 視覚障害者からの情報の表現手段

B-1 点字表記から漢字・かな・その他の記号による表記への変換

点字は、原則としてかなと対応するので、点字表記から通常の漢字・かな・記号による表記への変換は、かな・漢字変換の問題と共通するところが多い。この場合の最も大きな問題は、同音意義による曖昧性、すなわち、一つの点字表記が複数の異なる墨字表記に対応することである。一般に点字表記では、文節に近い単位での分かち書きが行われるので、べた書きの場合よりは形態素解析が容易である反面、点字に特有の表記（例えば助詞「は」・「へ」を「わ」・「え」と表記）による問題も生ずる。この場合の曖昧性は、形態素レベルでの同音意義（例：意思/意志/石・・・）と、文節内構文の多義（例：お国/奥に）に分けられる。

これらの曖昧性に対処して点字・墨字変換を行うに当たって、本研究では以下の三つの場合を検討した。

1) 完全自動変換による単一表記の出力

上記の曖昧性を完全に近い形で解消するには、単に文節内の言語処理だけでは不十分で、隣接の文節、文全体、さらには前後の文までを含む文脈の情報を利用した言語理解を行わねばならない。しかしながら、これらの処理を完全に行うことは、現在の自然言語処理技術では達成されていないので、ここでは文節内および後続文節との構文的整合性を利用した曖昧性の除去について検討した。具体的には、候補として得られた形態素列中の隣接する形態素の間の関係が構文的に許容し得るもののみを選択する。点字化された学術文書を対象として、上記の処理を行った結果、曖昧性の約 75% を解消し得ることを確認した。

2) 残存する曖昧性への対処

上記の処理によっても除去し得ない曖昧な箇所について、以下の二つの場合に分けて検討し

た。

- a. 晴眼者が介入して妥当な表記を選択する場合：通常のワードプロセッサの場合とは異なり、この場合の晴眼者は文書の著者ではないので、曖昧な箇所だけを出力するのでは不十分で、前後の文脈を参照することが必要となる。実験の結果、この場合の文脈としては 1 文ではほぼ十分であることを確かめた。
- b. 複数の表記の併記を許容する場合：曖昧な箇所について、異なる表記のすべてを表記して読者に直接判断させるのが最も厳密であるが、選択枝の数が余りに多い場合には非現実的になるので、形態素の出現頻度を考慮して高順位のものから 3 位までを併記する方法を検討した。これによりカバーし得ない曖昧性は約 1% 以下であることを確かめた。

B-2 視覚障害者のキーボード入力による通常の学術文書作成支援

これに関してはすでに各処で研究開発が行われ、複数のシステムが市販されている。したがってここでは、重度視覚障害者にとって使いやすく、最も多くのユーザをもつといわれる『AOK 点字ワープロ』をとりあげ、これを用いて視覚障害者が学術文書を作成する場合の問題点、および改良のために付加すべき機能について検討した。

B-3 視覚障害者の音声入力による通常の学術文書作成支援

晴眼者を対象とした音声認識・ワードプロセッシングと異なる点は、視覚障害者にとっての負担を軽減すること、具体的にはできるだけ少ない回数の発声で意図した墨字表記が作成できるようにすることである。このため以下の 3 点を検討した。

1) 有効な話者適応方法の検討

代表研究者らがすでに開発した話者適応の方法をさらに発展させ、内容が既知の訓練用の発声に基づき、個人の音素の特徴を能率よく学習することによって話者適応を行う方式を開発した。

2) 韻律を用いた同音異義語の識別

代表研究者らがすでに開発した音声の基本周波数パターンの特徴抽出手法を用いて、単語のアクセント型を認識し、同音異義語による曖昧性を解消する方式を開発した。

3) 合成音声による認識結果の確認

すでに作成した音声合成システムを用いて、文節ごとの認識結果から音声を再合成し、文書作成者にフィードバックする方式を検討した。なお、この手法はキーボード入力による文書作成時にも適用が可能であるが、A-1の項で述べた音声合成と異なる点は、句読点などの文字言語固有の記号をも発音して確認することである。

今後の課題と発展

以上、重度視覚障害者のための学術情報のインターフェイスの検討を行い、視覚障害者にとっての情報の入力と出力の機能を有するトータルシステムを念頭に置いて研究を進め、システムの重要な構成要素に関する技術について成果を挙げた。しかしながら、晴眼者にとってのワードプロセッサと同様、これらの要素技術は完全に無誤謬ではありえず、若干の試行錯誤を許容することは避けられない。むしろ重要なのは、視覚障害者がそのような試行錯誤を容易に、負担が少なく行えるような技術を開発することである。大部分の視覚障害者に残された機能のうち、情報の入出力に特に有利なのは明らかに音声であり、したがって本研究においてもこれに重点を置いた。重度視覚障害者の大学受験に際しては、現在、音声の使用が制限される場合もあるが、視覚障害者にとって音声の使用を可能とするような環境を整えることが重要であろう。

なお、本研究ではこれらの要素技術を視覚障害者による実際の使用場面で評価するまでには到らなかったが、これらの技術による障害者の負担の軽減を具体的、定量的に評価し、詳細な点での最適化を図ることは今後の課題である。

発表論文リスト

- 1) H. Fujisaki, K. Hirose and Y. Asano: Proposal and Evaluation of a New Type of Terminal Analog Speech Synthesizer, *Proc. 1990 Int'l Conf. on Spoken Language Processing*. Vol. 1, pp. 329-332 (1990. 11).
- 2) 坂部健一: 知識獲得に基づく意味のカテゴリー生成と未知語の意味推定, 東京理科大学基礎工学部電子応用工学科平成3年度卒業論文(指導教授藤崎博也)(1992. 3).
- 3) 亀田弘之, 藤崎博也: 日本語複合語の意味推定における文脈の役割, 電子情報通信学会『言語と知識の獲得』研究会資料, L K 92-3 (1992. 4).
- 4) 遠藤潤一: 品詞接続情報の定量化にもとづくかな文字文の形態素解析と点字・墨字変換への応用, 東京理科大学基礎工学部電子応用工学科平成4年度卒業論文(指導教授藤崎博也)(1993. 3).
- 5) 黒松哲也: 辞書情報・品詞接続情報・認知単位を用いた漢字かな交じり文の形態素解析, 東京理科大学基礎工学部電子応用工学科平成4年度卒業論文(指導教授藤崎博也)(1993. 3).
- 6) S. Ohno, K. Hirose and H. Fujisaki: Utterance Normalization Using Vowel Features in a Spoken Word Recognition System for Multiple Speakers, *Proc. 1993 Int'l Conf. on Acoustics, Speech, and Signal Processing*, Vol. 2, pp. 578-581 (1993. 4).
- 7) 広瀬啓吉, 藤崎博也: 視覚障害者用学術情報インターフェイスに関する研究, 東京大学工学部総合試験所年報, Vol. 52, No. 12, pp. 85-88 (1993. 9).
- 8) H. Fujisaki, K. Hirose and N. Takahashi: Manifestation of Linguistic Information in the Voice Fundamental Frequency Contours of Spoken Japanese, *Trans. Inst. Electronics, Information and Communication Engineers*, E76-A(11), 1919-1926 (1993. 11).
- 9) K. Hirose and H. Fujisaki: A System for the Synthesis of High-Quality Speech from Texts, *Trans. Inst. Electronics, Information and Communication Engineers*, E76-A(11), 1971-1980 (1993. 11).
- 10) 大野澄雄, 広瀬啓吉, 藤崎博也: 多数話者単語音声認識における母音の特徴を利用した入力音声正規化, 電子情報通信学会論文誌, J76-A(12) (1993. 12).

(1, 3, 6~10 は公表された学術論文 2, 4, 5 は未公表の卒業論文)