

特集

「RWC—実世界知能」

実世界における手話認識技術

Recognition Technology of Sign Language in the Real World

竹内 勝
Masaru Takeuchi

技術研究組合新情報処理開発機構
Real World Computing Partnership.
mtakeuch@crl.hitachi.co.jp, http://www.rwcp.or.jp/

小泉 敦子
Atsuko Koizumi

(同上)
koizumi@crl.hitachi.co.jp, http://www.rwcp.or.jp/

松田 純一
Junichi Matsuda

(同上)
j-matsud@crl.hitachi.co.jp, http://www.rwcp.or.jp/

佐川 浩彦
Hirohiko Sagawa

(同上)
h-sagawa@crl.hitachi.co.jp, http://www.rwcp.or.jp/

Keywords: sign language, gesture, non-manual sign, corpus, grammar.

1. はじめに

近年、交番や役所などの公共機関、デパートや外食産業などのサービス業務窓口において、手話によるコミュニケーションを目にする機会が増えている。また、手話による通訳を伴うテレビ番組が放送されている。手話の社会的な認知度が高まるにつれて、手話から日本語あるいは日本語から手話への翻訳を目的とした工学的な手話の研究が多くの研究機関で進められている [手話工学 97]。これらの研究にはバリアフリー社会実現のための基盤技術としての社会的意義が与えられてきた。インターネットの普及に伴い、情報格差の是正が要請される現在において、今後は産業的意義をもつものとする。

我々は平成4年に手話認識技術の研究開発を開始した。研究開発の特色は、実世界との関わりを強く意識している点にある。以下ではRWCの最終成果、マルチモーダル手話認識方式を概説し、認識対象と開発過程の点から特色を述べる。

2. マルチモーダル手話認識

RWCPの最終成果「マルチモーダル手話認識プロトタイプシステム」の外観を図1に示す。このシステムでは、手袋型入力装置により高精度かつ実時間で手指動作が入力される。また、2台のビデオカメラにより頭部の動きを入力する。CRTには入力内容に対応する日本語文が表示される [佐川 01]。

図2にマルチモーダル手話認識方式のブロック図を示

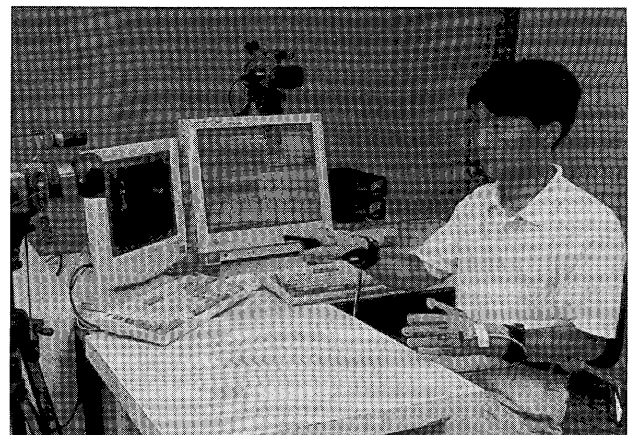


図1 マルチモーダル手話認識システム

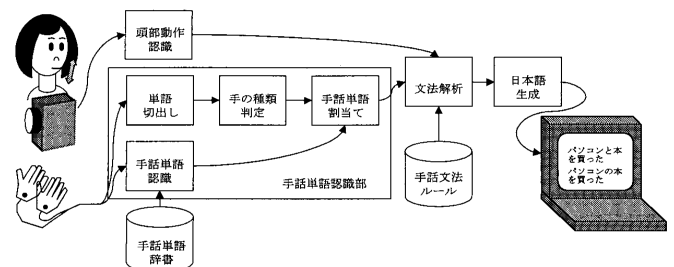


図2 マルチモーダル手話認識方式

す。

手袋型入力装置からは空間における手の位置と方向、指の関節角・開き角が1/30秒間隔で入力され、手話単語辞書を参照し、手話単語認識が行われる。それと同時に手指動作の切れ目の検出が行われ、単語が切り出される。また、利き手・非利き手・両手のどれで単語が表現されているか、つまり、手の種類の判定も行われる。そ

の結果を用いて手話文発話の連続する手指動作において、単語が発話されたと推定される部分に手話単語が当てはめられる。一方、手指動作の変化やカメラから入力される頭部動作から文法情報が抽出される。これらの単語情報と文法情報に対して手話文法規則が参照され、発話内容に対応する日本語文が生成される [Sagawa 00]。

実装は標準的な CPU 速度の PC1 台を使用して行い、リアルタイムな手話から日本語への通訳が行えることを確認した。なお、頭部動作認識方式には産業技術総合技術研究所で開発された高速な画像認識技術を利用した [Xu 00]。

3. 日本手話の認識

手話を母語とする聴障者が日常使用する日本手話は、日本語とは異なった言語である。日本手話では、手指動作（手の動き）および、非手指動作（頭部動作、視線、口形、眉の動き）によって情報を伝達する。特に手指動作の変化や非手指動作は文法的情報を伝達するために重要といわれている [神田 94, 木村 95]。日本手話の言語学的な研究は途上にあり、手指動作の変化の種類や非手指動作の種類、およびそれらの機能において明確になっていない部分が多い。

手話文法規則についてはさまざまな知見が発表されている。特に計算機への実装を主眼に置き、これらの知見を検証し、手話認識方式の開発を行っている。

手話では手指動作の変化によって文法情報を表現する場合がある。例えば体の前面右側で手話単語「男性」を表現し、続いて左側で手話単語「女性」を表現する。最後に手話単語「本」を右から左へ動かせば、「彼が彼女に本を渡す」という発話になる。逆に左から右に動かせば、「彼女が彼に本を渡す」という発話になる。つまり動作の方向が主語と直接目的語を決定する。このような表現は方向変化動詞と呼ばれている。

従来、動作の開始位置が主語を示し、終了位置が目的語を示すといわれていたが、実験結果から位置は重要ではなく、方向が重要であることを確認している。極端な場合、「男性」を提示した位置よりもはるかに外側から、体の中央に「本」を移動しても「彼が彼女に本を渡す」ことが表現される [佐川 98]。

さらに、手話文法規則を抽出を目的として、手話発話データの採取とラベル付けを行った。

自然な発話データを得るには即時的な自由発話が望ましいが、言語現象を分析するためには、肯定文と否定文、目的語が一つの文と二つの文、ある単語が単独で存在する場合と係り受け関係にある単語が存在する場合など、部分的に異なる例文を比較する必要がある。そこで、文法記述用データとして有用と思われる基本例文を日本語で作成し、それをもとに聴障者 3～4 名による手話表現を収集するという方法を取った。収集作業においては、

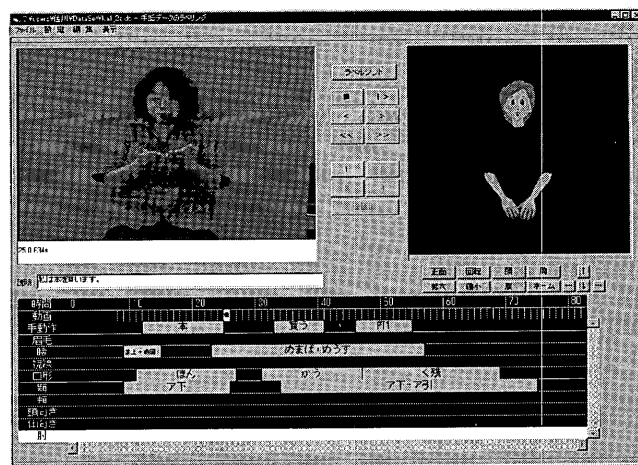


図3 ラベル付き日本手話コーパス

より自然な日本手話の表現にするため、聴障者と手話通訳者がペアとなり、どのような手話表現にするかを十分に議論させた。また、収集した手話をチェックし、不自然と感じられる場合は自然な手話表現になるまで再収集を繰り返した [小泉 01]。

ラベル付き日本手話コーパスは手話文の発話を撮影し、手指動作および非手指動作のラベルを付与したものである。現段階でのラベルの情報は、表現された手話単語と非手指動作を時間軸に沿って記述したものであり、いわば手話発話の書き起こしデータである。図3の手話文法解析支援ツールは、ラベル付与や手話文法の分析に使用する。手話発話の映像を下方のラベル情報と対応付けて画面に表示したり、手袋型入力装置の波形データを手話アニメーションで表示することができる。

約 2 500 発話へのラベル付与を行い、手話文法の分析に利用した。そして、少なくとも 50 種類の非手指動作が存在することを確認した。このうち頭部動作に関しては、手話文の解析に重要となる可能性がある動作は 9 種類、手話文中の機能としては、フレーズの区切り、文末提示、目的語提示、疑問詞など、数種類の機能が存在することを明らかにした。これらの解析結果に基づき、手話文法を記述し、手話構文解析による中間表現の生成と中間表現からの日本文生成方式を考案した。

4. 実際に発話される手話の認識

実際の発話においては、手指動作は発話時の意図・文脈・前後関係などによって変化し、必ずしも辞書や学習教材に示されたままの動作が提示されるとは限らない。このため、実用性のある手指動作の認識方式には動作変化に対する高度な柔軟性が要求される。また、これらの処理には実時間性も要求される。

手話単語認識は三つの段階を経て開発された。

初期の手話単語認識方式は DP マッチングを用いたものであった。辞書に記述された動作パターンとのマッチングを取るこの方式では実際の発話における手指動作の

変化に十分に対応することはできなかった。このため、構動素という動作の基本単位を考案し、手話単語認識を試みた。

「言う」という手話動作は、右手の五指をすべて曲げた形で、掌を前方に向け、口の前方に置き、その後、五指を開きながら、掌の方向を変化させずに、前方に直線移動する。構動素による手話単語認識方式では、手の形、手の方向、手の運動の各因子に対して、五指をすべて曲げた形、五指をすべて開いた形、掌の方向Z軸正、直線運動、運動方向Z軸正といった構動素と呼ぶシンボルに基づいて手話単語辞書を作成し、各構動素の認識結果から辞書を参照し、手話単語を決定する。複数の構動素を因子に割り当てることで手指動作の変化に対してある程度対応することができる。手指動作の変化がもたらす構動素の変化は手話単語ごとに異なる。このため現行のシステムでは単語ごとに、手の形、手の方向、手の運動の各因子の構動素は平均値や分散といった統計量に基づき記述している。例えば、手の形については各関節の曲げ角と指の開き角の平均値と分散が記述されている。これらの値は複数回の発話データを採取し決定される [Sagawa 98]。

手話単語認識方式を手話文認識に適用する場合、「渡り」と呼ぶ手話単語と手話単語の間の部分の影響で多数の手話単語の候補を生成する現象が起こる。これを「湧出し」と呼ぶ。「湧出し」は認識率の低下の大きな要因である。手話単語と「渡り」には動作特徴上の差異は見いだせなかった。しかし、「動作」、「停止」、「ポーズ」、「ノイズ」という動作特徴上の差異を用いて、手話単語を切り出し、「湧出し」を減少させる方法を考案した。「動作」は「手話単語」と「渡り」、「停止」は「動作」の境界、「ポーズ」は発話者が意図的に動作を停止させている部分、「ノイズ」は手をぶつけ合うような接触動作に起因する動きである。例えば「手話単語」+「渡り」が手話単語辞書のデータとマッチして「湧出し」が起こった場合でも、間に「停止」が検出され、「動作」+「停止」+「動作」と解釈されれば、「手話単語」+「渡り」による「湧出し」を取り除くことができる [佐川 99]。

手話動作が片手のみで表現されている(片手手話)か、あるいは両手を使用して表現されている(両手手話)かを判定することも手話文の発話における手話単語認識の高精度化のために重要である。微小区間における左右の手動作の速さの差の累積値に基づいて動作を行っている手の種類の判定を行う方式を考案した。前述の手話単語の切出し方法と同様にこの方式によっても手話文発話における手話単語認識の精度を向上させることができる [佐川 99]。

5. システム開発の経緯

実世界で使用される手話認識技術は発話された日本手

話を認識対象として設定する必要がある。そして、開発技術を積極的に実世界に出し、批判を仰ぎ、評価を行うべきであると考えた。

平成9年に双方向の手話認識プロトタイプシステムを作成し、計算機による手話認識というコンセプトを世の中に提示した。このシステムは手話単語を連続的に認識し、テンプレートを用いて日本語文へ変換する機能を有する。現在のシステムに比べると、発話内容への制限は非常に強く、ある程度の自由度はあるが決められたシナリオを通訳できるのみである。しかし、このシステムが基盤となり、新しい技術を開発するたびに逐次、システムの改良と実装レベルでの評価を繰り返し、日本手話を対象とする現在のシステムへと至っている。

平成12年には手話入力機能付情報提供端末を長崎県諫早市役所に設置して、手指動作認識技術に関する試験運用を実施した。このシステムは手話での質問をキーとして、情報検索を行う機能をもつ。試験運用の結果、想像以上に文型や語順の多様性、単語表現の地域性が見いだされた [Sagawa 00]。

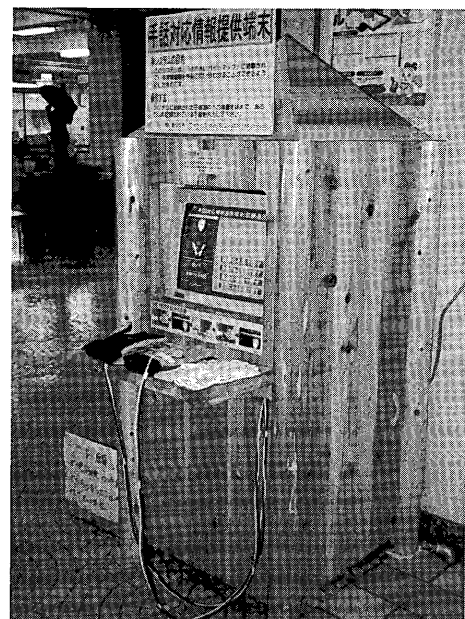


図4 手話入力機能付き情報提供端末

我々の研究開発の目的の一つに、手話を理解する人を増やすことがあげられる。そのためには、楽しく学べるのが一助となる。手袋型入力装置を使用すると学習者の手話動作の入力ができる。手話教育システムでは、手話動作の評価、お手本との比較、アニメキャラクターとの対話といった機能をすでに実現している。手話教育というタスクは正解が既知であるため、手話認識というタスクに比べて幾分簡単であり、早期製品化を実現したいと考えている [Sagawa 00]。

6. 今後の課題

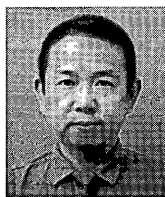
マルチモーダル手話認識システムの実用化には、(1) 文法規則の拡充、(2) 頭部動作以外の非手指動作処理方式の導入、といった課題が残るが、製品化を考慮する段階にあるものと考えている。手袋型入力装置の価格は数百万円である。製品化のためには、より安価な手指動作入力装置の導入・開発が必要となる。また、単語表現の地域性に対応するため、技術開発だけではなく、地域社会に密着したコンテンツ作成体制の整備も必要となる。

◇ 参考文献 ◇

- [神田 94] 神田和幸: 手話学講義, 福村出版 (1994)
 [木村 95] 木村晴美, 市田泰弘: はじめての手話, 日本文芸社 (1995)
 [小泉 01] 小泉敦子, 佐川浩彦, 竹内 勝: ラベル付き手話コーパスの作成, 第6回電子情報通信学会福祉情報工学研究会 (2001)
 [佐川 98] 佐川浩彦, 竹内 勝: 空間情報を利用した手話認識方式, 第14回ヒューマン・インタフェース・シンポジウム論文集, pp. 169-174 (1998)
 [佐川 99] 佐川浩彦, 竹内 勝: 手話認識における手話動作セグメンテーション方式, ヒューマンインタフェースシンポジウム'99論文集, pp. 749-754 (1999)
 [佐川 01] 佐川浩彦, 小泉敦子, 竹内 勝: 頭部動作によって表される手話文法情報を利用した手話文認識方式, ヒューマンインタフェース学会論文誌, 第3巻, 第3号, pp. 145-152 (2001)
 [手話工学 97] 手話工学研究会 編: 手話工学, 電子情報通信学会第3種手話工学研究会報告書 (1997)
 [Sagawa 98] H. Sagawa, M. Takeuchi and M. Ohki: Methods to Describe and Recognize Sign Language Based on Gesture Components Represented by Symbols and Numerical Values, Knowledge-Based Systems, Vol. 10, No. 5, pp. 287-294 (1998)
 [Sagawa 00] H. Sagawa, H. Ando, A. Koizumi, K. Iwamura, M. Takeuchi: Sign Language Recognition and Its Applications, Proc. 2000 RWC Symposium, pp. 143-146 (2000)
 [Xu 00] M. Xu, B. Raytchev, K. Sakaue, O. Hasegawa, A. Koizumi, M. Takeuchi and H. Sagawa: A Vision-Based Method for Recognizing Non-Manual Information in Japanese Sign Language, Proceedings of the Third International Conference on Multimodal Interfaces (ICIM2000) (2000)

2002年1月10日受理

著者紹介



竹内 勝

1984年東京工業大学理学部数学科卒業。1986年同大学院総合理工学研究科システム科学専攻修士課程修了。同年(株)日立製作所入社, 基礎研究所配属, 脳の情報処理機能, 遺伝的アルゴリズムの研究に従事。1995年より中央研究所勤務。手話アニメーション生成, 認識の研究に従事。2000年グッドデザイン賞(インタラクション部門)受賞。電子情報通信学会, 計測自動制御学会各会員。



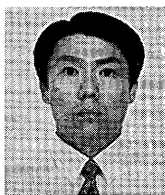
小泉 敦子

1982年国際基督教大学教養学部卒業。同年,(株)日立製作所入社, システム開発研究所配属。1993年同社中央研究所転属, 現在に至る。自然言語処理の研究に従事。情報処理学会会員。



松田 純一(正会員)

1984年東京大学理学部数学科卒業。同年(株)日立製作所入社, システム開発研究所勤務。1993年同社中央研究所勤務。自然言語処理, 機械翻訳, 多言語処理, ヒューマンインタフェースなどの研究開発に従事。情報処理学会会員。



佐川 浩彦

1989年東京大学工学部電気工学科卒業。1991年同大学院工学系研究科情報工学専攻修士課程修了。1991年(株)日立製作所入社, 同社中央研究所配属, 現在に至る。手話認識・生成の研究に従事。2000年ヒューマンインタフェース'99学術奨励賞, 2001年日本ITU協会賞ユニバーサルアクセシビリティ賞受賞。情報処理学会, ヒューマンインタフェース学会, 日本手話学会各会員。