

手話自動翻訳システムにおける パターン認識・合成

大木 優, 佐川浩彦, 崎山朝子, 大平栄二, 藤澤浩道

(株)日立製作所 中央研究所

手話は、聴覚障害者の日常会話言語の一つであり、主に手の動きによって単語や文を表現する言語である。本報告では、筆者らが開発中である手話自動翻訳システムについて述べる。手話自動翻訳システムは手話と日本語とを自動翻訳するシステムである。(1) 手の形や位置データをデータグローブを使って入力し、手の動作を認識し、手話を日本語に翻訳する。(2) 入力された日本語を手話に翻訳し、3次元コンピュータグラフィックスを使って手話アニメーションとして表示する。

Pattern Recognition and Synthesis for Sign Language Translation System

Masaru Ohki, Hirohiko Sagawa, Tomoko Sakiyama, Eiji Oohira, Hiromichi Fujisawa

Central Research Laboratory, Hitachi, Ltd.

Sign Language is one of the communication means which hearing handicapped people use. Words and sentences are mainly represented by gestures of hands. In this report, we show a sign language translation system which we are developing. The system translates Japanese sign language into Japanese and vice versa. (1) Hands shape and position data are inputted by using data gloves. Inputted hand actions are recognized, and is translated into Japanese sentences. (2) Japanese text is translated into the sign language represented as 3-D computer graphic animation of sign language gestures.

第1章 緒言

手話は、聴覚障害者が日常会話言語として使っている言語であり、聴覚障害者にとって負担の少ない言語である。手話の特徴は、手の動きや身振りなどのジェスチャーによって情報を伝達することである。手話と日本語では、単語の表現方法以外に語の順序などの文法的な側面も異なるので、手話を知らない健聴者には、手話を見たとしても伝えたい情報を理解することは難しい。そのため、聴覚障害者と健聴者との間で情報を伝達する場合、手話と日本語を通訳する手話通訳者を必要とする。たとえば、聴覚障害者が病院に行くときには、手話通訳者に同行を依頼することが多いと言われている。しかし、緊急の場合など、手話通訳者の同行が困難である可能性が高いので、手話通訳者を介さなくても、会話をしたいというニーズが強い。このニーズにこたえる手段の一つとして、著者らは、コンピュータによって手話と日本語との間を翻訳する手話自動翻訳システムを開発している[1,3,4,5,6]。手話に関する研究は、手話が聴覚障害者の重要なコミュニケーション手段であることから、大学などでも行われている[8,10,12]。

著者らが開発中の手話自動翻訳システムは手話認識部と手話合成部からなる。

(1) 手話認識部

手話から日本語への翻訳を行う。手話を表現する手の形や位置データをデータグローブを使って入力する。入力した手の動作をパターン認識し、手話から日本語に翻訳する。

(2) 手話合成部

日本語から手話への翻訳を行う。日本語を

入力すると手話に翻訳し、3次元コンピュータグラフィックスを使って手話アニメーションを表示する。

手話アニメーションは、手話自動翻訳システムだけではなく、聴覚障害者への情報伝達手段として使うことが期待できる。現在社会において情報を伝える手段のほとんどは健聴者を対象としており、文字、画像、音が中心である。聴覚障害者のより広い社会参加を促進するためには、文字以外に聴覚障害者にとって負担の少ない手話による情報伝達も重要になってくると考えられる。

本報告では、聴覚障害者の現状について述べ、手話自動翻訳システムの手話認識と手話合成の方式について述べる。

第2章 聴覚障害者の現状と福祉施策

2.1 聴覚障害者の現状

日本全体で聴覚障害者は約35万人であるといわれている。表1.1に聴覚障害者を始めとする障害者の状況を示す。なお、聴覚障害者は、聴覚・言語障害に区分される。1級に近いほど障害の度合いが高い。

聴覚障害者の日常のコミュニケーション手段は、手話以外に口話、筆談がある。文献[9]によると、それらの特徴は、表1.2となる。一方、米国には、聴覚障害者とスムーズにコミュニケーションできるのであれば、あらゆるメディアを使ってコミュニケーションするというトータルコミュニケーションの理念[2,9]が20年前から提唱されている。この理念に従えば、これらのどれかを使うべきだと限定することは、適切ではない。表1.2の手段を状況に応じて、適切に組み合わせ

表1.1 日本の障害者数（昭和62年、単位：千人）

内訳	総数	1級	2級	3級	4級以下
聴覚・言語障害	354	28	90	65	171
視覚障害	307	107	66	30	105
肢体不自由	1,460	186	291	246	735

表 1.2 聴覚障害者のコミュニケーション手段の特徴

方法	特徴
手話	<ul style="list-style-type: none"> ・リラックスした気分で相手の話を受けとめることができる。 ・発信内容を即時に理解できる。 ・長時間コミュニケーションしても疲れない。 ・語彙数が少ないので間違って読まれることがある。
読話	<ul style="list-style-type: none"> ・方言がある。 ・同じような口形があるため、会話の内容や前後関係が理解できていないと読話は難しい。
筆談	<ul style="list-style-type: none"> ・日常使い慣れていない言葉は理解しにくい。 ・多くの聴覚障害者が使用している。 ・重要な話などに適する。 ・解かりやすく書く必要がある。

て使うべきである。

トータルコミュニケーションの理念に従えば、聴覚障害者に一方向のみで情報を伝達する場合でも、文字、手話、画像などを使って情報を伝達すべきと考えられる。手話アニメーションは文字と共に聴覚障害者への情報伝達メディアである。メディアにはそれぞれ長所や欠点があるので、手話アニメーションと文字の利用も状況に応じて使うべきである。

2.2 聴覚障害者の福祉施策

米国では障害者の社会参加が進んでいる。1986年には、障害者も電子事務機器を使えることを保証する法律も制定され、電子事務機器を使えないことを理由にして障害者の雇用を拒否してはならなくなった。1988年には、「電子機器アクセシビリティ指針」が公布され、連邦政府はこの指針に沿った機器しか購入できなくなった。さらに、雇用・交通・建物・通信における障害者の差別撤廃を求める「米国人障害者法」が成立した。

一方、日本では、米国の「電子機器アクセシビリティ指針」に対応して、1990年に「情報処理機器アクセシビリティ指針」が通産省から公表された。このように、国内外とも、電子機器や情報機器における障害者への対応はますます重要となっている。

ろうあ学校での教育では、最近まで、手話の教

育は禁止され、読話や発語による口語法の考え方で教育が行われていた。しかし、読話や発語が簡単ではないことなどから手話が注目され、福祉の領域で手話が用いられてきている。1988年に手話通訳士の認定制度が決まり、1989年から手話通訳士の試験が行われるようになった。このような経緯から、手話は聴覚障害者の言語として認められてきた。

第3章 手話自動翻訳システム

手話は、日本語と異なり、以下のような特徴を持つ。

- (1) 主に手の動きで意志を伝える。
- (2) 表情や身振りも言語的構成要素である。

手話は、表情や身振りでも重要な情報を伝える。たとえば、手話単語を注目させるために手の動作以外に、あごを動かしたりする。

- (3) 語彙が日本語と異なる。

たとえば、手話では「夏」と「暑い」が同じ動作である。一方、「家に上がる」と「エレベータで上がる」の「上がる」は手話では異なった表現である。

- (4) 一つの動作で複数のことを表現することがある。

右手と左手で別々のことを表すことにより、同時に複数のことを表現することができる。たとえば、「飛んでいる飛行機を見る」

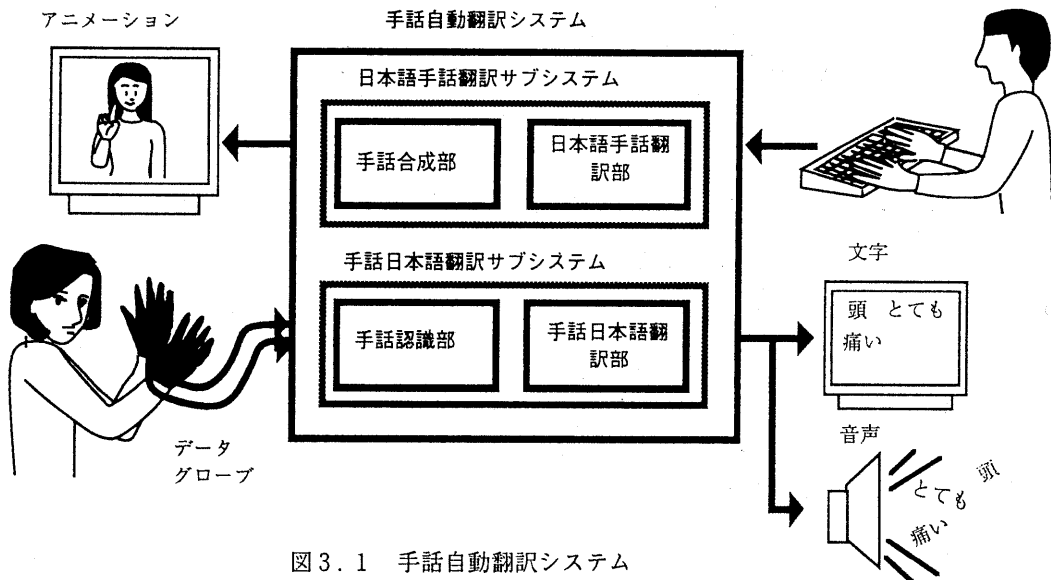


図 3.1 手話自動翻訳システム

ということを一つの動作で表現できる。

(5) 助詞を使わない。

助詞の代わりに、空間的な位置関係を利用する。「彼が彼女に本をあげる」と言うことを手話で伝える場合は、まず、右手で「彼」（親指のみを立てた動作）を表し、左手で「彼女」（小指のみを立てた動作）を表す。そして、「彼」を表した右手の位置で「本」（両手の掌を付け、親指側から開く動作）を表し、そのまま左手の位置に動かす。「本」を右から左に動かしたことが、「～から～へ本を移動する」ことを表す。

以上のように手話と日本語は、表現形態と言語体系が異なる。そのため、手話自動翻訳システムは、パターンとしての表現形態を変換する以外に、手話と日本語の言語としての違いも翻訳する必要がある。

著者らが開発中の手話自動翻訳システムの構成を図 3.1 に示す。手話自動翻訳システムは、手話を日本語に翻訳する手話日本語サブシステムと、日本語を手話に翻訳する日本語手話サブシステムからなる。

手話日本語翻訳サブシステムは、手話を認識す

る手話認識部と、認識した手話を日本語に翻訳する手話日本語翻訳部からなる。手話認識部では、データグローブ^{*)}を使って手の動きを入力し、パターン認識を行なう。

日本語手話翻訳サブシステムは、日本語を手話に翻訳する日本語手話翻訳部と、手話を 3 次元コンピュータグラフィックスを使ってアニメーションとして表示する手話合成部からなる。以下では、手の動きをパターン認識する手話認識部と手話のパターンをアニメーションとして生成する手話合成部について述べる。

第 4 章 手話認識

4.1 手話認識における課題

手話通訳を実現するための手話認識における課題は以下のとおりである。

- (1) 手の動きの認識
- (2) 表情や身振りの認識
- (3) 実時間の認識

手話通訳システムでは、実時間で通訳する

^{*)}米国 VPL Research Inc. 社が開発したものであり、DATA Glove は VPL Research Inc. 社の商標。

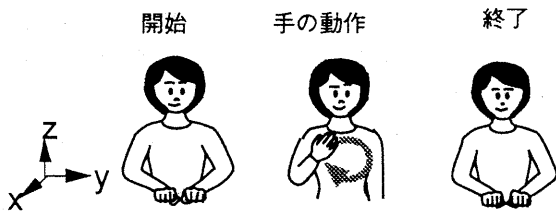


図 4.1 [具合] という手話

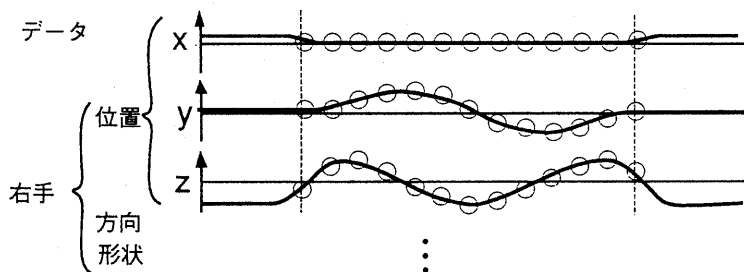


図 4.2 [具合] という手話の右手の位置の変化

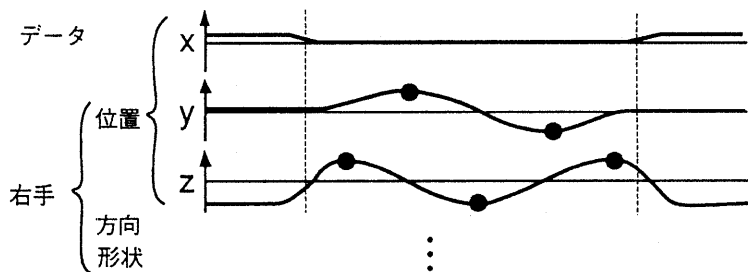


図 4.3 位置の変化の極小点について [具合] という手話を特徴抽出した結果

必要がある。そのため、手や表情、身振りの認識を実時間で行う必要がある。

(4) 連続手話認識

音声認識と同様に、手話で文を伝える場合、なんらかの区切りを使って単語の区切りを明確にする場合と、単語を区切らない場合がある。連続手話認識は、手話単語の区切りをシステムが自動的に認識するものである。

(5) 不特定手話者認識

不特定手話者認識は、特定の人の手話を認識するのではなく、不特定の人の手話を認識

するものである。音声認識の不特定話者認識に相当する。

これらの課題の中で、必要な課題は、(1)と(3)である。次節以降では、(1)と(3)を解決する方式について述べる。

4.2 手話認識方式

著者らが開発中の手話自動翻訳システムでは、手の動作を入力するためにデータグローブを用いている。データグローブには、一つの指に2本ずつ、片手に10本の光ファイバが取付けられている。この光ファイバを使って指の曲げ具合を検出する。また、手の位置は磁気センサーで検出する。得られるデータは、指の関節の曲がり具合、手の位置、手の向きを表すデータで、片手で合計16個、両手にすると32個のデータになる。そして、これらのデータは動作パターンとして1秒間に30回の頻度でコンピュータに入力する。

データグローブから入力された動作パターンをそのまま使うと、データの量が多すぎるため、手の動作を短い時間で認識することが難しくなるとい問題がある。これを解決するために、入力された動作パターンと標準パターンとを比較する照合処理の前に、手の動きの特徴抽出を行う。特徴抽出によってデータ数を減らすことにより、照合処理を高速にする。

(1) 特徴抽出

データグループで取り込まれた動作パターンを解析して、手話パターンの特徴抽出を行なう。このようにすると、手の動きの特徴的な動作パターンが得られ、データを圧縮することができる。

(2) 照合

特徴抽出した動作パターンを使って、あらかじめ登録しておいた手話の単語の標準的な動作パターンと照合する。現在、照合方式としてDP (Dynamic Programming) 照合[11]を用いている。この処理では、コンピュータに登録してあるすべての単語の標準的な動作パターンと照合する。そのためデータ量が多いと、単語数が増大した場合、認識する時間が膨大になる。ここでは、入力された動作パターン以外に標準的な動作パターンも特徴抽出によりデータを圧縮している。そのため、データの量を減らしているため、高速に照合することができ、認識時間を短くすることが可能となっている。

4.3 特徴抽出

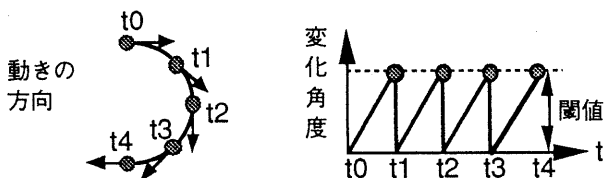


図 4.4 ベクトル変化の大きい点について「具合」という手話を特徴抽出した結果

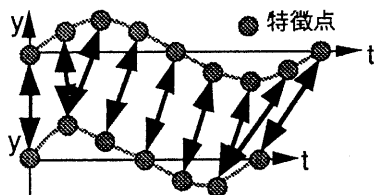


図 4.5 特徴抽出した点による照合

本節では、手話認識方式の特徴である手の動きの特徴抽出について述べる。手話の手の動作は、以下の二つに大きく分けることができる。

- (1) 指文字のように静的なパターン
- (2) 動的なパターン

本研究では、パターンの特徴として、以下の特徴を選び、それらの特徴点と特徴点間の時間長によって手話のパターンを表す。

- (1) 位置の変化の極小点
- (2) 動きのベクトル変化の大きい点

これらの特徴抽出の方法について説明する。図 4.1 は胸の前で、円を描く手話である。これは、「具合」を示す手話である。データグループから入ってくる手の動きは、X、Y、Z方向で見ると、図 4.2 のように変化している。1秒間に手の位置や形状のデータが30回送られてくる。

図 4.2 における変化の極小点は、図 4.3 となる。位置座標から見て極値を取ったところが特徴点である。この方法では、基本的な動きの特徴を抽出することができる。

しかし、その方法だけでは、図 4.1 に示したような弧を描く動きの特徴抽出ができない。そのため、動きのベクトル変化を加算し、ある量を越えたならば、それを特徴点とする。円弧の場合、図 4.4 に示すように t_1 や t_3 などの特徴点を抽出することができる。

照合には、入力されたパターンと標準パターンとも特徴抽出を行なったパターンを使う。図 4.5 に示すように、二つのパターンともパターンが圧縮されているので、照合する点の数が少ない。その結果、照合時間の短縮が図られる。

特徴抽出した場合としない場合の結果は図 4.6 に示すとおりである。17単語の手話データの認識実験で、特徴抽出しない場

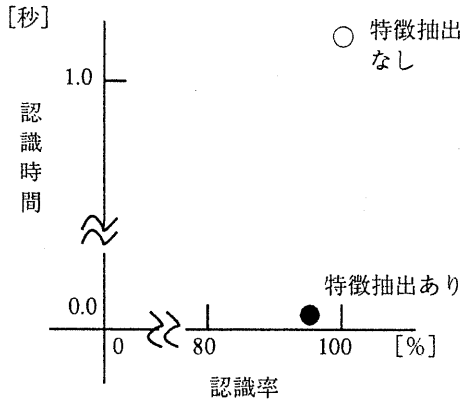


図 4.6 照合の性能

合は、100%の認識率であるが、HP社製ワークステーションHP9000/720 (57MIPS) を使用して1単語当りの認識時間は1.23秒である。特徴抽出すると、認識率は97%であるが、1単語当りの認識時間は0.076秒となる。

第5章 手話合成

手話合成部は、日本語から手話へ翻訳する部分である。手話は3次元CG (コンピュータグラフィックス) を使って手話アニメーションとして表示する。手話合成部のシステム構成を図 5.1 に示す。

手話アニメーションを生成する場合の重要な課題は、以下の二つである。

- (1) 個々の手話単語をアニメーションとして

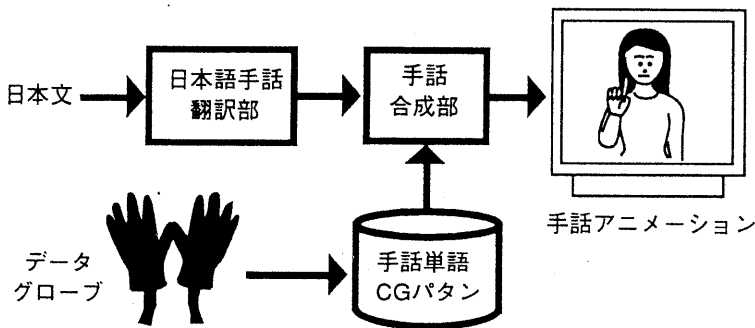


図 5.1 手話合成部

動かすために3次元CGのデータを作る必要がある。

- (2) 個々の手話単語アニメーションをつなげて、一つの手話の文にするために、単語間の動きを滑らかにする必要がある。

本システムでは、課題 (1) を解決するために、データグローブで入力された手の位置や形のデータをそのまま3次元CGを動かすために利用する。

課題 (2) については、単語と単語間のアニメーションを自動的に補間をすることにより解決した。日本語から手話に翻訳する場合は、日本文を日本語手話翻訳部で手話単語の見出し列に翻訳する。その手話単語の見出しに相当する手話単語のCGパターンを呼び出すことにより、一つの手話単語のアニメーションを生成する。単語間の補間は、前の単語のアニメーションが終了した手の位置から、後の単語のアニメーションが開始される手の位置までの動きを自動的にアニメーションとして生成することにより行う。その結果、手話の文が、滑らかな連続的なアニメーションとして表示される。

本手話合成方式の特徴は、手話単語CGパターンを組み合わせるだけで、手話の文をアニメーションとして表示することができることである。そのため、手話アニメーションの作成や変更が容易である。

手話を表示する場合、コンピュータグラフィックスではなく、実写 (ビデオ) を使うことが考えられる。実写は、人間が手話を行っているため、人間味があり、暖かな印象を受けるなどの利点がある。しかし、自動翻訳システムに使用する場合、任意の日本語を翻訳する場合、実写では、任意の手話を連続的に表示することは困難であ

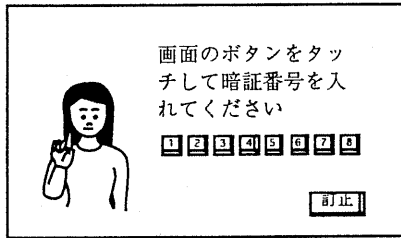


図 5.2 住民票自動交付システムの手話の説明

る。また、一般に、ビデオの修正や編集は困難である。さらに、時間がたってしまった場合、同一人物でも外見が変わってしまう可能性があるため、変更はますます難しくなる。

手話アニメーションは、手話通訳システムだけでなく、聴覚障害者への情報を伝達する手段の一つとして使うことが考えられる。たとえば、住民票自動交付システムの操作を聴覚障害者に説明するために、手話アニメーションを使うことができる(図5.2)。文字による画面の操作説明の他に、手話による操作説明を表示する。文字と手話の両方で説明を行うことによって、聴覚障害者により分かりやすい操作の説明ができる。住民票自動交付システムに限らず、公共施設などにおいて、聴覚障害者へ情報を伝達する必要がある状況は多い。このような状況において、手話アニメーションは、聴覚障害者への情報メディアとして広く使うことが期待できる。

第6章 結言

本報告では、聴覚障害者の現状について述べ、著者らが開発中である手話自動翻訳システムについて述べた。手話認識については、さらに大語彙化を目指して認識速度を速くする必要がある。手話合成については、聴覚障害者の情報メディアとしての応用を検討していきたい。

参考文献

[1] 大平栄二, 崎山朝子, 阿部正博, 佐川浩彦

: 手話合成システムの基本検討, 第46回情報処理学会全国大会, 8P-3, 1993年。

- [2] 小川仁, 神田和幸: 手話通訳の基礎, 第一法規, 1992。
- [3] Sagawa Hirohiko, Sakou Hiroshi, Abe Masahiro : Sign Language Translation System Using Continuous DP Matching, IAPR Workshop on Machine Vision Applications, pp.339 - pp.341, 1992.
- [4] 佐川浩彦, 酒匂裕, 阿部正博: 連続DP照合を用いた手話通訳システム, ヒューマンインタフェース研究会, 情報処理学会, 44-12, 1992年。
- [5] 佐川浩彦, 阿部正博: 手話通訳システムにおける大語彙化に関する検討, 第46回情報処理学会全国大会, 8P-2, 1993年。
- [6] 崎山朝子, 大平栄二, 佐川浩彦, 阿部正博, 新井清: アニメーションによる手話合成方式の検討, 第46回情報処理学会全国大会, 8P-4, 1993年。
- [7] 全日本ろうあ連盟「手話研究委員会」: わたしたちの手話, 財団法人全日本ろうあ連盟, 1987。
- [8] 寺内美奈, 長嶋裕二, 三原浩樹, 長嶋秀世, 大和玄一: アニメーションによる日本語手話表現に関する基礎的検討, ヒューマンインタフェース研究会, 情報処理学会, 41-7, 1992。
- [9] 野沢克哉, 聴覚障害者のケースワークII, 1991。
- [10] 松本崇, 鎌田一雄: 手話ワープロ構築に関する基礎研究, ヒューマンインタフェース研究会, 情報処理学会, 50-6, 1993。
- [11] 森健一: パターン認識, 電子情報通信学会, 1990。
- [12] 渡邊茂晃, 井上隆也, 藤重栄一, 黒川隆夫: 日本語・手話翻訳の自動化に関する技術的課題, Hyman Interface, Vol. 8, pp. 363 - pp. 370, 計測自動制御学会, 1993。