

8P-2 手話通訳システムにおける大語彙化に関する検討

佐川 浩彦, 阿部 正博

日立製作所 中央研究所

1はじめに

手話は聴覚障害者の重要なコミュニケーション手段であり、聴覚障害者と健聴者のコミュニケーションは手話通訳者を介して行なわれる。しかし、近年、聴覚障害者のプライバシー保護や手話通訳者の不足、手話通訳者の職業病が問題となってきた。

このような背景のもとに我々は、自動手話通訳システムの開発を進めている[1]。[1]では、連続的な手話文から手話単語を認識し、文への変換を行なう手話通訳システムのプロトタイプの開発について報告した。しかし、この時点での単語数は26と非常に少なく、実用的なシステムにはまだまだ多くの課題が残されていた。

手話における単語数は、基本的な単語だけでも約3000語あるといわれている[2]。しかし、病院や警察のような限定された状況を考えると、使用される単語数も1000語程度あれば十分ではないかと考えられる。そこで、我々はまず、1000語の認識を目標としてシステムの開発を進めている。

本稿では、システムが認識可能な語彙数を620語に拡張して手話単語認識実験を行ない、現状の認識方式による認識精度の評価および大語彙化における問題点の検討を行なった結果を報告する。

2手話認識方式

連続的な手話文から高速に手話単語を認識するために、圧縮連続DP照合法[1]という照合法を用いる。この手法は、音声認識において一般的な手法である連続DP照合法[3]を改良し高速化した手法である。

圧縮連続DP照合法では、まず、標準手話パターンおよび入力手話パターンをその動的な特徴に基づいて圧縮する。手話パターンの圧縮では、特徴点として、

1. 速度が極小となる時刻

2. 速度ベクトルの方向変化がある閾値を越えた時刻

を検出する。そして、これらの点における手話パターンのバターンベクトルと特徴点間の時間長でパターンを表すことによって圧縮を行なう。

次に、このようにして圧縮したパターン同士を連続DP照合によって照合する。連続DP照合法は、非線形に伸縮するパターンをある制約(DPバス)のもとで時間の伸縮を許しながら連続的に照合していく方法である。これを圧縮したパターンの照合に利用する場合、圧縮したパターンは時間軸が非線形に圧縮されているため、通常のDPバスを使用したDP照合では、パターン間の対応がうまくとれない可能性がある。そこで、DPバスとして通常より対応点間の制約を緩くした図1のようなバスを採用する。また、バスの重みとして圧縮パターンの各特徴点の時間長を用いる。

圧縮連続DP照合の結果、図2のように圧縮パターンの各特徴点の対応が求められる。

Enlargement of Vocabulary on Sign Language Translation System
Hiroyuki SAGAWA, Masahiro ABE
Central Research Lab., Hitachi,Ltd.

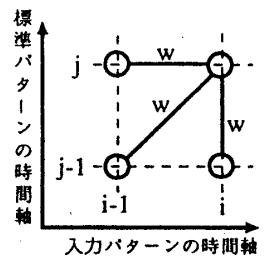


図1：圧縮連続DP照合に用いるDPバス。格子点(○で表示)は各時刻の特徴点が対応することを示す。また、バス(直線で表示)は(i-1,j), (i-1,j-1), (i,j-1)のいずれかの対応が有効になった場合にしか(i,j)という対応は有効にならないことを示す。

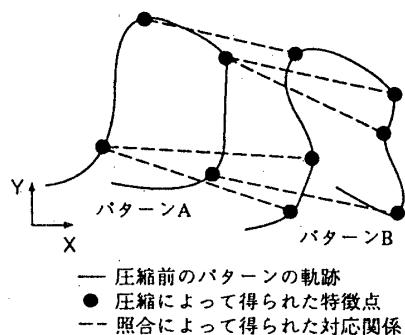


図2：圧縮パターン間の特徴点の対応例

3実験方法

認識実験に使用する手話単語としては、文献[4]、文献[5]に掲載されている手話単語のうち病院での会話に関係のある単語、および今回実験に使用した手話データの登録者が所属している手話サークルで必修単語とされている単語、計620語を採用した。

これらの単語について各10回づつデータをとり、5回分のデータを標準手話パターン作成用に、残り5回分のデータを実験用入力手話パターンとした。データは全て同一話者による。

照合を行なった後、求められた距離の時系列から極小値を検出することによって単語の検出を行なう。ただし、検出処理は0.5以下の距離に対してのみ行なった。この値は、これまでの実験によって経験的に決定した値である。

4実験結果

認識結果より求めた認識率を表1に示す。

また、検出された正しい単語候補の距離の分布と、誤った

表 1: 認識率

1位認識率	97.1%
5位以内認識率	99.4%

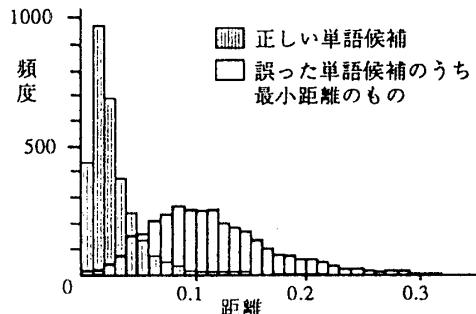


図 3: 正しい単語候補の距離と、誤った単語候補のうち最小距離の候補の分布

単語候補のうち最小距離の候補の分布を図 3 に示す。

さらに、第 1 候補に正しい単語候補が検出されなかったデータについて、解析を行なった。誤認識した単語については、その原因として、

- 原因 1 動作の組合せの中に全く同じ手話が含まれる場合
- 原因 2 動きや形、位置が似ている場合
- 原因 3 両手手話の片手の動きが別の手話になる場合
- 原因 4 両手手話の片手の動きが別の手話に似ている場合
- 原因 5 1 つの動作が複数の意味を表す場合

の 5 種類があることが分かった。図 4 に、上述の原因 2 および原因 4 によって誤認識された単語の例を示す。

5 考察

全体の認識率は、1 位認識率で 97.1%、5 位以内認識率では 99.4% と非常に高い認識率が得られている。このことから、手のデータのみでも手話単語の認識がかなり高い精度で行なえることが分かる。

正しい単語の距離と誤った単語候補のうち最小距離の候補の分布をみると、正しい単語候補は 0.05 未満の距離に集中しており、その頻度は全単語の 87.4% になる。逆に誤った単語候補については、距離が 0.05 以上の頻度は全単語の 90.9% であり、この点からも本認識方式は有効であると考えられる。しかし、今回の実験では、

1. 辞書パターンとテストパターンはほとんど同じ状況で入力されたデータである
2. 入力した状況が異なるデータについて完全に補正を行えなかった

ため、新たなデータや連続手話文の認識等でパラメータの変動が生じた場合、候補の分布は今回の結果に比較して広がると考えられる。このため、認識率が低下することが予想される。

誤認識の原因としては上述のように大きく分けて 5 種類の原因が考えられるが、このうち、原因 5 については辞書を作る際の問題である。また、原因 1 および原因 3 についても、全く同じ動作であるため、誤認識は避けられないと考えられる。原因 2 および 4 による誤認識の場合は、図 4 に示されているように、見かけ上類似点が少ないような単語に誤認識している単語も幾つか見られる。これは、現在の認識方式では、パターンの識別に本来必要なないパラメータも含めて認識を行なっているため、認識において手話の特徴がうまく利用さ

単語(正)	単語(誤)	距離	単語(正)	単語(誤)	距離
風邪	良い	0.022	悲しい	不思議	0.143
悲しい	規則	0.319	賛成	過去	0.387

図 4: 原因 2 あるいは原因 4 のため誤った候補の例 (イラストは文献 [6] による)

れていないことによると考えられる。そのため、手話の特徴抽出と照合における距離計算については改善の余地が残されている。

また、今回は計算量や認識時間について具体的な解析は行なっていないが、本認識方式では単語数に比例して認識時間が増大する。そのため、並列化等による高速化が必要である。ちなみに、HP9000/model720 上の実測値では 1 単語の認識に約 1 分要していた。

6 おわりに

本論文では、我々が開発を行なってきた手話認識方式について、認識可能な語彙数を 620 語に増加して単語認識実験を行ない、認識精度の評価および大語彙化における問題点の検討を行なった。

その結果以下の結論を得た。

1. 単一の単語認識ではあるが、1 位認識率で 97.1%、5 位以内認識率では 99.4% の高い認識率が得られた
2. 新たなデータの認識においてはパラメータの変動により認識精度が低下する可能性がある
3. 手話動作の特徴が認識にうまく利用されていないと考えられる場合がある
今後の課題としては、
 1. 連続手話文への適用
 2. 動作の特徴に基づく距離計算
 3. データの変動の調整方法
 4. 認識時間の高速化

が挙げられる。

参考文献

- [1] 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”, 情報研究, HI-44-12, 1992 年 9 月
- [2] 田上 隆司, “手話のすすめ”, 講談社現代新書, 1983 年
- [3] 岡 隆一, “連続 DP を用いた連続単語認識”, 日本音響学会音声研究会, S78-20, pp.145-152, 1978 年
- [4] 飯塚 千代子, “ともに歩むために おはえようみんなの手話会話篇”, 国際放送
- [5] “わたしたちの手話(会話編 1)”, 全日本ろうあ連盟, 1977 年
- [6] “わたしたちの手話 1 ~ 10”, 全日本ろうあ連盟, 1969 ~ 1986 年