

Title	隠れマルコフモデルを用いた手話単語認識システム
Author(s)	伊藤, 徳広
Citation	
Issue Date	2000-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1346
Rights	
Description	Supervisor:堀口 進, 情報科学研究科, 修士

隠れマルコフモデルを用いた手話認識システム

伊藤 徳広

北陸先端科学技術大学院大学 情報科学研究科

1997年2月15日

キーワード: Sign language, HMM, Basic move, Gesture, recognition.

1 序論

近年、聴覚障害者の社会参加にともなって、聴覚障害者と健聴者がコミュニケーションを行なう機会が増えている。聴覚障害者の様に音声によるコミュニケーションが出来ない場合、コミュニケーション手段として筆談や手話などに頼ることになる。しかし筆談ではその速度に大きな問題があるうえ、聴覚障害者、健聴者双方に大きな負担が掛かる。一方、手話によるコミュニケーションでは、速度的には問題ないものの手話自体を理解できる人が少ないため、どんな状況でも使える訳ではない。また、手話通訳士による通訳のサービスもあるが、通訳士の数が少ないため、利便性に欠ける。そのため、計算機による手話通訳システムの構築が望まれている。

現在、手話の入力方法として採用されている方法は2種類ある。一つは画像を用いる方法であり、もう一方は手形状入力装置などを体に取り付け、実際の角度などを採取する方法である。画像を用いる場合、機器を装着する必要が無いため、話者の動きを制限しないという特徴がある。しかし、指先の角度など細かなデータを取り出す事は困難であり、多くの手話単語を認識するシステムに使用するのは難しい。一方、装着型の入力デバイスを使用する場合、装置は特殊になるが、得られるデータの精度は非常に良くなる。この特徴は、より多くの手話単語を認識するシステムを構築する上で非常に重要である。装着型デ

バイスを用いた手話単語認識例としては、DP マッチングを法により、特定話者で 620 もの単語を 90 %以上の認識率を得られている。しかし、DP マッチングは個人の癖の影響を強く受けてしまうため不特定話者には向いていない。これに対し、近年、新しい手話認識法として音声分野で使われていた隠れマルコフモデル (HMM) を用いた手話単語認識方法が試みられている。HMM は学習によってパラメータを推定するのでデータの揺らぎに強い特徴を持っている。しかし、学習に多量のサンプルデータが必要となるため、単語毎ごとに HMM を作成した場合、それに伴って学習サンプルも増加してしまうという問題がある。

本論文では、HMM 使用における学習サンプル増加を抑える手法を提案し、HMM を用いた不特定話者に対応出来る手話単語認識を提案する。手法としては、手話単語を手形・基本動作・運動平面という部分に分け、HMM で認識する部分は基本動作のみとする。これにより HMM のモデル数は基本動作の数だけとなる。手形状入力装置としては Cbyberglove を、3次元位置姿勢入力装置としては FASTRACK を用いる。まず、HMM で基本動作の認識実験を行なう。つづいて、手形・基本動作・運動平面を組み合わせた手話単語認識システムを構築し、その性能評価を行なう。

2 手話単語の音韻表記

手話単語は大きく分けると、手形と動作、そして動作面の 3つの要素をもつ。手形は手話動作中の指の形を、動作はどのような線形状で手を動かしているかを、そして動作面は腕が動いている位置 (例えば体の正面や水平面など) を表している。手話の音韻表記はこの様に、一つの手話単語の要素を幾つかに分割する事によって少ない記述記号で多くの手話単語に対応している。

3 単語認識手法

従来 of HMM を使った手話単語認識では、一つの手話単語ごとに認識を行っていた。しかし、この手法では単語数が増加すると学習サンプルも増加することになり、単語数増加の大きな障害となる。本研究では、手話の音韻表現を参考にすることにより、手話単語を部分的認識し、その結果を統合することにより単語を認識するシステムの構築を試みた。手話単語は手形・動作・運動平面の 3要素に分け、HMM は動作のみを認識することとした。認識対象の動作は手話単語を構成する動きの基本的単位となるもので、以後これ

を基本動作と呼ぶ。また手形についてはベイズ法を使い、その形状を認識することにし、運動面に関してはKL法で得られたデータの分散が最大となるベクトル情報を代用する事にした。

4 基本動作認識実験

HMMは学習によってそのパラメータを変化させるため学習に用いるサンプルの数や種類によって、その認識率は大きく変わる。そこでまず、学習サンプル数の変化が認識率にどの様に影響するかを調べた。その結果、特定話者での学習・認識実験では、「直線」「半円」「停止」が70%前後、往復が60%弱でほぼ安定となった。しかし「円」や「その他」はまだ不安定な認識率であった。

5 手話単語認識実験

基本動作認識実験で学習したHMMとベイズ法、それにKL法による運動面の情報を単語ごとに集め、単語認識用辞書を作成した。作成した単語辞書を用いて、単語の認識実験を行なった。その結果、単語辞書を作成した話者での認識実験(特定話者)では最高で62.8%の認識率が得られた。一方、HMMの学習にも単語辞書作成にも関わっていない全くの未知話者では認識率は最高で19.6%とかなり低い値となった。このような結果になった要因としては、HMMによる基本動作の区分けが少な過ぎる事が挙げられる。また、手形や運動平面に関しても、個人の癖などをあまり考慮に入れていない事が誤認識の要因になったと考えられる。

6 結論

本論文では基本動作にHMM、手形にベイズ認識法、運動面にKL法を用い、その情報を統合して手話単語を決定する認識システムを構築し、検証を行なった。その結果、基本動作が少な過ぎたためか、HMMでの動作認識率が悪く、特に、未知話者では単語認識率が非常に悪い結果となってしまった。この結果より、手話280単語を識別する為には、もう少し細かな動作分けが必要である事が分かった。また、手形や運動平面に関しても、個人に左右されない手法を検討する必要があると思われる。