

手話の音韻構造を考慮した 手話認識における音韻重みの最適化

杉 果穂 酒向 慎司 北村 正
名古屋工業大学工学部情報工学科

1. はじめに

手話は視覚言語の 1 つであり、手や指の運動から成る手指動作と、視線や表情などで表される非手指動作の 2 つを用いて意思伝達が行われる。神田氏らが定めた日本手話音韻表記法[1]によれば、手話は主に手の形状、位置、運動の 3 要素から成り立つと言われている。

このような背景から、自動手話認識においても音韻構造を考慮することが望ましいと考えられ、有賀により手話の言語的特性を考慮した手法が提案された[2]。この研究では、要素間の重要性を重みによって調整することができ、認識対象単語全体に対して適切な重み(位置:運動:局所=0.25:0.60:0.15)が設定されていた。しかし、手話単語ごとに、3 要素の重要度は異なると考えられる。そこで本研究では、手話単語ごとに適切な重みを付与することによって認識率が向上するか検討する。

2. 音韻構造に基づいた手話認識

従来研究では、手話単語の動画像から手の形状、位置、動きの特徴抽出を行い、HMM(隠れマルコフモデル)を用いて学習と認識を行う手法が提案されている。先に述べたように、単語ごとに重要となる音韻要素が異なるため個別の重みを設定することが望ましい。しかし、適切な重みを単語別に設定することは現実的ではない。そこで有賀らの研究でも言及されているように単語間の類似性に注目する。つまり、一部の単語間に類似した共通の重みが存在すると考え、事前に定めた単語重みのクラスタリングによって代表的な重みを分類する手法を提案する。

2.1. 音韻重みのクラスタリング

従来研究の方法に基づき各単語ごとに認識率が最大となるさまざまな重みの組み合わせから最適な重みを定める。これらの重みの集合をクラスタリングすることで代表的な重みを決定する。本研究では、 k -means アルゴリズムを用い、各クラスタの重心を代表重みとする。

3. 実験

3.1 実験条件

クラスタリングによって得られた代表重みを用いて単語認識の性能が向上するか検討する。最適重みは 3 要素の重みをそれぞれ 0.1 刻みで変化させ、66 通りの重み付けを試みた。クラスタリングでは、クラスタ数を 1~7 で行い、それぞれの場合での代表重みにおいて単語認識率を計算した。ベースラインとして従来研究の認識

表 1. 実験条件

単語数	100 単語×3 セット
実験データ	学習用:100 単語×6(3 名×2 回)
	認識用:100 単語×2(1 名×2 回)
	Leave-one-out 法(4 セット)

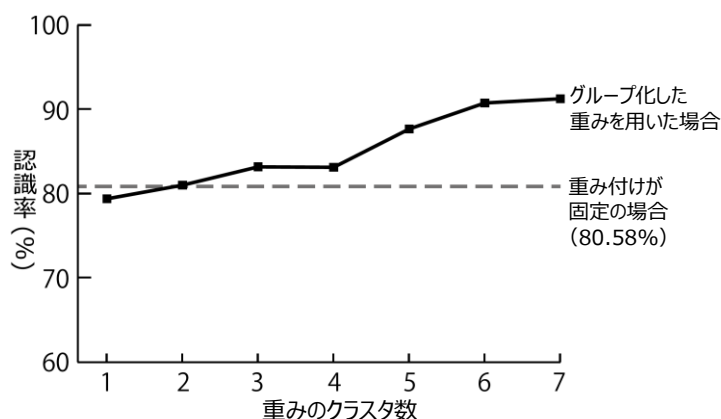


図 1.クラスタリングされた代表重み別の単語認識率

率(重み付けが一定である場合)の結果と比較する。実験に用いた単語セットを表 1 に示す。

3.2 実験結果

図 1 より、クラスタ数が 2 以上のとき従来法に対して認識率が向上していることがわかる。つまり 3 要素の重み付けを、認識させたい手話単語に応じて変化させることにより、認識率が向上するということが確認できた。

4. むすび

本研究では、代表的な重み付けを行うことにより認識率が向上することが確認できた。この結果から単語ごとに重要視する要素が異なることが分かった。これらの知見を踏まえて、事前に重みを求めることなく最適な重みの自動推定を行うことが今後の課題となる。

謝辞

本研究は科学研究費補助金(課題番号:25350666)の補助を受けた。

参考文献

- [1] 神田和幸, 中博一, “日本手話の音韻表記法”, 手話学研究第 12 巻, pp.31-39, 日本手話学会, 1991.
[2] 有賀光希, 酒向慎司, 北村正, “HMM に基づく人物や動作の多様性を考慮した手話認識手法”, 電子情報通信学会技術研究報告(WIT), Vol.110, No.53, pp.55-60, 2010