

手話認識における手話の実データを用いない手型認識の検討

粟田 渉平 酒向 慎司 北村 正

名古屋工業大学大学院情報工学専攻

1 はじめに

手話の通訳支援を目的としたコンピュータによる手話認識の研究が進められている。手話認識において、手話単語のモデルの構築方法は確立しておらず、手話単語単位でモデルを構築する方法では、大語彙化に伴い単語モデルが増大したり、学習データに存在しない未知語を認識できないという問題がある。

波多野らの研究 [1] では、手話辞書の表記法に基づいて手話単語を手指動作の 3 要素である手形状（手の形・向き）、手の位置、手の動きの組み合わせによって定義し、単語認識を行った。3 要素の中でも手形状は、種類が多いことに加えて手首や指の角度など自由度が高いため、認識が難しく、手形状の認識率の低さが課題であった。また、手話データの不足に対応するため、実データではなく各要素を表現した擬似的なデータを収集し、手話単語を認識した。一部の擬似データに実データを加えて手形状認識を行ったところ、認識率が向上することが示唆された。このことから、擬似データを使用するには、実データにより類似した擬似データが必要であると考えられる。そこで本研究では、手型の認識において、実データを使用する場合と擬似データを使用する場合を比較する。また、手型認識精度の向上のため、Kinect v2 から得られる手の先と手の中心の関節座標を利用して手の傾きを推定し、テンプレートの候補を削減する手法を検討する。

2 手型の認識

手の周辺の深度画像のテンプレートマッチングを用いて手型の認識を行う。手話動作中で、手の速さが最も遅いフレームの深度画像を認識対象とする。テンプレートの選択も同様の方法で行う。一つの手型パターンにつき、掌の向きが前・後・上・下・横の 5 方向のテンプレートを用意し、認識対象との類似度が最大の手型を認識結果とする。用意したテンプレート画像の例を図 1 に示す。

3 手の傾きの推定

肩の中央の座標を基準とした手の中心の座標と手の先の座標を利用して、手が垂直であるか、平行であるかの分類を行う。手の傾きを求め、閾値によって分類する。手の傾きが垂直な状態であれば、掌の向き「前」「後」「横」のテンプレートを使用して、平行な状態であれば掌の向き「上」「下」のテンプレートを使用してテンプレ

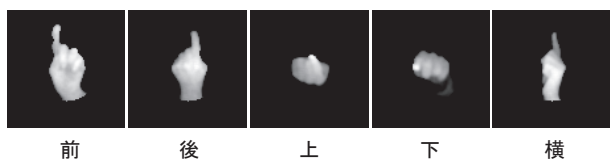


図 1 テンプレート画像の例 (1 型)

表 1 手話単語認識実験：実験条件

手話単語数	245 語 × 2 名 × 3 回
手型	21 種類
対象画像サイズ	120×120 pixel
テンプレート画像サイズ	90×90 pixel
テンプレート画像の枚数	手型 × 掌の向きにつき 6 枚 15 度ずつ回転させる

表 2 手型の認識実験：実験結果

	手の傾き推定		
	なし	あり	手の傾きが既知
擬似データ	33.6 %	23.6 %	30.1%
実データ	53.0 %	42.1 %	45.0%

プレートマッチングを行う。

4 実験

4.1 実験条件

擬似データと実データを使用した手型認識実験を行った。また、手の傾きの推定を利用した手型認識実験を行った。表 1 に実験条件を示す。

4.2 実験結果

実験結果を表 2 に示す。手の傾きの推定により認識率は低下した。使用するべきテンプレートを手の傾き推定により除外したことが認識率低下の原因となった。また、手話動作中の認識対象として選択した画像の掌の向きが、辞書の表記と一致せず適切なテンプレートを使用できない場合もあった。

また、実データの認識率よりも擬似データの認識率が低い理由として、一部の手型で実データに出現するような手の形状の擬似データを用意することができなかったことが挙げられる。手の動作は単語動作の流れや骨格によって制限されるため、表出される掌の向きや指の形が手話辞書の表記の通りとは限らない。収集する擬似データの掌の向きを増やしたり、手話の動作を伴った手の形状を手型の擬似データとして収集することで、このような実データとの差を小さくすることができると考えられる。

5 今後の課題

今回用意した擬似データは実データの量よりも少なく、擬似データの有効性の検証には、様々な規模の擬似データによる実験が必要である。擬似データを拡張し、擬似データの有効性を検証することが今後の課題である。

参考文献

- [1] 波多野ほか. Kinect v2 による手話動作の 3 要素に基づく実時間手話認識. 信学技報, 福祉情報工学, 2016.