手話認識用データベースの構築と手話認識における動作画像領域の重要度の検討

宇都宮　彼方†　岡安　優奈††小澤　辰典†††　田中博††　西村広光†
† 神奈川工科大学情報学部情報メディア学科　†† 神奈川工科大学情報学部情報工学科
††† 神奈川工科大学大学院情報工学専攻

1. はじめに

画像認識を利用した手話認識システム[1]の開発を行ってきているが実用に十分な性能に至っていない．本検討ではシステムの認識精度向上のため，意図的に情報量を削減させた手話動画を用い動作画像領域の重要度の検討を試みた．また手話動画データベースを構築した．

2. 手話動作画像領域の重要度の検討

本検討は欠損した情報にある特徴に着目し，意図的に情報を削減させた手話動画を用いて認識実験を行った．

2.1 人間による評価

手話動作のどの情報が重要であるのかを調べるため，人間が手話認識を行う上で重要とする情報を調査した．評価方法はフレームに対し部位情報の削減，モザイク処理，フレームの削減の処理を行った 25 単語の手話動画を，日常的に手話を扱う方に提示した．

2.2 システムによる評価

システムでの認識において手話動作の小さな動きが必要な情報であるのか調査した．手法としてフレームごとの手首の重心を 1 つ前のフレームと比較し，変化量が小さいフレームに対し前フレームの重心を維持することで小さな動きを削減した．評価は「アトピー」，「おしっこ」，「ガン」，「コンタクト」，「喘息」の 5 単語の手話動画を用いた．

3. 実験結果

表 1 には人間による評価結果を示す．

表 1　人間による評価結果

	正答	誤答	未解答
部位情報の削減	45	1	4
モザイク処理	43	1	6
フレーム情報の削減	49	0	1

結果から人間の手話認識において部位情報の削減による表情の欠損は影響を受けないことが確認できた．

表 2 は以下の条件でのシステムによる評価結果を示すものである．

1. 削減を行う前の動画
2. 横の小さな動きを削減した動画
3. 縦の小さな動きを削減した動画

表 2　システムによる評価結果

1 の正答率	2 の正答率	3 の正答率
73%	57%	7%

結果からシステムでの認識には，小さな動きであっても重要な特徴なのだと確認することができた．

4. 指標データベースの作成

手話認識システム[1]の精度向上にはデータベースが必要不可欠であるため，表 3 に示す手話単語を収録した手話データベースを構築した．

表 3　収録手話単語 55 種

薬を飲む	頭痛	汗だく	骨
ダイエット	酔う	冷や汗	喘息
顔が赤い	呼吸	車椅子	毒
おしっこ	盲腸	吐き気	ハゲ
カテーテル	発熱	昼寝	病気
インフルエンザ	命	風邪	肥満
ズキンズキン痛い	腰	糖尿病	捻挫
イライラ	汗	医者	ガン
アトピー	近視	限界	骨折
発達障害	裸	検出	喫煙
コンタクト	鼻水	体調	摘出
たくましい	腎臓	禁煙	涙
ギリギリ痛い	虫歯	耳鼻科	眼鏡
産婦人科	顔色	脳卒中	

手話単語は資料[2]に示す"SmartDeaf"の医療区分に基づいたものであり，表 4 にデータベースの規模を示す．

表 4　収録

単語数	55
手話者人数	9
サンプル数	30
総データ数	14,850

データベースには複数回手話動作を行う手話動画が収録されており，1 回の動作ごとに分割するためのフレーム数を記録したファイルを作成した．そのため総データ数 14,850 個の手話動画を容易に出力することが可能である．

5. まとめ

手話認識システム[1]の精度向上に向け，情報量の削減に着目した動作画像領域の重要度の検討を行った．結果から表情から手話の認識に繋がる情報は少なく，手の小さな動きが認識に影響することが確認できた．

参考文献

[1] 小澤辰典ほか，“DP マッチングから得られた認識対象のクラス絞り込みによる HMM を用いた手話認識精度の向上”，信学総大 2017
[2] 株式会社ケイ・シー・シー，“SmartDeaf”，https://www.smartdeaf.com，[参照日 2018.01.11]

Copyright © 2018 IEICE