

HMMを用いた動きのある連続指文字認識

大野 雄士朗[†] 鈴木 秀智[†] 太田 義勝[†]
[†] 三重大学大学院工学研究科情報工学専攻

1. はじめに

現在、難聴に苦しむ人は全国で約30万人、軽度の難聴者は600万人以上と言われている。聴覚障害者は、音声言語によるコミュニケーションが困難であり、手話を用いる事で情報交換を行っている。近年、聴覚障害者と健聴者との会話の支援や、人とコンピュータの間での対話を目的とするヒューマンインターフェースの研究が盛んに行われている。また、手話の一種に指文字がある。指文字とは、手の形を文字言語に対応させた、視覚言語の一要素である。これは、五十音を表現することができ、語彙の数が格段に少ない手話に比べ、固有名詞を始めとした、手話単語にない単語の表現が可能になっている。しかし、これらを理解できる人は大変少なく、聴覚障害者と健聴者の円滑なコミュニケーション環境が充分でないというのが現状である。手話および指文字の認識の研究は古くから行われている。手話については手話単語ごとに構築したHMMを用いて認識する手法[1]が主流である。指文字については掌や指の形状に基づく認識[2]がよく研究されている。

2. 従来手法と問題点

本研究ではHOG特徴量とAdaBoostを使った大分類と、大分類によって分類されたグループごとの細分類による静止指文字認識[3]を行っている。この実験では静止指文字41種類に対し、84.2%の認識率を得ている。また、これまでの研究により、動きのある指文字を含めたすべての指文字に対して、動作特徴による分類を行い、認識結果(図1)を得た。問題点として、引く動作と静止を分類する有効な特徴量を得られていないこと、一文字の認識のための特徴量を抽出しており、連続的に指文字認識を行うことを想定していないことが挙げられる。また、グループ7の認識率が0%である原因として、すべての文字がグループ3と誤認識されていることが挙げられる。

グループ	適合率(%)	再現率(%)
1	98.4	97.3
2	100	92.7
3	63.9	92.8
4	88.9	84.2
5	69.2	100
6	88.9	88.9
7	0	0

図1: 動作特徴分類結果

3. 提案手法

連続的に指文字を認識する方法として、隠れマルコフモデル(Hidden Markov Model, HMM)とビームサーチがある。本研究では、撮影動画の各フレームから、手領域を抽出し、特徴量としてHOGと全フレームからの移動座標(ΔX , ΔY , ΔZ)を抽出し、ベクトル量子化を行ったものを出力シンボルとし、状態数、初期状態遷移確率を決定し、文字ごとのHMMを作成することを提案する。 ΔZ 座標の抽出により、引く動作の特徴量を、また各フレームからのHOGにより手の形に関する特徴を得ることができる。

4. 実験データ

実験動画の撮影にはMicrosoft社のKinectを用いる。Kinectは赤外線センサーを内蔵しており、Depthデータ(図2)を得ることが出来る。これにより引く動作の識別を行うことが可能である。実験データには640×480[pixel]の2Dカメラの撮影による動画像および人物と手領域追跡を行うDepth動画像を用いる。

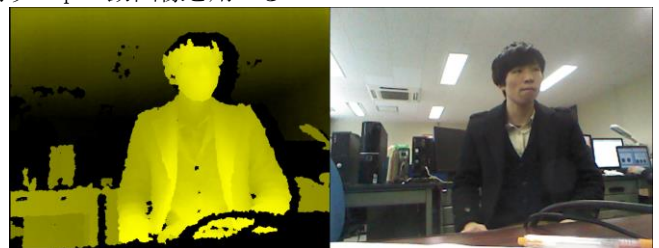


図2: デプス画像の例

5. 評価方法

一文字に対して学習を行い、一文字の認識および単語の認識を行う。一文字の認識には交差検証を用いて適合率と再現率により評価を行う。

6. まとめと今後の方針

連続指文字認識のためのモデル作成という目的でHMMによる文字モデルの作成、特徴量、引く動作の識別のための撮影機器、実験環境を検討した。今後の方針として、ベクトル量子化のクラス数の決定、HMMモデルの作成が挙げられる。

参考文献

- [1] 松尾, 他: “HMMを利用した画像処理による手話認識のための特徴抽出および状態分割”, ヒューマンインターフェース学会論文, Vol NO. 1, 2013立命館大学
 [2] 三宅, 他, “距離画像を用いた動きのある指文字認識に関する研究”, 修士論文, 筑波技術大学, 2012
 [3] 菊田, 他 “類似文字を考慮した段階的な指文字認識”, 修士論文, 三重大学, 2012. 02.