

手指の動きをともなう指文字の非接触認識手法の検討

若月大輔¹⁾, 三宅太一²⁾, 内藤一郎¹⁾

筑波技術大学 産業技術学部 産業情報学科¹⁾

筑波技術大学大学院 技術科学研究科 産業技術学専攻²⁾

キーワード：指文字認識, 距離画像, 赤外線 TOF カメラ, 聴覚障害者

1. まえがき

音声認識に関する技術が進歩し、音声入力機能を備えた情報機器が増えてきた。音声入力はインタフェースに触れることなく機器を操作することができ、視線をそらせない、手が使えないような場面で活用されている。また、直接インタフェースの操作に不慣れ、あるいは不可能な高齢者や障害者の入力手段としても注目されている。

しかし、音声入力は全てのユーザにとって必ずしも有効な入力インタフェースではない。聴覚障害者の多くは発話障害をもち、明瞭な発音が困難な者も多い。健聴者と比較して不明瞭、かつ個人差が大きいため先に述べた音声入力をまったく活用できない場合もある。今後、音声認識技術がさらに進歩し、音声入力による操作を前提としたインタフェースが一般化すると、聴覚障害者が利用できない機器が登場してくる可能性が高い。

そこで、音声入力を代替する手段として、指文字による入力インタフェースを実現するために、情報機器に指文字を認識させるための基礎的な検討を行った。先行研究でも指文字認識を試みた報告があるが、それらの多くは濁音や半濁音などの動きをともなう指文字を対象としていない。本研究では、手型が静止した指文字に加えて動きをともなう指文字も認識するための方法を提案し、その精度について評価を行った。

2. 動きをともなう指文字とは

指文字は書記言語（日本語の場合はひらがな）の各文字に対して、手型が1対1で対応している。そのうち、図1に示すような濁音、半濁音、長音、小書き文字、促音は手型を固定したまま上下、横、および手前方向に移動させて表現する。本研究ではこれらの指文字を、動きをともなう指文字と呼び、認識させるための方法を提案した。



図1：動きをともなう指文字

3. 指文字認識の概要

動きをともなう指文字を考慮した認識処理の流れを次に示す。

- (1) 指文字の距離画像撮影
- (2) 距離画像中の手領域抽出
- (3) 手領域の手型識別
- (4) 手領域の動作検出

処理(1)では距離画像センサとして、高速に距離画像を撮影できる赤外線 TOF カメラを使用した。撮影した指文字の距離画像から処理(2)で手領域を抽出する。ここでは距離画像から手の体積を計算し、手領域のみを自動で抽出できる方法を考案し実装した[1]。処理(3)の手型識別では、抽出した手領域の距離画像について HLAC で特徴ベクトルを求め、SVM で手型を識別させる[2]。また、抽出した領域について動きをともなう指文字かどうか処理(4)で検出して、手型と動作を組み合わせる指文字を認識する。

4. 動きをともなう指文字認識

手型を動かさない指文字と動きをともなう指文字の距離画像を比較すると、手型は同じであっても距離画像に違いがある。図2に「ひ」と「び」の例を示す。これは対象の動きに起因するぶれ（動作ノイズ）であると考えられる。動きをともなう指文字の距離画像には必ず動き方向の動作ノイズが発生するため、これらを画像の特徴と考えると、前節の処理(3)の手領域の手型

識別における学習データとして活用する方法を考案し、動きをともなう指文字を認識させた。

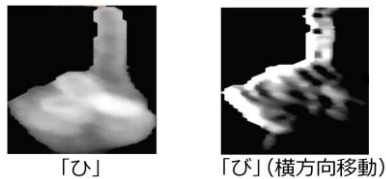


図2：動きをともなう指文字の距離画像

前節の処理(4)では、手領域の重心の3次元位置の履歴に対して主成分分析を行い、第一主成分より動き方向、そのスコアから動いた量を推定して手領域の動作の有無を検出する方法を考案し実装した。図3は「き」を濁音方向に動かした場合である。動き方向とその量から求まる動きベクトルと、濁音などの動き方向の単位ベクトルとの内積値が大きい方向を選択する。

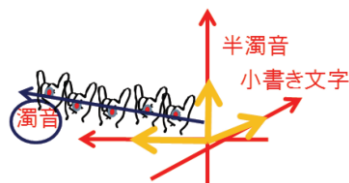


図3：指文字の3次元空間中の動作検出

手領域の手型識別結果と動作検出結果を組み合わせ、動きをともなう指文字の認識を行う。例えば、手型が「き」か「ぎ」と識別とされたとき、動作検出で動きがないと検出された場合は「き」、濁音方向の動作が検出された場合は「ぎ」として認識する。

5. 認識実験とその結果

動きをともなう指文字の認識精度を調査するために、動きのない指文字として「あ」「か」「さ」「は」行、および「や」「ゆ」「よ」の23文字、動きをともなう指文字として「が」「ば」行、「ゃ」「ゅ」「ょ」13文字を対象にして認識実験を行った。被験者1名について各指文字を20枚ずつ撮影したものを10セット用意し、そのうち5セットを学習データ、他の5セットを評価用データとして認識させた。まず、動作検出(前述3.(4))を使用せず、動作ノイズを考慮した学習データで認識させた場合の精度を調査した。その結果、平均で約0.85の認識率であった。特に、濁音や半濁音などの指文字の認識率が高く、動作ノイズを考慮した距離画像を用いた場合に動きをともなう指文字を認識できることがわかった。次に、動作検出も組み合わせた場合

の評価を行った結果、平均認識率が0.90に改善した。これは動作検出によって動きのない静音の指文字と、動きをともなう濁音などの指文字の判別精度が改善されたことが理由である。これらの実験から本方法は動きをともなう指文字を含めた指文字認識が可能であることが明らかになった。

6. まとめ

本研究では、音声入力を代替する入力手段として、指文字入力インターフェースを実現するために、動きをともなう指文字を認識するための基礎的研究を行った。距離画像から求まる体積から手領域を自動抽出する方法、動作ノイズを考慮した手型識別、ならびに手領域の3次元的な動作検出を行う方法を提案した。動作ノイズを考慮した方法では濁音や半濁音などの動きをともなう指文字の識別が可能になることを示し、さらに動作検出を組み合わせることで認識精度を改善できることを明らかにした。

今後の課題は、今回認識対象にできなかった「の」「も」「り」「ん」の手型そのものが動く指文字の認識である。入力インターフェースを構築するには、指文字で表現される文の認識が必要になるため、指文字間のわりや単語などの認識について研究を進めたい。

本研究は平成24年度教育研究等高度化推進事業A競争的教育研究プロジェクト事業の助成を受けて実施した。成果の一部を、筑波技術大学テクレポ、HCGシンポジウム2012、ならびに第8回日本聴覚障害学生高等教育支援シンポジウムにて公開した[1][2][3]。

本テーマに関する成果の学会発表等

[1] 三宅太一, 若月大輔, 内藤一郎. ”

三宅太一, 若月大輔, 内藤一郎 距離画像を用いた動きのある指文字の非接触認識手法の検討 HCGシンポジウム2012. 2012;():p.270-275.

[3] 三宅太一, 若月大輔, 内藤一郎. 距離画像を用いた動きのある指文字の非接触認識～指文字入力インターフェースの実現を目指して～. 第8回日本聴覚障害学生高等教育支援シンポジウム・ランチセッション「聴覚障害学生支援に関する機器展示」. 2012;():p.50.