

運筆音を用いた文字認識の検討

屋内 律澄[†] 川波 弘道[†]
[†] 津山工業高等専門学校

1. はじめに

運筆音とは運筆音は文字を記述する際に筆記具によって発生する擦過音である。現在、紙に書かれた文章を認識する光学文字認識や人間の声を認識する音声認識の研究は盛んに行われている。しかし、運筆音を用いた認識の研究は少ない[1]。筆者らは、運筆音により光学機器が不要な文字認識が可能になると考え、その有効性を検討する。

本稿では、文字の違いと文字を書く素材によって運筆音がどの程度異なるか調査した結果を報告する。

2. 運筆音による文字認識

2.1 データ収録方法

スチールデスクにおいた紙の上と、スチールデスクの表面に直接、ひらがなをシャープペンシルで運筆した音を収録した。また、自由に記述すると筆癖にも個体差が生じると考えられるため、図1上段のようなひらがな練習用プリントに印刷された文字をなぞることで個体差の抑制を行った。運筆音は SONY, PCM-M10 を用いて、48[kHz], 16bit で収録した。

2.2 特徴量

次の3つの特徴量を用いて類似度を調査する。なお、DTWによるフレームの対応付けを事前に行っている。

- (1). 振幅スペクトル: 波形に短時間フーリエ変換を適用し、各フレームの振幅スペクトル列を比較する。
- (2). 振幅スペクトルにおけるセントロイド: 波形に短時間フーリエ変換を適用し、各フレームの振幅スペクトルのセントロイドを比較する
- (3). メル周波数ケプストラム係数(MFCC) 20 次元

3. 実験

紙の上で記述した「あ」の特徴量を基準とし、新たに紙の上で記述した「あ」から「お」までの運筆音とデスクの上に直接記述した「あ」の運筆音との類似度を求めた。各比較対象の類似度を表1に示す。今回の実験では振幅スペクトルにおけるセントロイド、MFCC は同じ条件で録音された「あ」の音と高い類似を示しているが別の文字に対する類似度も0.3以上を示しており、誤認識を行う可能性があると考えられる。そのため、振幅スペクトルが比較的頑健である傾向が見られる。また、対象が紙と机とで類似度は著しく低下している。

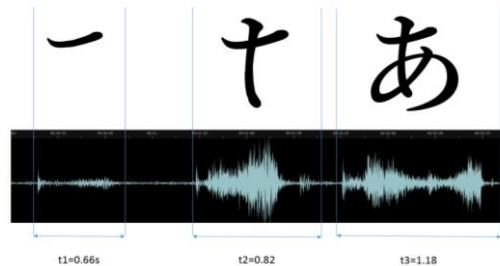


図1 運筆音における有音区間と無音区間

表1 各文字条件における類似度

比較文字 (記述環媒)	振幅スペクトル	振幅スペクトル (セントロイド)	MFCC
あ(紙)	0.7332	0.9186	0.8151
あ(机)	0.0652	0.1086	0.5447
い(紙)	0.0482	0.8852	0.0134
う(紙)	0.2787	0.8122	0.3501
え(紙)	0.1243	0.4162	0.4486
お(紙)	0.0164	0.5264	0.2295

4. まとめ

本稿では、運筆音を用いた文字認識の検討を行った。今回の実験では6パターンにおける比較を起こしたが、ひらがな全種類やひらがな以外の認識を行う場合、計算量が多くなる。また、運筆音を用いる認識の場合、文字を書き込む媒体によって発生する音が異なることがあるが、認識を行うために媒体の識別を不要とすることが望ましい。

そこで、閾値を設定し図1のように有音区間と無音区間に分割することで、ひらがなの一画を記述するのにかかる時間を特徴量として機械学習の設計を行うことを検討する。この手法のメリットとして、音を分割する時点で画数を求めることができる。画数ごとにグループ化することで、画数の異なるグループの計算が不要となり、高速な計算につながると考えられる。

参考文献

- [1] 北野真示・西野隆典・成瀬央, 「HMMを用いた運筆音からの手書き数字認識」, 信学技報, vol.113, no.346, PRMU2013-88, pp.121-125, 2013