

オンライン手書き数式の行列構造認識

日暮 大輝†

慶應義塾大学大学院
理工学研究科開放科学専攻

櫻井 彰人‡

慶應義塾大学理工学部

1 目的

タブレット端末を利用した手書き数式認識で、行列構造を認識する。また、行列の構造認識に適した数式認識アルゴリズムを提案する。

2 既存研究の紹介

数字の隙間の大きさの比を用いて閾値で判断し、行列の構造認識を行う手法が豊住らにより提唱された[1]。この手法では、列ごとに行為の長さが異なったり、行ごとに列の長さが異なったりする誤認識をすることがある。

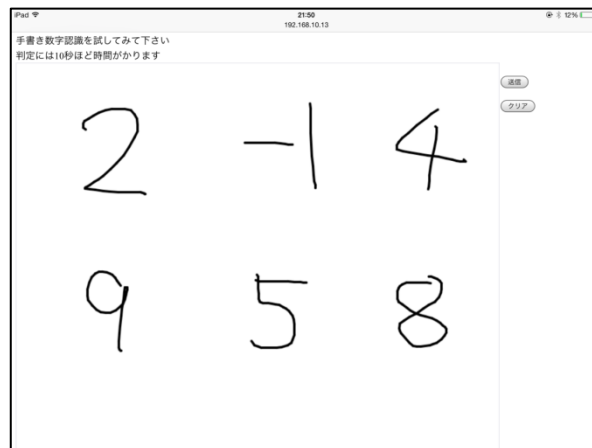


図2 行列の入力例

3 提案手法

本研究では、既存研究の行列モードに対応する部分を提案する。図2の $\begin{bmatrix} 2 & -1 & 4 \\ 9 & 5 & 8 \end{bmatrix}$ を例とする。

3.1 ストロークをクラスタリング

各ストロークの中心座標をx軸に射影する。最短距離法でクラスタリングを行い、生成されたクラスターを左から順に1列目、2列目、3列目としていく。同様の処理をy軸でも行い、行を決定する。クラスタリングの終了条件は、結合するときの最短距離の差が最大になるときとする。すなわちクラスター数は少なくとも2になる。

図3は入力例をx軸に射影してクラスタリングした際にできるクラスターである。表1はクラスターの結合回数とそのときの最短距離、直近の最短距離同士の差をまとめたものである。結合回数6回と7回ときの最短距離の差が128.5で最大である。そのため結合回数6回までクラスタリングする。この場合、正しく3列の行列だと認識できた。同様の手順で2行の行列と認識できた。

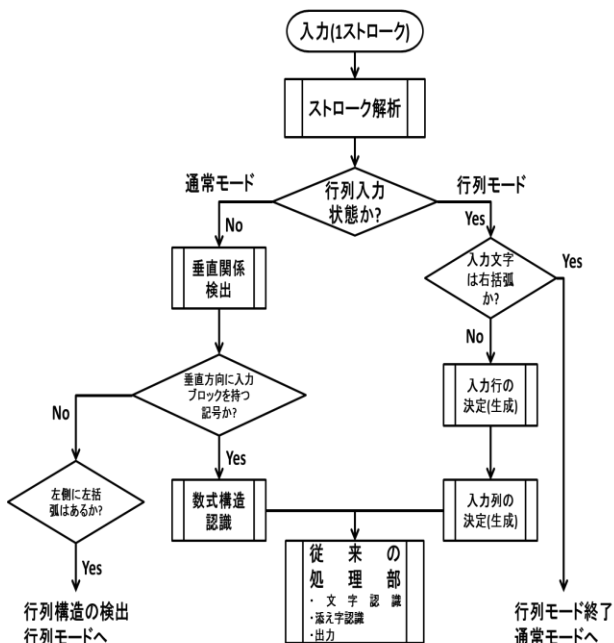


図1 豊住ら[1]のアルゴリズム

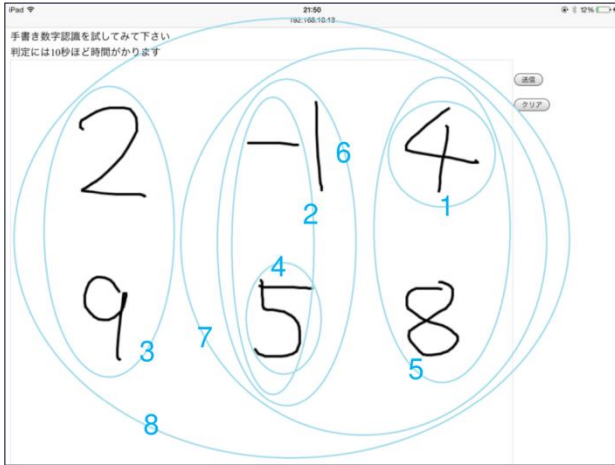


図3 最短距離法クラスタリングにより生成されたクラスター

表1 クラスターの結合回数と最短距離

クラスターの結合回数	1	2	3	4	5	6	7	8
最短距離	2.0	7.0	11.5	13.5	15.5	52.5	181.0	256.0
最短距離の差		5.0	4.5	2.0	2.0	37.0	128.5	75.0

3.2 クラスターから行列構造の認識

このクラスタリングの終了条件だと、クラスター数は2以上になるため、1行の行列や1列の行列では例外処理が必要となる。 $\begin{bmatrix} -1 \\ 4 \\ 8 \end{bmatrix}$ を例とする。クラスタリングにより、列方向は3つのクラスターに分けられ、 $\begin{bmatrix} - & & 1 \\ & 4 & \\ & & 8 \end{bmatrix}$ と認識される。このように行列に空白の成分がある場合は1列あるいは1行の行列であるとする。ここで各ストロークのクラスタリングは終了する。以降は各クラスターで数式認識を行う。

3.3 行列の成分ごとに数式認識

1ストロークが1つの数字を表現していると仮定して予測する。その後以下の2つの結合条件を満たす2ストロークを結合して1つの数字を表現しているとする。

結合条件1

連続する2ストロークが1つの数字を表現していると仮定して予測し、予測確率が85%を超える

結合条件2

重なった部分がある

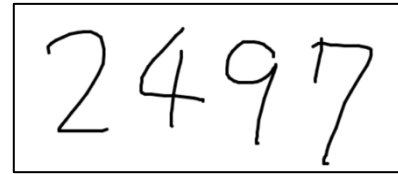


図4 数式の入力例

表2 予測した数字と予測確率

ストローク番号	1	2	3	4	5	6	
1ストロークずつ	予測した数字	2	[1	9	1	7
	予測した確率	95%	23%	94%	98%	93%	41%
2ストロークずつ	予測した数字		4	4	0	+	7
	予測した確率		28%	96%	30%	23%	55%

図4の2497を例とする。第2ストロークと第3ストロークが1つの数字を表現していると仮定して予測したときの予測確率は96%で、閾値の85%を超えている。そこで第2ストロークと第3ストロークを結合する。

また、第5ストロークと第6ストロークは重なった部分があるため、結合する。その結果、2497と正しく認識できた。なお、数字の予測は多クラス確率推定SVMを用いた。クラスは0,1,2,3,4,5,6,7,8,9,+,-,[]の14クラスとして、3284件の学習データを元に学習した。学習データは10人以上の学生がタブレットまたはスマートフォンで入力して集めた。

4 まとめ

$\begin{bmatrix} 12 \\ -1 \end{bmatrix}$ や $\begin{bmatrix} 46 \\ 23 \end{bmatrix}$ のような2行1列の行列を $\begin{bmatrix} 1 & 2 \\ - & 1 \end{bmatrix}$ や $\begin{bmatrix} 4 & 6 \\ 2 & 3 \end{bmatrix}$ のように2行2列の行列として認識してしまう。これを改善する工夫が求められる。

また、数式認識は現在数字認識しか対応していないが、分数や指数などの各種構造認識に対応したり、より多くの機械学習手法と比較して認識精度をあげたり、改善の余地がある。

5 参考文献

[1]豊住健一, 鈴木隆広, 森健策, 末永康仁, "オンライン手書き数式認識システムにおける行列要素の位置関係に基づく行列認識機構", 電子情報通信学会誌, vol.J86-D-II, no.9, pp.1278-1285, 2003