

二次元移動方向に基づく 空中手書きアルファベット文字認識に関する研究

西田 好宏*

An Aerial Handwritten Alphabet Character Recognition Using Two-dimensional Moving Direction

Yoshihiro Nishida*

This paper describes a method to recognize a character handwritten in the air. This recognition method uses the motion direction instead of positions of the device. It also doesn't use the information of pen up and down. It selects and orders character candidates with DP (dynamic programming) matching. We prototyped an application software that recognize a character by using mouse movements in order to study algorithm and usability. In this time, we added categories of Alphabet character. We achieved correct recognition rate about 90% for categories of Hiragana and Alphabet characters by using two-dimensional moving direction.

Keywords: 文字認識, 空中手書き入力, DP マッチング, 二次元移動方向, アルファベット

1. はじめに

クラウド・コンピューティングの普及により、いつでも、どこでも、ふと思いついたアイデアなどをメモとしてクラウドに保存したり、出来事をつぶやきあったりする機会が増えている。当然、そのためにはテキストデータの入力が必要で、両手が空いていなくても片手で簡単に文字を入力することができたら、益々便利になると考えられる。また、デジタルサイネージが普及し、単に情報を流すだけではなくインタラクティブな操作や検索が可能になっているが、不特定多数の人が操作するため非接触で操作や文字を入力したいという要求がある。

また一方では、KINECT センサが発売されてジェスチャーや人物のトラッキングを実現することで、空中での手の動きを二次元平面上にマッピングすることが容易になって来ている。

そのような背景の中、絶対位置情報を利用せずに相対的な移動方向情報のみを利用して、1文字単位で一筆続け書き文字として認識する方法を試みている。また、空中で書いた筆跡の代わりにマウス入力を用いて、マウスで書いた文字をリアルタイムで認識するアプリケーションを作成し、ひらがなと数字の認識率が向上することを確認した。^{[1][2]}

今回、新たに小文字アルファベットと大文字アルファベットを辞書データに追加し、高い認識率が得られる筆記方法、および辞書データについて検討した。

* 電気電子情報工学科

2. 文字認識エンジン

空中手書き文字入力を実現するためには、下記の3つの処理が必要である。

- ① 動きを検出するセンサ処理
- ② 動きから文字を認識する文字認識エンジン処理
- ③ 認識した文字の変換や確定などの処理

本紀要は、②の文字認識エンジン処理に着目し、二次元平面上に書いた一筆続け書き文字の認識に関するもので、従来のひらがなと数字に加え、新たに小文字アルファベットと大文字アルファベットの認識を可能にした。

2-1. アルゴリズム

本研究では、D Pマッチングを採用している。D Pマッチングとは、一度計算した結果をうまく再利用して、効率的に計算することである。D Pマッチングは、理論的明快さ、アルゴリズム実装の容易さ、学習が不要なうえ、少ない計算量で最適なマッチングが求まるという特長を持つ。そのため、時系列パターン認識において広く利用されている。

D Pマッチングを実現するアルゴリズムの具体例を説明する。D Pマッチングは長さの異なる2組の連続するデータ間の隔たりを「エラー値」としている。「エラー値」には、各データ間の隔たりと伸縮の相違を数字化し、演算、累積加算を行い、最小となるパターンを出力する。

このD Pマッチングを入力コードとすべての辞書データに対して行い、比較した結果から、最も「エラー値」が小さかった辞書データの文字を入力文字と判定する。また、比較データには方向ベクトルを用いているため、方向ベクトルの角度の差に応じて、「エラー値」も大きくなるように設定している。例えば、図1の方向ベクトルにおいて、「a」と「b」の「エラー値」より「a」と「c」の「エラー値」を大きく設定している。

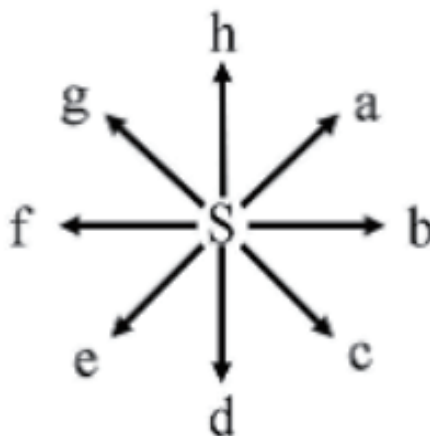


図1. 方向ベクトル

2-2. 評価用アプリケーション

文字認識エンジンを評価するため、文字の大きさ、位置に関係なく、マウスを用いて一筆書きで書いた文字を認識するアプリケーションを開発した。

アプリケーションの画面例を図2に示す。検出した方向コード列と辞書データのコード列をDPマッチングで比較し、エラー値が一番少ない文字を第一候補、次に第二候補、第三候補として表示している。

そもそも空中で文字を書く場合は、人の動作そのものが正確でなく、再現性に乏しい。そのため、認識率の僅かな向上よりも、むしろ間違った時の正しい候補の選択方法や修正方法を便利にして、トータルとしていかに簡単に意図した文字が入力できるかが重要である。そこで、認識結果（第一候補認識率）以外にも、第二候補、第三候補の検出結果を表示して、その選択を簡単に構成を想定している。そのため、本来の第一候補で認識する認識率だけでなく、第一候補から第三候補までの認識を含めた合計認識率を算出、評価を行った。

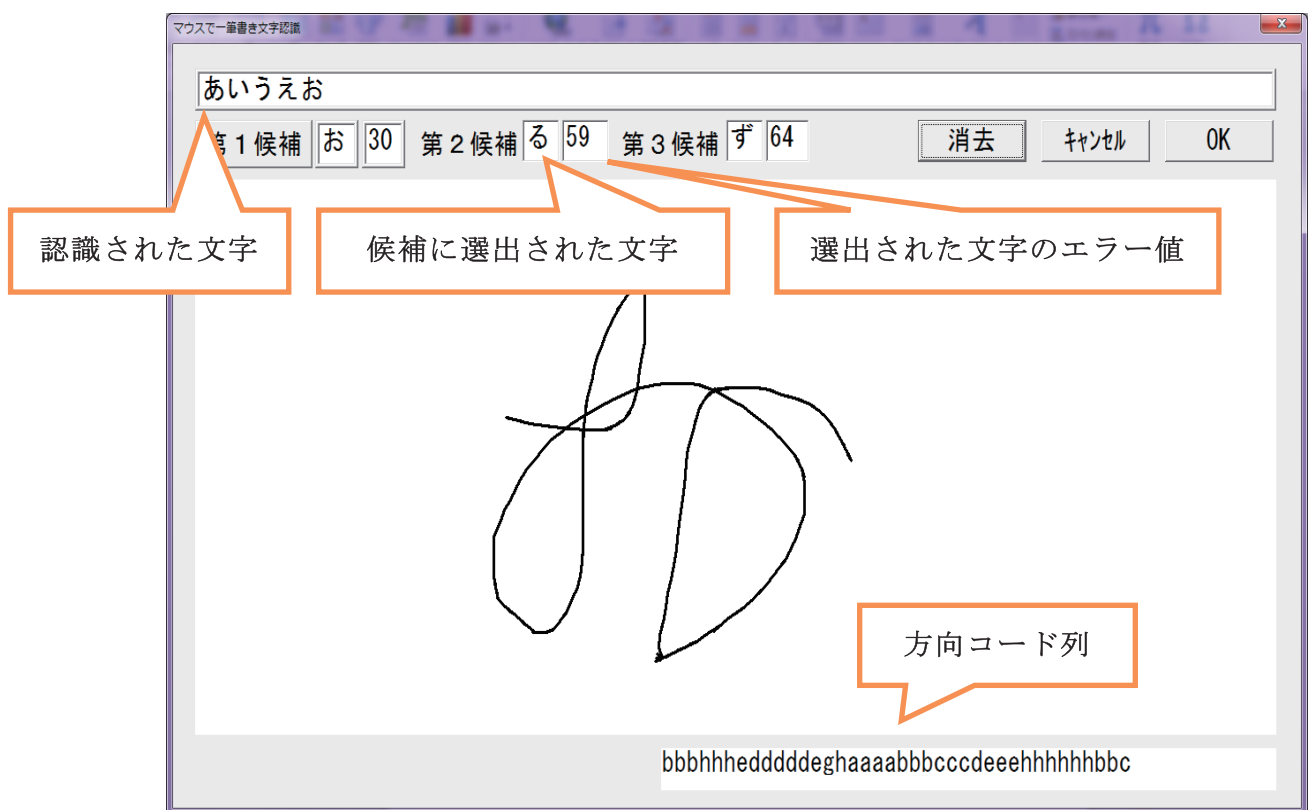


図2. 一筆書き文字認識アプリケーション

3. 提案するアルファベット認識

3-1. 入力モード

認識する文字種類を追加する場合の方法として、入力モードを追加して文字種類に応じて変更することが考えられる。例えば、一般的な携帯電話の文字入力では、ひらがな、数字、アルファ

ベットなどの入力モード、パソコンの場合には、半角英数、全角かな等、複数の入力モードがあり、入力文字に応じて変更して入力している。

しかしながら、入力モードの変更は操作が煩雑になるという問題があるため、今回は入力モードを設けずに、ひらがな、数字、アルファベットのどの文字が入力されても認識できる構成とした。

3-2. 小文字、大文字の区別

上記の通り入力モードを設けない構成にしたため、アルファベット入力において小文字と大文字の区別が必要になる。しかし、「C」「M」「N」等のように筆記体の小文字と大文字の書き方が類似している文字が多いという問題がある。そこで、今回は大文字アルファベットについては小文字アルファベットを連続で二度書くことにより大文字アルファベットと認識するようにした。大文字アルファベットの入力例を図3に示す。

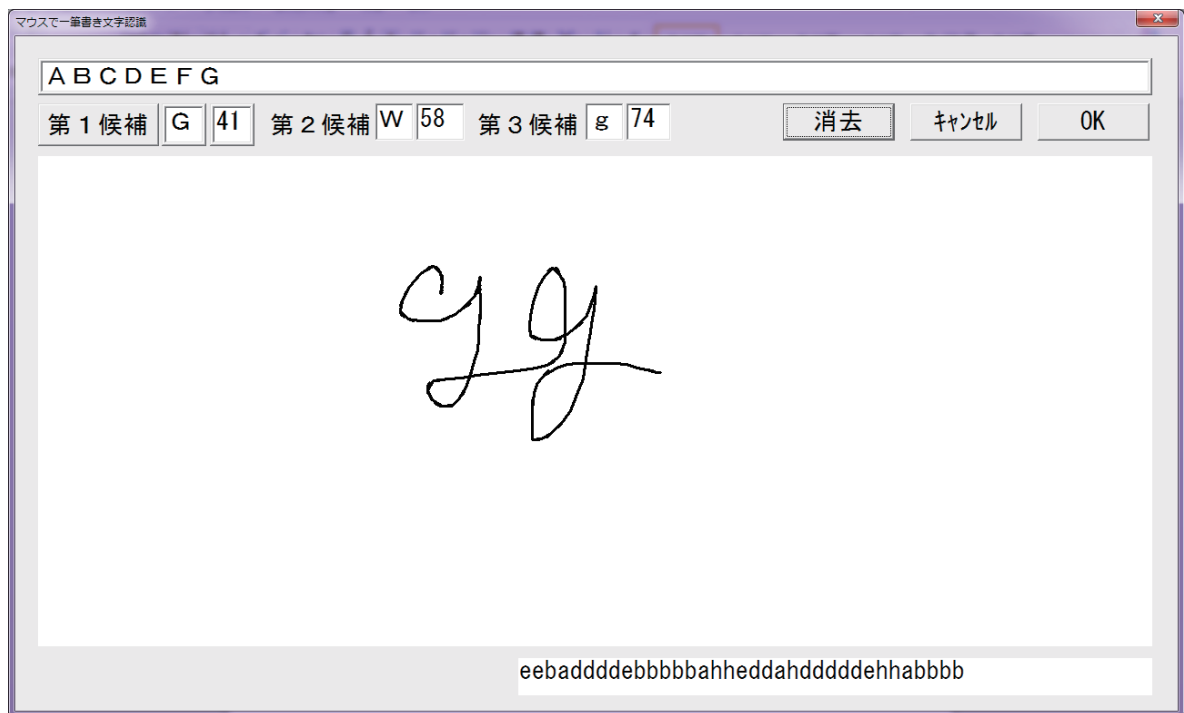
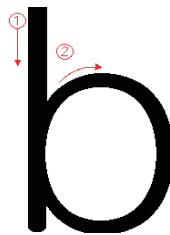


図3. 大文字アルファベットの入力例

3-3. 辞書データ

ひらがなの辞書データは、「書き順シート」「筆順ドリル」「ひらがな練習帳」をもとに作成しており、書き方の個人差は比較的少ない。一方、アルファベットは筆記体で書く人が少なくなり、アルファベットの書き方の個人差は非常に大きいと考えられる。今回は、ブロック体のアルファベット小文字・書き順一覧[3]のデータを基本にして、被験者の書き方により辞書データを修正した。図4にアルファベット「b」の場合の筆記例とその辞書データ例を示す。



b の辞書データ :

ddddddddddddddhhaaaabbcddddddeefffffff

図4. 「b」の筆記例とその辞書例

4. 実験

実験は4人の被験者で行った。実験の基本的な条件は下記の通りである。

- 一文字につき10回書き、また選出された候補の文字を記録として残す。
- 文字の大きさ、筆跡する位置は、アプリケーション上、マウスなら自由である。
- 出来る限り丁寧に書く。
- 実験をする前に、数回試し書きをしてから実験を始める。

4-1. 被験者Aと辞書A

被験者Aが被験者Aの書き方により修正した辞書Aを用いた時の認識結果を表1に示す。アルファベットの方がひらがなよりも高い認識率が得られ、アルファベットの第一候補での認識率は89.0%、第三候補までの合計認識率は99.6%が得られた。また、ひらがなも加えた場合の第一候補での認識率は85.6%、第三候補までの合計認識率は97.3%が得られた。

表1. 被験者Aと辞書Aの認識結果

文字	第一候補	第二候補	第三候補	合計
あ	9/10	1/10	0/10	10/10
い	8/10	2/10	0/10	10/10
う	10/10	0/10	0/10	10/10
a	6/10	4/10	0/10	10/10
b	10/10	0/10	0/10	10/10
c	10/10	0/10	0/10	10/10
A	7/10	1/10	1/10	9/10
B	9/10	0/10	1/10	10/10

C	10/10	0/10	0/10	10/10
アルファベッ ト認識率	463/520 (89.0%)	42/520 (8.1%)	13/520 (2.5%)	518/520 (99.6%)
ひらがなを含 む認識率	1023/1220 (85.6%)	122/1220 (9.1%)	31/1220 (2.6%)	1176/1220 (97.3%)

4-2. 被験者B, C, Dと辞書A

次に、被験者B, C, D, が辞書Aを用いた時の認識率を表2に示す。辞書Aは被験者Aの書き方に合わせているため、アルファベットの第一候補での認識率は84.3%、第三候補までの合計認識率は96.9%と、被験者Aに比べて共に低くなってしまった。

表2. 被験者B, C, Dと辞書Aの認識結果

文字	第一候補	第二候補	第三候補	合計
a	22/30	3/30	4/30	29/30
b	30/30	0/30	0/30	30/30
c	30/30	0/30	0/30	30/30
小文字 認識率	686/780 (87.9%)	66/780 (8.5%)	15/780 (1.9%)	767/780 (98.3%)
A	25/30	4/30	0/30	29/30
B	18/30	4/30	6/30	28/30
C	29/30	1/30	0/30	30/30
大文字 認識率	629/780 (80.6%)	86/780 (11.0%)	31/780 (4.0%)	746/780 (95.6%)
アルファベッ ト認識率	1315/1560 (84.3%)	152/1560 (9.7%)	46/1560 (2.9%)	1513/1560 (96.9%)

4-3. 被験者B, C, Dと辞書X

そこで、被験者B, C, Dが辞書Aを用いた時に共通して認識率が低い文字を抽出し、「a」「d」「e」「n」「o」「x」「D」「G」「I」「O」「T」「V」「X」「Z」の文字を修正した辞書Xを作成した。

被験者B, C, Dが辞書Xを用いて実験した認識結果を表3に示す。辞書を修正したことで、アルファベットの第一候補での認識率は約6%向上し90.6%、第三候補までの合計認識率は2%向上し98.9%になった。

表3. 被験者B, C, Dと辞書Xの認識結果

文字	第一候補	第二候補	第三候補	合計
a	26/30	3/30	1/30	29/30
b	28/30	0/30	2/30	30/30
c	28/30	5/30	0/30	30/30
小文字 認識率	718/780 (92.1%)	48/780 (6.1%)	11/780 (1.4%)	777/780 (99.6%)
A	25/30	3/30	2/30	30/30
B	27/30	3/30	0/30	30/30
C	29/30	0/30	0/30	29/30
大文字 認識率	696/780 (89.2%)	53/780 (6.8%)	28/780 (2.2%)	766/780 (98.2%)
アルファベッ ト合計認識率	1414/1560 (90.6%)	101/1560 (6.5%)	28/1560 (1.8%)	1543/1560 (98.9%)

4-4. 被験者Aと辞書X

さらに、被験者Aが辞書Xを用いた時の認識結果を表4に示す。アルファベットの第三候補までの合計認識率は辞書Aの時と同じ99.6%であったが、アルファベットの第一候補での認識率は辞書Xが被験者B, C, Dの書き方に合わせて修正したにもかかわらず、辞書Aの時より約3%向上し92.3%が得られた。また、ひらがなも加えた場合の第一候補での認識率も約2%向上し88.0%、第三候補までの合計認識率も約1%向上し98.1%が得られた。

表4. 被験者Aと辞書Xの認識率

文字	第一候補	第二候補	第三候補	合計
あ	9/10	1/10	0/10	10/10
い	8/10	2/10	0/10	10/10
う	9/10	1/10	0/10	10/10
a	10/10	0/10	0/10	10/10
b	10/10	0/10	0/10	10/10
c	10/10	0/10	0/10	10/10
A	10/10	0/10	0/10	10/10
B	6/10	3/10	1/10	10/10
C	10/10	0/10	0/10	10/10

アルファベッ ト認識率	480/520 (92.3%)	33/520 (6.3%)	5 /520 (1.0%)	518/520 (99.6%)
ひらがなを含 む認識率	1067/1220 (88.0%)	106/1220 (8.5%)	16/1220 (1.6%)	1189/1220 (98.1%)

5. まとめ

センサやデバイスに依存せずに簡易で汎用性の高い空中手書き文字入力を実現するため、ペンのストローク毎のアップダウン情報および絶対位置情報を利用しない構成とし、相対的な移動方向情報のみを用いて1文字単位で一筆続け書き文字を認識する方式を提案している。

今回、アルファベットを追加するに当たり、覚えやすく、高い認識率が得られる方法として小文字アルファベットは書き順通りに一筆書きで、大文字アルファベットは小文字アルファベットを二度、一筆書きする方法を採用した。

被験者Aの書き方に合わせた辞書Aを作成後、被験者A, B, C, Dの4人の被験者に対して実験を行った。その結果、第三候補までの合計認識率が95%を超える高い認識率が得られた。しかし、被験者B, C, Dに共通して認識率が低い文字や、第一候補に選出されにくい文字があることから、実験の際に取った筆記データを基にしてアルファベットの辞書データを修正した辞書Xを作成し、再度、4人の被験者に対して実験を行なった。その結果、被験者B, C, Dだけでなく被験者Aの認識率が向上し、被験者4人でアルファベットの第一候補に選出される認識率は90%以上、第三候補までの合計認識率は99%が得られた。

今後、認識した文字の変換や確定、候補文字の選択等の文章を作成するためのIME処理の検討などを行って、実用化を目指したい。

参考文献

- [1] 西田好宏, 小倉一孝, 三浦浩一, 松田憲幸, 瀧寛和, 安部憲広: 移動方向情報のみを利用した空中手書き文字認識; ヒューマンインタフェース学会誌, Vol. 12, No. 3, pp. 289-296, 2010 年
- [2] 西田好宏: 移動方向情報とストローク比率を利用した空中手書き文字認識; ヒューマンインタフェースシンポジウム, HIS2010, pp. 311-314 (2010).

(平成 24 年 3 月 31 日受理)