距離画像を用いた決定木による指文字認識*

慶島 淳一[†]・小野 智司[†]・中山 茂[†]

Finger Character Recognition Based on Decision Tree Using Distance Image*

Junichi KEISHIMA[†], Satoshi ONO[†] and Shigeru NAKAYAMA[†]

1. はじめに

手指の伸縮により文字を表現する指文字体系は、手話が存在しない単語、固有名詞などを相手に伝える場合や、手話単語の説明に用いられる。手話や指文字、ジェスチャなどの認識システムは、キーボードやマイクの利用が困難な、または、はばかれる状況において、視覚障害者のみならず健常者にとっても自然で有用な入力方式である。

手指は多くの関節を含むため、ジェスチャや指文字など様々な形状を表現できる一方、形状の認識、モデル化が難しい部位である。データグローブなどの接触型センサを用いると各関節の角度を取得でき、指文字の認識を正確に行える。しかし、装置の着用により被観測者に負担をかけ、ケーブルにより動作が制限されてしまう。カラーグローブなどを装着する方式は、自由な動作が可能なのものの、装置の脱着の手間がかかる。

長嶋ら[1]と広瀬[2]はカメラから取得した画像をもとに日本語指文字を認識する方式を提案している。これらの方式は、一色の単純な背景で手形状のみを撮影する必要がある。また、長嶋らの方式は12文字、広瀬らの方式は34文字の指文字を認識できるものの、「え」と「さ」、「う」と「ら」などの類似形状を識別できない(Fig. 1)。

著者らは3次元スキャナを用いるテンプレートマッチング(3-Dimensional Template Matching: 3D-TM)による認識方式を提案した[3]。3D-TMは、複雑な背景のもとであっても、41文字の指文字を高い精度で認識できる反面、認識に時間がかかる、手形状を記録していないユーザ(未登録者)の指文字の認識が困難であるという問題点を持つ。

本速報では、従来の2次元画像を対象とした決定木による指文字認識方式を、3次元スキャナにより取得した距離画像に適用できるよう拡張した方式を提案する。本方式

Fig. 1 Examples difficult to be distinguished

は、複雑な背景下でも、3D-TMと同程度に高精度で、3D-TMよりも高速に認識を行うことができる。また、手形状から抽出した特徴情報をもとに認識を行うため、未登録者の指文字の認識を高精度で行うことができる。

なお、現状では3次元スキャナのデータの取得に0.6秒の時間が必要となるため、動作を伴う指文字と手話の認識は困難であり、手話と指文字の双方を実用的な精度、処理時間で認識するシステムを開発するためには、3次元スキャナとカメラの併用が必要である。3次元スキャナとカメラの併用方法については、別の論文で検討するものとし、本速報では、3次元スキャナを用いた指文字認識方式についてのみ述べる。

2. 提案する方式

2.1 手形状抽出

提案する指文字認識方式は、3次元スキャナで取得した手形状を距離画像に変換し、特徴抽出を行った後に決定木により指文字の識別を行う。なお、本速報における距離とは、スキャナから見た場合の奥行き方向の距離とする。

Fig. 2(a)の入力に対し、遺伝的プログラミングを用いて自動生成した肌色抽出フィルタ[4]を適用した結果をFig. 2(b)に、3次元スキャナで取得した形状データを

Fig. 2 Examples of an input and distance image

* 原稿受付 2005年10月19日

[†] 鹿児島大学 工学部 Faculty of Engineering, Kagoshima University; Korimoto 1-21-40, Kagoshima 890-0065, JAPAN

Key Words: finger character recognition, 3-dimensional scanner, decision tree, distance image.

Table 1 Features

ID	Feature	Value
f_1	Direction	Up, Side, Down
f_2	Number of SFs	0,1,2,...,5
f_3	Finger direction	Up, Side, Down
f_4	Width between SF roots	Wide, Narrow
f_5	Width of SF tip	Wide, Narrow
f_6	Angle between FF and MF	Wide, Narrow
f_7	Height	High, Mid, Low
f_8	Exist SF tip just above Wr.	Yes, No
f_9	Larger area bordered by SF.	Left, Right, Equal
f_{10}	Side	Front, Back
f_{11}	Crossing FF and MF	Yes, No
f_{12}	Area of the nearest part	Large, Small
f_{13}	Distance between nearest finger and SF root	Near, Far

SF: Stretched finger, T: Thumb, FF: Forefinger,
MF: Middlefinger, Wr: Wrist,

Table 2 Experimental results

	3D-TM	DI-DT
Recognition rate ($P_1, \dots, P_5$)	92.7%	89.2%
Recognition rate ($P_6, \dots, P_{10}$)	48.7%	71.1%
Recognition time [s]	17	0.2

2.3 指文字識別

つぎに、抽出した特徴情報をもとに、決定木に従って指文字の識別を行う。決定木は、先行研究 [1,2] および5人の被験者 $P_1, \dots, P_5$ から取得した距離画像 205 枚をもとに作成した。作成した決定木を Fig. 3 に示す。作成した決定木は、非終端ノード数が 30、終端ノード数が 41 となった。深さ 1 のノードでは手の向きによって、深さ 2 のノードでは指の数によって分類を行い、深さ 3 以降のノードで、その他の属性情報に基づいて各指文字を認定する。Fig. 3 における灰色の非終端ノードは、従来の方式 [1,2] では識別を行えなかった指文字を、距離情報を利用することで識別するノードである。

3. 評価実験

10 人の被験者 $P_1, \dots, P_{10}$ から取得した距離画像 410 枚を用いて、3D-TM [3] と提案する方式 (Distance-Image-based Decision Tree: DI-DT) の性能を比較した。手形状は、手を固定せずに、スキャナとの距離を 60~70cm 程度に大まかに調整して取得した。評価実験に用いたデータには、決定木の作成に用いたデータは含まれていない。

実験結果を Table 2 に示す。 $P_1, \dots, P_5$ の平均認識率は 3D-TM が DI-DT よりも高い。しかし、 $P_6, \dots, P_{10}$ の平均認識率は、3D-TM では 48.7% であった指文字識別率を、DI-DT は 71.1% まで大幅に改善できている。これは、3D-TM は形状に基づいて指文字認識を行うため、手のサイズや指の開き具合が学習データと異なるような、未登録者の指文字の認識は困難であるのに対し、DI-DT と決定木の組合せによる指文字認識は、抽出した特徴に基づいて指文字認識を行えるためである。しかし、 $P_6, \dots, P_{10}$ に対しては、DI-DT であっても「え」や「ろ」などの識

(a) Error examples of finger character 'え'

(b) Error examples of finger character 'ろ'

Fig. 4 Error examples

別が難しい指文字において、認識に失敗することがあった。これは、Fig. 4 に示すように、同じ指文字であっても、被験者ごとに指の開き具合、曲げ伸ばし具合、手首の角度が大きく異なっていたためと考える。今後、特徴情報の設定および決定木の改善により、未登録者の認識率の改善を図る。

また、Table 2 より、提案する方式は 3D-TM よりも高速であり、0.2 秒という実用的な時間で認識を行えることがわかる。

4. おわりに

3次元スキャナから取得した距離画像を用いて指文字を認識する方式を提案した。提案する方式は、距離画像から取得した属性情報を用いて指文字を認識するため、従来のカメラ画像を入力した方式では識別が不可能な指文字も認識することが可能である。また、肌色領域や動きを含む複雑な背景のもとでも手形状を適切に取得できる。実験により、本方式は、未登録者による指文字を 3D-TM よりも高い精度で認識できること、および、3D-TM の 85 分の 1 の時間で認識できることを確認した。

今後、より多くの被験者からのデータの取得、および決定木学習 [5] の採用により、よりロバストな指文字認識を実現する決定木を生成する。

参考文献

- [1] 長嶋, 小野寺, 長嶋, 寺内, 大和: 指文字認識に関する基礎的検討; 信学技報, Vol. SP92-64, pp. 23-30 (1992)
- [2] 広瀬: 細線化画像を用いた指文字認識; 情処研報, コンピュータビジョン, Vol. 83, No. 1 (1993)
- [3] 東山, 小野, 王, 中山: 3次元テンプレートマッチングによる指文字認識; 電学論 C, Vol. 125, No. 9, pp. 1444-1453 (2005)
- [4] 慶島, 前薗, 小野, 大橋, 中山: 遺伝的プログラミングによる画像フィルタ設計; 第12回電子情報通信学会九州支部学生会講演会, D2-18, p. 136 (2004)
- [5] J.R. Quinlan: *C4.5 Programs for Machine Learning*, Morgan Kaufmann Publishers (1993)