

ステレオカメラを用いた指文字の認識手法

田村研究室 外山 貴之 宮崎 信二

1. はじめに

本研究では、ステレオカメラを用いて指文字の認識を試みる。手話で表現できない人物名や固有名詞・外来語を伝える技能を指文字といい、手話とは独立に必要な不可欠な技能である。例えば、指文字の学習システム[1]を考えると、グローブ等の器具を必要とし、拗音など動きを必要とする指文字を対象としていない。また、手話認識のための複雑な背景下で形状推定[2]を行う手法が研究されている例もあるが、輪郭線のみを使用している為、裏表や複雑な形を認識するのは困難であり、多くの指文字を認識するには新しい認識システムが必要となる。

本研究では、指文字の認識手法を模索する上でステレオカメラを使用する。指文字には手を前後に移動をすることで拗音になるものがある。このような指文字を識別するためにステレオカメラを用いて奥行き情報を取得することで解決を図る。

2. 認識手法

2.1 対象とする指文字

指文字には、あ行～わ行の 50 音やアルファベット等があり、国によって異なる。清音には大きく分けて動きを要しない静文字(41 文字)と、動きを要する動文字 (5 文字) がある。また、元になる清音との組み合わせで表現する動文字である濁音、半濁音、拗音、促音、長音がある。本研究ではすべての静文字と、ステレオカメラを用いることにより、奥行き情報を取得できることを確認するため拗音を認識させる事を試みる。

表 1 指文字表

清音	46 文字	静文字	41 文字※
		動文字	5 文字
濁音	20 文字	半濁音	5 文字
拗音	3 文字※	促音	1 文字
長音	2 文字		

※本研究の対象

2.2 認識手順と手法

図 1 に認識手法の手順を示す。

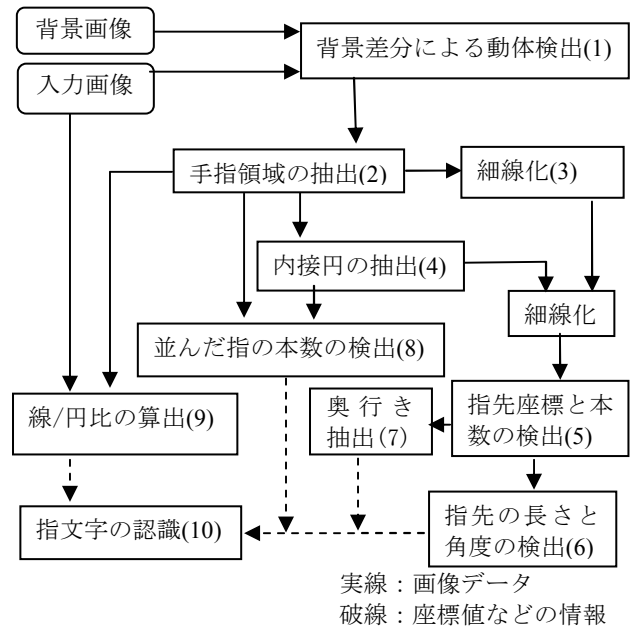


図 1 認識手法の手順

図 1 の各手順について説明する。

(1) 背景差分による動体検出

入力画像と最初に撮影した画像とで背景差分を行い、二値画像を出力する。動体があるときは動体を白画素とし、それ以外を黒とする。抽出された白画素に番号を付けてラベル画像を作成する。

(2) 手指領域の抽出

指定範囲の明度領域を抽出してノイズ除去や小領域削除・孔領域削除を行い、領域の形を整える。(1)の画像と AND を取り手指領域とする。

(3) 細線化

(2)の手指領域を、Hilditch のアルゴリズム[4]で細線化する。

(4) 内接円の抽出

手指領域のうち最大領域に内接する円を求め、手の平に近似する。その面積と中心、x 座標と y 座標の最大値と最小値を求める。

(5) 指先座標と本数の検出

(3)の細線化された二値画像の白画素から 8 近傍をみて端点を検出し、指先とする。

(6) 指の長さや角度の検出

(4)で得られた円の中心から、指先までの距離と角度を求め、指の長さや角度とする。

(7) 手の前後運動の検出

右画像と左画像から得られた、指先の座標から、y座標が±4画素の範囲内で一致する対応点のx座標の差を視差として算出する。この指先の視差により、手の奥行きを求める。拗音では、手全体を後ろへと引くため、時間的に前後の画像より視差の増減を求め手の前後運動を検出する。

(8) 並んだ指の本数の検出

(5)では離れた指の本数を検出したが、図2の様に隙間無く並んだものは分離できない。

そこで、(4)で求めた円と手指領域を用いて、上下左右4方向の円の外周より外側へ±5画素の位置にある各点を中心に指幅を算出する。

(9) 線/円比の算出

輪郭線の長さ/円の面積＝「線/円比」と表記する。円/線比は、手の表裏で異なる指文字を判別するのに使う。入力画像と(2)の画像を用いて、マスク演算し、指文字領域のみを抽出し、輪郭線抽出を行う。図2の指文字の外周に違いは無いが、『う』の様に指が前にあるものは輪郭線が長いため、線/円比に違いがある。

(10) 指文字の認識

例として図2の『う』を認識するために必要な情報は指の本数や角度、線/円比である。(5)では指の本数が出ない為、(8)より本数を求める。角度は1.58~1.80[rad]で線/円比0.6のすべての条件を満たすものを『う』とする。これ以外の指文字も同じような条件で認識が可能である。



図2 指文字の『う』(左画像)と『と』(右画像)

5. 実験

著者2人による実験を行い、各指文字に対して30秒以内に認識できた場合を成功とみなし、1つの指文字に対して10回実験を行う。なおステレオカメラの上下にライトを設置することで

影の影響を最小限に押さえ、カメラから30cm離れた場所で認識することとする。

実験の対象とする指文字は、静文字41文字と拗音「ゃ」とする。

表3 実験結果

100%	あ、す、て、ね、ふ、へ、 る、や、ゆ、る、れ
80%以上	い、き、こ、た、ち、つ、や
50%以上 80%未満	か、く、け、し、せ、と、な、 に、ひ、ほ、ま、み、む、よ、わ
50%未満	う、え、お、さ、そ、 ぬ、は、め、ら、ろ

実験結果より、著者2人の認識率の平均は表3の様な結果が得られた。認識率が高かった指文字は指の本数が多いもの、または、指の特徴が得やすいものである。認識率が低かった指文字は、類似の指文字がある指文字、すべての指を曲げ形状が円に近い指文字である。他に、認識させる際に30cmという指定範囲に入っていたとしても、指文字を行うときの光の反射具合や背景色、背後にパソコンなどの光源があるときも認識率が低下してしまうことがあった。

初心者が実験を行った場合30cmという指定範囲に入っていない、カメラの視野からずれてしまう等が考えられる。

5. まとめ

本研究では、ステレオカメラを用いる事で昨年度の研究[2]で用いた様なグローブ等の特殊な器具を使用することなく認識させることができた。

今回認識できなかった動文字を認識させる場合は、移動距離等を計算することで問題解決が図れると考えている。

参考文献

- [1] 田海慶太, “USBカメラを使用した指文字の学習用システムの試作”, 情報処理学会全国講演論文集 Vol. 69th, No. 2, PP. 281-282, 2007
- [2] 浜田康志, 島田伸敬, “手話認識のための複雑背景化で高速に運動する指文字の形状推定”, 電子情報通信学会論文誌 D Vol. J90-D No. 3 pp. 617-627, 2007
- [3] David A. Forsyth, Jean Ponce(大木剛:訳)「コンピュータビジョン」共立出版 2007
- [4] 安居院猛, 長尾知晴「画像の処理と認識」昭晃堂 1996
- [5] ちよっと便利帳 みんなの知識委員会 <http://www.benricho.org/yubimoji/50on.html>