

RGB-D カメラを用いた指文字認識

青木俊和[†] 野村昌輝[†] 高塚崇文[‡] 田村仁[†]

日本工業大学工学部創造システム工学科[†] 日本工業大学 大学院工学研究科情報工学専攻[‡]

1. はじめに

手話は聴覚障害者と健常者がコミュニケーションをとる際に重要な役割を担っている。しかし、健常者のなかで手話を使える人は限られているため、広く一般にコミュニケーションをとることは困難である。

これに対し、手話の自動認識や通訳システム開発の試みが多くなされている。

本研究では、手話の認識を行う前に指文字を認識する。そのためのインターフェースとして RGB-D カメラを用いる。今回は ToF 方式を採用しているカメラや、ステレオカメラよりも安価な kinect を使用し、そこから得られる距離画像の座標情報を用いて、従来よりも効率的な認識を行う。

2. kinect について

kinect は、複数のセンサとプロセッサ、USB2.0 インターフェースによって構成されているデバイスである。本研究で使用した kinect のセンサは 8 ビットの VGARGB 画像カメラやクラス 1 赤外線レーザによるパターン投光器と赤外線カメラで構成される深度センサである。計測可能距離は 0.8m~3.5m の間を理想とする。

2.1 kinect の精度と改善

最初に認識を行う予備実験として、0.6~3m の範囲での kinect の精度を確認した。kinect の距離画像は、通常状態では輪郭線が上手く取れず、指先の欠けやデータの孔がでてしまった。Davide らの実験[1]によると、kinect の精度は 65~75m が最も精度が高く、平均誤差は 2.8mm 以下である。このため、75cm 以上での指先の認識を行おうとすると、通常状態では困難である。そこで、より認識精度を上げるために、距離画像に輪郭線の改善を行い、認識にどのような影響を与えるのか検証した。

また、kinect の構造上の問題として、赤外線レーザと赤外線カメラの位置がずれているので、影が発生してしまう。この問題に対しては kinect に搭載されている RGB 画像カメラを使用して補正を行う。

2.2 輪郭線の改善

輪郭線の精度をより正確の確かめるために、水平方向と垂直方向 Sobel フィルターを使い、輪郭線のみを抽出した。この処理により、図 1 の状態では輪郭線が不安定なことが分かった。

この輪郭線を改善するために、10 フレームの画像を Buffer に格納し、平均化した。輪郭線、指先の欠けは図 1 に比べて改善したが、処理が重く他の処理と併用ができなかった。

別の方法として、距離画像にスムージングを施した。通常画像と比べ輪郭線は改善されたものの、データの孔と影の影響で手領域の抽出がうまくいかなかった。

また、図 1 の通常画像に写っている最も黒い部分が影である。



図 1 左:通常画像 中:Smooth 右:RingBuffer

2.3 データの孔と影の補正

データの孔と影の部分の距離情報は 0 になっているので、その箇所を補正する処理を考える。そこで、距離画像と RGB 画像の 2 つを用いて、距離情報が 0 の箇所を補正していく。その前準備として、深度センサと RGB カメラの位置がずれているため、距離画像と RGB 画像の座標を合わせる。その後、距離情報が 0 の X,Y 座標を取得し、RGB 画像上の同座標から似た色を 1 画素ずつ走査していき、同じ色であれば隣の距離情報に置換していく。色の判別には、各色のユークリッド距離を計算し、同色かどうかを判別する。

Finger language recognition using the RGB-D camera

[†]Toshikazu Aoki, Masaki Nomura, Hitoshi Tamura · Innovative Systems Engineering Nippon Institute of Technology

[‡]Takahumi Koutuka · Graduate School of Engineering Nippon Institute of Technology

今回は RGB のみではなく、HSB, HSL の計 3 つの表色系を使い、影補正を試みた。



図 2 影補正の図(左上:RGB 右上:HSB 左下:HSL)

3. 手の検出方法

本研究では、深度センサから得られる深度情報を用いて手領域の検出を行う。

最初に余分な距離情報を排除するため、指定の距離以降の深度情報を遮断してから手領域を検出する。その後、最も近い距離にある物体を手と想定し、その物体の領域のみを抽出する。

4. 円形度による手の形状判別

抽出した手領域の画像にラベリング処理を施し、最も面積の大きいラベルの輪郭抽出を行い、式(1)から円形度を求める。

$$C = \frac{4\pi S}{L^2} \quad \text{式(1)}$$



図 3 深度情報から抽出した各手の画像

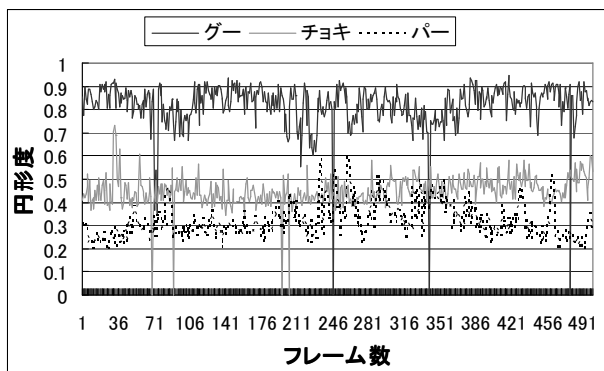


図 4 フレーム毎の各手の円形度折線図

5. 細線化による指先の特徴点検出

抽出した手領域の画像の閾値を 50 に設定し 2 値化処理を施す。その後、細線化を行い、指先の端点を特徴点として検出した数を求める。



図 5 細線化した手の特徴点検出例

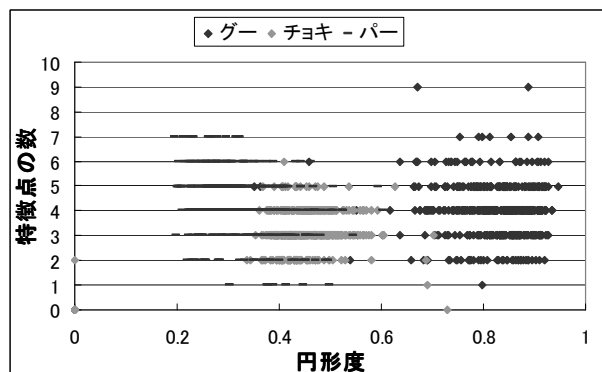


図 2 各手の円形度と特徴点の数の分布図

6. 実験結果

抽出した手領域の円形度を求め、グー・チョキ・パーの各手を判別することはできたが、パーをチョキと誤認することが多く、グー・チョキの判別と比べ精度は低かった。そこで、特徴点の数と円形度を合わせた結果、円形度のみの場合より精度は向上した。

7. まとめ

当初は影補正を行った後、円形度と細線化による指先の特徴点の計 3 つを合わせて行う予定だったが、現時点では影補正、円形度と細線化の処理が別々のままである。

今後は、3 つの処理を合わせ、再度実験を行うどのように手が検出されるかを検証する。

参考文献

[1] Davide Minnai, 伊藤 大生, 大川 涼, 高塚 崇文, 田村 仁
「RGB-D カメラによって得た頭部情報の三次元モデル化法」
『情報処理学会全国大会講演論文集』
Vol.74 No.2 pp.245-246, 2012