

温度画像処理による手話認識¹

吉富康成²

京都府立大学人間環境学部⁵

永山しづえ³

神田通信機⁶

杉山雅祥⁴

滋賀富士通ソフトウェア⁷

1 緒言

近年、福祉活動の発展に伴い、手話の出来る健聴者が幾分増えてはきたがまだ少なく、健聴者とのコミュニケーションに手話を用いる人々にとってインタラクティブな手話インタフェースの実現が望まれている。このため、これまで可視光画像を用いた手話認識システムの研究[1,2]が進められてきたが、可視光画像から色情報を用いて手領域を抽出しているため、照明条件や背景の影響を受けるという問題点がある。そこで、著者らは、この問題点を解決するために、明るさ等の影響を受けない赤外線による2次元温度分布画像(「温度画像」と表記)を手話認識に用いる研究を進めてきた[3]。

本研究では、温度画像を用いた手の追跡及び手話認識を検討した。

2 赤外線画像の特徴

本研究では、赤外線を用いて手話温度画像を作成している。温度画像は、下記の式で表される放射発散度 $W(W/cm^2)$ と温度 $T(K)$ の関係(Stefan-Boltzmannの法則)を利用して作成できる。

$$W = T^4 \quad (1)$$

ここで、Stefan-Boltzmann 定数は $5.6705 \times 10^{-12} W/cm^2K^4$ である。一方、放射率は物質とその表面状態に依存し、理論は1(完全黒体での値)である。この法則により、赤外線を用いて皮膚の温度が測定できる。人間の皮膚の放射率は $0.98 \sim 0.99$ であり、理論値に極めて近い。一方、日常生活で体の近くにある物質の放射率は 0.98 より低いので、放射率を1に設定し温度測定領域を皮膚の温度域(約 $303 \pm 7K$)とすることにより、背景から皮膚領域だけを抽出できる。皮膚の温度範囲の

1: Sign language recognition using thermal image processing

2: Yasunari Yoshitomi, 3: Shizue Nagayama, 4: Masaaki Sugiyama

5: Faculty of Human Environment, Kyoto Prefectural University

6: Kanda Tsushinki Ltd.

7: Shiga Fujitsu Software Ltd.

個人差は少ないと考えられるので、測定対象に対して同一の測定条件を用いることができる。また、赤外線は物体自体から放出されるため、照明は必要なく昼夜いつでも皮膚領域を抽出できる。

3 温度画像の作成

赤外線画像装置を用いて人の皮膚のみを抽出した手話の温度画像を作成し、8ミリビデオに録画した後、動画をMpeg2ファイルに落とし、1秒間に10フレームずつの256階調の静止温度画像(図1)を作成した。

4 温度画像からの手領域抽出と追跡

作成した静止温度画像を用いて、手がどのような動きをしたかを調べるために手の追跡を行う。この静止画を使用し、手候補領域を抽出する。まず温度画像のノイズを除去した後、メディアンフィルタを用いてスムージング処理をし、画面の平滑化を行う。続いて、線形変換を施し、固定閾値により2値化する。これにラベリングを行い、同じラベルの面積を測り、面積の大きいものから10個取り出し、これを手候補領域とする。この際、ある場面において顔に手がかかっていた場合、テンプレートマッチング法を用いて顔から手候補領域を抽出する。また、フレーム間差分による手候補領域の絞り込みを行う。そして、これらの手領域候補から前フレームで判断した手領域との形状誤差、重心間距離、外挿した重心との距離の3つの評価項目の結果による多数決によって手領域を特定し、次の画像



図1 人の皮膚のみを抽出した静止画

の処理に移る。すべての静止画像の手領域を特定し終えれば、左右の手それぞれの重心の軌跡を作成し、手話の手追跡を終了する。

5. 手話認識

手話の手追跡に成功したもののみ手の軌跡の時系列データを作成し、教師データとした。そして、未知手話の手の軌跡を作成した。同じ意味の手話でも所要時間がまったく同じではない。そこで、教師データにおける所要時間を未知データと同じにする時間の規格化を行った。この時、未知データの各フレームの時刻と対応する時刻の教師データの座標位置は内挿により決定した。このようにして求めた座標と未知手話の各フレームでの手の重心座標との距離を求め、この距離の総和が最も小さい教師手話の意味を認識結果とした。

6. 実験

手話を認識するために予め教師データを作成する。その後、未知の画像に対して認識を行う。まず、実験のために次のような条件で70種類の手話を選択した。

- ・ 軌道を描き、大きな動きのあるもの
- ・ 名詞
- ・ 頭の動きの意味のあるものは除く
- ・ 最初に型を作っているものは除く
- ・ カメラ方向の動きに意味のあるものは除く
- ・ 手の交差のあるものは除く

そして、これらの手話を2回ずつ録画した。各手話の1回目を教師データ作成用とし、動画から作成した静止画像から手の軌跡を求めた後、軌跡の時系列データを作成した。次に、1回目で追跡が成功した手話と同じ意味の2回目の手話の画像を用いて、手の軌跡を作成した。そして、1回目の手話の軌跡との各時刻での位置のずれを基に手話を認識した。70種類の手話のうち、1回目で手の追跡が成功したものは43種類で、追跡成功率は61%となった。また、この43種類の手話のうち2回目でも手の追跡が成功した手話40種類に対する手話認識成功は30種類で、認識率は75%であった。図2図3に手話追跡成功例を示す。

7. 結言

温度画像処理を用いた手話認識法とその実験結果を報告した。手の追跡成功率は61%、追跡成功の場合の手話認識率が75%だった。今後の性能改善方策として、テンプレートマッチング条件の再検討、認識時において手の形状等評価項目の追加などがあげられる。



図2 手話“人々”の追跡成功例



図3 手話“テレビ”の追跡成功例

参考文献

- [1] 渡辺賢,岩井儀雄,八木康史,谷田正彦,“カラーグロブを用いた指文字の認識”,信学論(D-II),vol.J80-D-II,No.10,pp.2713-2722,1998.
- [2] 今川和幸,呂山,猪木誠二,松尾英明,“肌色領域により隠れて見える場合を考慮した手話動画からの手の実時間追跡”,信学論(D-II),vol.J81-D-II,No.8,pp.1787-1795,1998.
- [3] 稲田美保子,吉富康成,田中幸枝,“温度画像処理による手話における指文字解析”,計測自動制御学会九州支部,第19回学術講演会予稿集,pp.129-132,2000.