温度画像処理による手話認識

（情報処理）杉山雅祥

１.序論

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。以上のことから、これまでに可視光画像を用いた手話認識システムの研究が進められているが、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、照明条件や背景を一定にしなければならないという問題点を持っている。本研究では、この問題点を解決するために光の影響を受けない赤外線による 2 次元温度分布画像を用いた手の追跡及び手話認識を行った。この方式を用いれば、例え暗闇においても手話を認識する事ができる。本研究は、OS を UNIX としたワークステーションを用いた既報の研究[1]をシステム試作に結びつけるため、パソコンで動作するプログラムの開発と、精度向上、発展を目指したものである。

本研究では、開発環境として使用 OS を Microsoft 社の Windows2000、開発言語として同社の Visual C++6.0 を使用している。

2.方法

図 1.人の皮膚のみを抽出した静止画

最初に動画を入力し、静止画像を作成する。この画像の作成は、赤外線画像装置(ニコン製サーマルビジョンシステム LARD-3ASH)を用いて、256 階調の手話の温度画像を作成し、8 ミリビデオに録画した後、動画像をファイルに落とし、1 秒間に 10 フレームずつの静止画を作成した。この温度画像(図 1)は、 Stefan-Boltzmann の法則を利用して、人の皮膚のみを抽出し作成されている。

図 2 に、本法の処理フローを示す。静止画像を使用し、手領域候補を抽出する。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、同じラベル領域の面積を図り、この面積の大きいものから 10 個取り出しそれぞれを手候補領域する。この時、同時に顔部分の処理も行う。通常、手候補領域を取り出した時点で顔は削除するだけだが、手が顔にかかっている、もしくは顔の中に入っていると判断した場合は、テンプレートマッチングにより前回特定した手領域の位置から顔にかかっている手の位置を予想し、顔領域と重なる手候補領域を抽出する作業を行っている。また、フレーム間差分により手候補領域の絞り込みを行う。

最後に、前のフレームで判断した手領域と全ての手候補領域とを形状誤差、重心間距離、外挿した重心との距離で判断し、特定される手候補領域が 2 つ以上一致していれば手領域を決定する。決定した手領域が、時刻 t と時刻 t+1 の間でどのような位置関係になったのかを判断し、手追跡をするという手法をとっている。

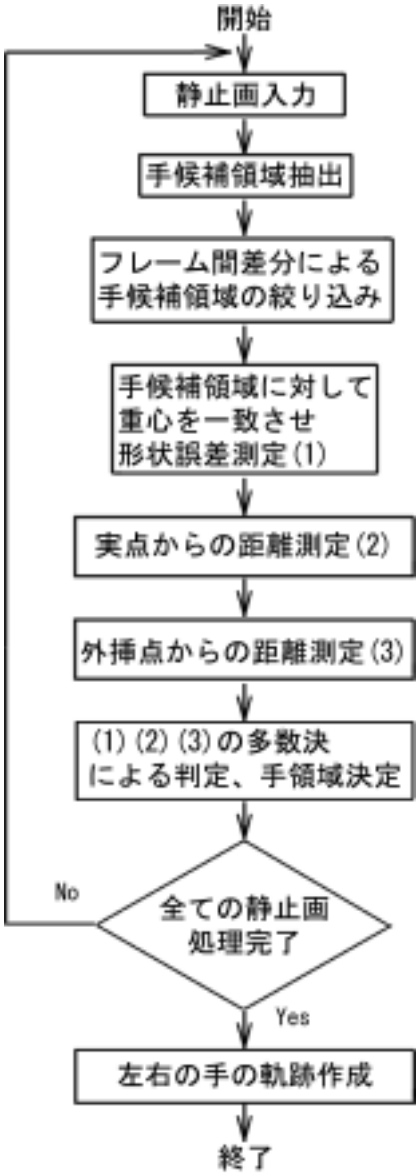

図 2.手追跡処理のフローチャート

図 3.手話"人々"の追跡成功例

3.実験

3.1 条件

予め手話を認識するために学習を行い、その後、未知の画像に対して手話認識を行う。まず、実験のために次のような条件で 70 種類の手話を選んだ。

・軌道を描き、大きな動きのあるもの
・名詞
・頭の動きに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

次に、これらの手話を 2 回ずつ録画した。各手話の 1 回目を学習用とし、動画から作成した静止画像から手の軌跡を追跡した後、軌跡の時系列データを作成した。次に、追跡が成功した手話の中から 11 個をランダムに選び、認識用とした 2 回目の手話の画像を入力し、手の軌跡を追跡した。最後に学習結果と照らし合わせ、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。

3.2 結果と考察

手の追跡を学習する部分において、70 種類の手話を入力した結果、43 パターン成功した。図 3 が成功例での手の軌跡である。失敗したものは 27 パターンあり、追跡過程で左右の手を取り違えることと、手が顔にかかった時のテンプレートマッチングでの手の予想ミスが主な原因である。続いて認識部において 11 個の手話のパターンを未知の手話として入力した結果、9 個のパターンの手話を正しく認識できた。

4.結論

今回の実験の結果、追跡率は約 61%、追跡成功の場合の認識率が約 82%という値が出た。これにより、温度画像処理による手話認識の可能性を示したと言える。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などがあげられる。

参考文献

[1]永山しづえ, " 温度画像による手話における手追跡 " ,宮崎大学工学部情報工学科,平成 12 年度卒業論文,1-27,2001