

5C-8 手話画像認識のための顔構成要素の抽出に関する検討*

長嶋 祐二 岡 靖雄 長嶋 秀世
工 学 院 大 学

1 まえがき

我々は、聴覚障害者と健常者の円滑なコミュニケーションを実現させるために、手話-日本語翻訳システムの構築を目指している。手話の調動は主に手指の型や動きにより意味情報が伝達されるが、顔表情や視線の方向によって意味の強調や限定を行ったり、アナホリック的指示などを行う。したがって、画像処理による手話認識の一部として、顔表情の認識処理が必要となり、そのためには顔構成要素(目・鼻・口など)を抽出しなければならない。顔構成要素の抽出についてはこれまでも検討がなされているが、手話動作には様々な首振り運動を伴うものがあるので、これらについて考慮した処理が必要となる。

そこで本報告では、首振り運動の中の首を横に傾ける動作に注目し、顔の傾きを検出する事により、傾いた顔画像からも顔構成要素を抽出可能な方法について検討を行った結果を述べる。

2 モーメントによる顔の傾きの補正

入力画像は人物正面上半身画像の濃淡画像(大きさ256×256画素、256階調)であり、背景は暗幕として、以下、画像の左上を原点、横軸右方向をX軸正方向、縦軸下方向をY軸正方向とする。

入力画像から体領域を抽出した後、体領域のY軸に対する周辺分布により、顔領域を抽出する(図1)。

この顔領域の中心線(眉間と鼻の頭を通る直線)の傾き角をモーメントにより検出する。顔の中心線が通り、顔の中心領域に位置する鼻の頭を以下のように求める。

①顔領域抽出画像を各濃度レベルの面積が等しくなるように8段階に量子化し5×5のメディアンフィルターをかける。

②濃度レベル8の等濃度画像において、顔領域抽出画像の重心 $G_{F_x, y}$ を中心に半径 r の円形に切り抜く。ここで、 r は重心 G_F から上下、左右、斜めの8方向の



図1: 顔領域抽出画像

背景までの距離の平均の50%とする。

③等濃度画像のラベリングを行い、各領域の重心 G_{N_x, y_i} と面積 S_i を求め以下の条件を満たさない領域を削除する。

- $\|G_F - G_{N_i}\|_2 > \epsilon_r$
- $S_i < S_{th}$

ただし、 $\|\cdot\|_2$ はユークリッド距離、 $\epsilon_r = 20, S_{th} = 3$ とした。

④残った領域で重心が背景から一番遠い領域を鼻の頭の候補とし、背景までの最短距離を h_{min} とする。 $h_{min} > r$ ならば、その領域の重心を鼻の頭とし、濃度レベル8で鼻の頭が見つからなければ、濃レベル7も同様に調べる。

慣性モーメント[1]は、濃度値がある軸からどのように分布しているかを定量的に表すことができる。図形の原点(重心 $G_{N_{g_x, g_y}}$)を通り、X軸と θ の角度をなす直線 $Q: y = x \cdot \tan \theta$ の回りの慣性モーメントは次式で表される。

$$M_\theta = M_{20} \sin^2 \theta + M_{02} \cos^2 \theta - M_{11} \sin 2\theta \quad (1)$$

ここで、 M_{pq} は重心 $G_{N_{g_x, g_y}}$ を原点とする座標形における $(p+q)$ 次の重心モーメントであり次式で表せる。

$$M_{pq} = \sum_y \sum_x (x - g_x)^p (y - g_y)^q g(x, y) \quad (2)$$

$g(x, y)$: 濃度値

ここで、重心は鼻の頭とする。

慣性モーメントは顔領域抽出画像を全部用いると、髪型などの影響を受けやすいので、鼻の頭 $G_{N_{g_x, g_y}}$ を中心

* A STUDY ON EXTRACTION OF FACIAL PARTS FOR JSL RECOGNITION
Yuji NAGASHIMA †, Yasuo OKA † and Hideyo NAGASHIMA †
† Kogakuin University

に半径 R (ここで R は鼻の頭から背景までの最短距離の 80% とした。) の円形に切り抜いた画像で慣性モーメントを求める。慣性モーメントが最大となる θ_{max} を顔の傾き角とし、鼻の頭 $G_{N_{g_x, g_y}}$ を通り傾き θ_{max} の直線を仮の顔の中心線とする。結果を図 2 に示す。顔の中心線が X 軸と垂直になるように座標系を回転させる。



図 2: 顔の傾き角検出結果

3 目・口の中心点の検出

濃度レベル 1 の画像に対してラベリングを行い、各領域の重心 $G(g_x, g_y)$ 、面積 S_i を求め以下の条件を満たさない領域を削除する。ある 2 つの領域の重心を、 $G_1(g_{x1}, g_{y1})$ 、 $G_2(g_{x2}, g_{y2})$ とし、 G_1 、 G_2 の中心線からの距離をそれぞれ r_1 、 r_2 、 G_1 、 G_2 間のユークリッド距離を d_{12} 、顔領域抽出画像の重心を $G_F(f_x, f_y)$ とする。

- $S_{min} < S_i < S_{max}$
- $|r_1 - r_2| < \varepsilon_r$
- $|g_{y1} - g_{y2}| < \varepsilon_y$
- $d_{min} < d_{12} < d_{max}$
- $g_{y_i} < f_y$

ここで、 $S_{min} = 5$ 、 $S_{max} = 200$ 、 $\varepsilon_r = 30$ 、 $\varepsilon_y = 20$ 、 $d_{min} = 30$ 、 $d_{max} = 70$ とした。

残った領域の数が 2 個の場合、その領域の重心を目の中心点とする。残った領域の数が 3 個または 4 個の場合、眉毛と目であるので下側にある 2 つ領域の重心を目の中心点とする。その他の場合は、濃度レベル 2 で同様に目の中心点を探す。

目の中心点が求められたら、左右の目の中心を通り左右の目を結ぶ直線に垂直な直線を顔の中心線とする。

次に、口の中心点を求める。顔領域抽出画像を Sobel オペレータで微分する。顔の中心線上で左右の目の中心から、下方向に、(左右の目の距離) $\times 1.2$ だけ離れた点を、仮の口の中心点 $G_{M_x, y}$ とする。 G_M を中心に横 W_{M_x} \times 縦 W_{M_y} のウインドウをつくり、ウインドウ内の微分画像を判別分析法により 2 値化する。ここで、 $W_{M_x} =$ (左右の目の距離)、 $W_{M_y} = W_{M_x} \times 0.7$ とす

る。ウインドウ内のエッジ画像の重心の Y 座標を口の中心点 G_M の Y 座標とする。

4 目・口の輪郭の抽出

微分画像において、目の中心点を中心に横 W_{E_x} \times 縦 W_{E_y} のウインドウをつくり、ウインドウ内の微分画像を判別分析法により 2 値化する。ここでは、 $W_{E_x} = 25$ 、 $W_{E_y} = 15$ とした。エッジ画像を、 N 回拡大 N 回縮退処理し、面積最大の領域を凸閉包 [2] し、これを目の輪郭とする。ここでは、 $N = 3$ とした。

口の輪郭は、微分画像において口の中心点 G_M を中心に横 W_{M_x} \times 縦 W_{M_y} のウインドウをつくり、目と同様に輪郭を求める。但し、 $N = 5$ とした。輪郭抽出結果を図 3 に示す。

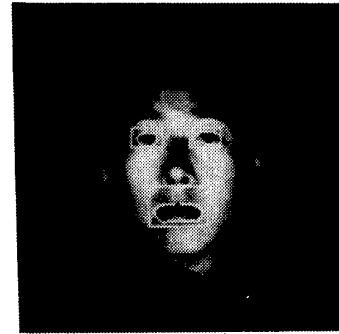


図 3: 目・口の輪郭抽出結果

5 実験結果

顔を横に最大 45 度傾けた 8 人 74 枚の正面画像に対して、目・口の輪郭抽出処理をした結果 69 枚の画像で目・口の輪郭が抽出でき抽出率は 93% であった。また、顔の振り向き角が上下左右にそれぞれ、5、10、15、20、30 度の画像で目・口の輪郭抽出処理をした結果、左右方向に 10 度、上方向に 20 度、下方向に 15 度以内で正しく目・口が抽出できた。

6 あとがき

本報告では、モーメントにより顔が横に傾いていても、顔の傾きを補正し、目・口の位置を求め目・口の輪郭を抽出することが可能となった。

今後は、顔が振り向いているときに、目・口を抽出する方法を検討し、目・口の輪郭を用いて顔表情認識を行っていく。

参考文献

- [1] 南, 中村: 画像工学, コロナ社.
- [2] 谷内田: コンピュータビジョン, 丸善株式会社.