手話認識における指型の認識に関する検討

5C−9

小野寺 卓　長嶋 祐二　長嶋 秀世
工学院大学

1. まえがき　聴覚障害者のｺﾐｭﾆｹｰｼｮﾝ手段の1つとして用いられている手話を，画像処理によって認識する研究が盛んに行われている[1]．

手話は両上肢を上半身の前で様々に変化させ，その位置や形，動き等の身振りに意味情報を結び付けた視覚言語であり，その認識には両上肢の動きを捉えることが重要となる．

本報告では，手話動作を構成する様々な要素の中の，手首から先の手掌及び五指で作られる指型に注目し，画像処理によりこれを抽出する方法について検討した結果を述べる．

2. 動き領域の抽出　運動する上肢を画像処理で抽出する手法としては，手掌や指，肘等に目立ちやすいﾏｰｸをつけてこれを追跡することが従来から行われてきた．ここではﾓﾉｸﾛで撮影を行い，衣装と肌露出部との間に生ずる輝度値の差という簡単な特徴に基づいて，特別なﾏｰｸづけ等を行わずに手の抽出を行う．対象者は手首まで袖のある長袖で，ﾓﾉｸﾛで撮影したときに肌よりも低輝度で写る模様等のない単色系の衣服を着用していると仮定する．

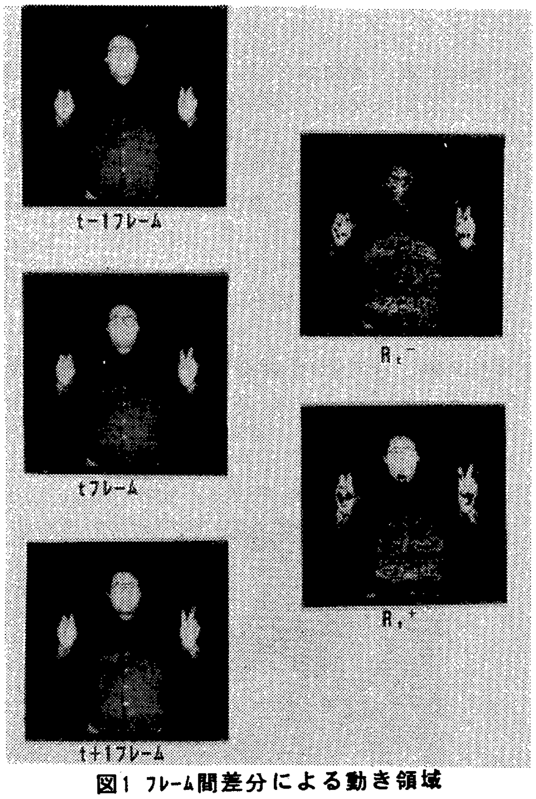

図1 ﾌﾚｰﾑ間差分による動き領域

連続撮影された上半身動画像の第tﾌﾚｰﾑを$d_t[x,y]$と表す．一般に動画像の動き情報はﾌﾚｰﾑ間差分によって示される．ここでは，次のような方法で動き領域を抽出する．

①ﾌﾚｰﾑ間差分処理　各ﾌﾚｰﾑを，顔や手等の肌露出部からなる高輝度領域と，衣服や頭髪，背景領域からなる低輝度領域とに分けて考え，低輝度領域に対する高輝度領域の動きを捉えるために，次式のような条件を満たす領域をを抽出する．

$$R_t^+ = \{d_{t+1}[x,y] \mid d_{t+1}[x,y]-d_t[x,y] > th\} \quad (1)$$
$$R_t^- = \{d_{t-1}[x,y] \mid d_{t-1}[x,y]-d_t[x,y] > th\} \quad (2)$$

ただし $th > 0$

これにより，第tﾌﾚｰﾑを基準として，第t+1, t-1ﾌﾚｰﾑで動きのあった高輝度領域R_t^+, R_t^-が得られる．R_t^-はt−1→tﾌﾚｰﾑの間に消滅した高輝度領域を，R_t^+はt→t+1ﾌﾚｰﾑの間に現れた高輝度領域を表している．また，thは動きとは見なされない小さな輝度値の変化や，カメラ系の雑音を抑制するための閾値である．図1に抽出されたR_t^+とR_t^-の例を示す．

②判別分析法による閾値の決定　R_t^+とR_t^-より求めた適当な閾値によって原画像を二値化して，高輝度領域に含まれる肌露出部を抽出する．ここではR_t^+とR_t^-に含まれる輝度値が，肌露出領域とその他の領域との二つのﾓｰﾄﾞを形成していると仮定して，各ﾌﾚｰﾑ毎に判別分析法を用い，閾値th_tを得る．この値は各ﾌﾚｰﾑに固有のものだが，これを用いて各ﾌﾚｰﾑ毎に二値化した場合，th_tの値に多くの不連続点が存在するために正しい結果が得られない．これらの不連続点は，動作開始時やある動作から別の動作に移行する時に，動き領域が消滅するために生じたものである．

そこで，適当なﾌﾚｰﾑ長を設定して，この範囲内でのth_tの平均値を全ﾌﾚｰﾑ共通の閾値として用いる．これにより，動きのないﾌﾚｰﾑに対しても，近似的に動きのあるﾌﾚｰﾑの閾値を適用することができる．

3. 手の抽出　この結果には手や顔等の肌露出領域とその他の雑音が含まれている．ここから，ﾌﾚｰﾑ間の連続性に注目して移動する手領域を抽出する．

初めに，各ﾌﾚｰﾑに対して八近傍連結に注目したﾗﾍﾞﾙづけを行う．ﾗﾍﾞﾙづけされた各孤立領域は，それぞれ手や顔領域等を表しており，これらがﾌﾚｰﾑ間で同一のﾗﾍﾞﾙを持つようにﾏｯﾁﾝｸﾞを行う必要がある．

ここでは，一つの領域が隣接ﾌﾚｰﾑ間で大きく移動することはないという仮定に基づいて，ﾌﾚｰﾑ間のﾗﾍﾞﾙのﾏｯﾁﾝｸﾞを行う．具体的には，第tﾌﾚｰﾑにおけるﾗﾍﾞ

A Study of Recognition of Finger Sign on Recognition of JSL
Taku ONODERA, Yuji NAGASHIMA and Hideyo NAGASHIMA
Kogakuin Univ.

ﾙiがつけられた孤立領域を構成する画素集合を R_t^i と表したとき，第t+1ﾌﾚｰﾑにおいて

$$R_t^i \cap R_{t+1}^j \neq \phi \qquad (3)$$

なる条件を満たす R_{t+1}^j のﾗﾍﾞﾙをjからiに更新するという操作を基本とする．この時，特殊な状態として，二つ以上の領域が一つに融合する場合と，一つの領域が二つ以上に分離する場合とがある．これらは，手と手，手と顔領域が重なる動作の時に生ずるが，それぞれの状態を同時に記録し，融合，分離の後でも整合性を保つようにする．

次に，各領域の重心の速度ﾍﾞｸﾄﾙの大きさを基に，動きの少ない領域を求め，これらより領域が現れてから消滅するまでの寿命が短いものを除去する．以上で移動する手領域と静止した顔領域が抽出される．

4.指型の認識 抽出された手領域に対して，伸展させた指領域の抽出処理[2)3)]を用いて指型を認識する．

この処理は，**①縮退/拡大による単独伸展させた指領域の抽出 ②微分勾配の方向を用いた密着伸展させた指同士の境界線の抽出**の二つの処理からなる．

①の処理では，二値化された手領域から，他の指とつけずに伸した指（単独伸展させた指）を縮退/拡大により抽出する．

縮退/拡大処理は，領域周辺部の微細な突起構造や孤立雑音を除去し，領域の形状を滑らかにする前処理として一般的に用いられているが，ここでは，単独伸展させた各指領域を手掌領域からの突起構造と見なして，縮退/拡大によりこれを消滅させることにより抽出を行う．この処理では，任意方向に伸びる突起構造を除去可能な無方向性の縮退/拡大を用いることにより，指の方向や画像平面内での手の回転の影響を受けずに抽出が可能である．

抽出された各指領域は指ﾍﾞｸﾄﾙとして表し，その指す方向や手掌との位置関係，各々の挟角等を用いて，指型の構造を記述する．この処理結果を図2に示す．

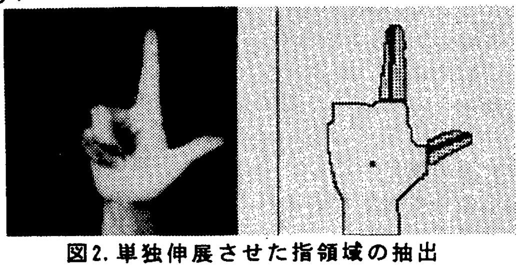

図2.単独伸展させた指領域の抽出

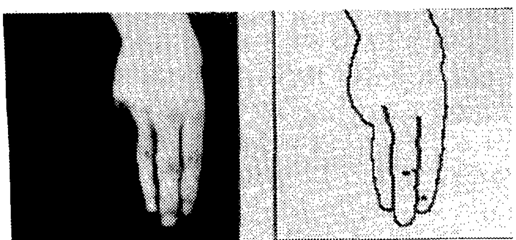

図3.密着伸展させた指同士の境界線の抽出

②の処理は，複数の指を揃えて伸した状態（密着伸展させた指型）を，各指の間に生ずる暗い筋（指同士の境界線）を抽出することにより認識する．

この処理では，指の境界に生ずるｴｯｼﾞを，指領域の内側から外側に向かって輝度値が落ちる斜面と見なし，その向きを微分ｵﾍﾟﾚｰﾀにより得られる微分勾配の方向としてとらえ，この値が互いに向き合った状態で隣接する二領域の境界を追跡することにより，密着伸展させた指同士の境界線として抽出する．この処理では，照明強度や個人の肌色の違い等の影響を受けやすい，微分強度による通常のｴｯｼﾞ抽出処理と比較して，より安定な抽出が実現できる．この処理結果を図3に示す．

5.実験結果 簡単な動作を撮影し，本手法により手領域の抽出と指型の認識実験を行った結果を示す．

ﾃﾞｰﾀはSONY製XC-57で撮影し，追記型光磁気ﾋﾞﾃﾞｵﾃﾞｨｽｸに蓄積した後，512x512x8ﾋﾞｯﾄのﾌﾚｰﾑﾒﾓﾘを介してﾜｰｸｽﾃｰｼｮﾝに転送し処理を行う．動作は，両手を上げ，胸の前で交差させ，再び同じ経路で戻すというもので，各動作の前に一瞬の停止状態を含んでいる．以上の動作を約3秒で行い，これを30ﾌﾚｰﾑ/秒の速度で撮影した．

実験の結果，図4に示すように両手の融合状態を含め，その位置と速度が追跡ができた．また速度の変化より，手が静止していると思われる1ﾌﾚｰﾑをとらえ，指型抽出処理を行ったところ，正しく認識できた．

6.あとがき 本報告では，手話動画像認識の基礎となる，動画像からの手領域の抽出，及び指型の認識処理について述べた．実験の結果，衣服と肌露出部の輝度値の差という簡単な特徴を用い手の運動が追跡でき，動作ﾌﾚｰﾑ中の指型の認識処理が可能となった．

今後は，ｶﾗｰ画像を用いてより精密に肌露出部を抽出することや，実際の手話単語認識へ適用することを検討する．

参考文献 1)安達:手話通訳システムの研究動向，信学研資，NLC92-5，pp.33-40(1992) 2)小野寺，長嶋，長嶋:縮退/拡大処理による指文字認識，信学秋期大会，D-288(1992) 3)長嶋，小野寺他:指文字認識に関する基礎的検討，信学技報SP92-64，HC92-41，pp.23-29(1992)

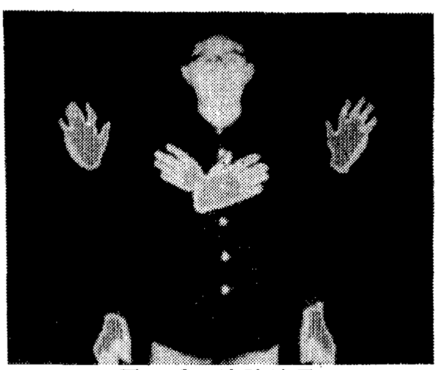

図4.手の追跡結果