

動的輪郭モデルによる対話型コミュニケーションシステム自動化の研究

知能情報工学講座 梅原 末世 (指導教員 渡邊 睦教授)

1. 研究の目的

今日、障害者・高齢者の社会参加の機会増大が強く求められている。渡邊研究室では『口周辺領域の明度分布の解析を用いた顔画像処理による対話型コミュニケーションの研究』という新しい視点からのアプローチによるコミュニケーションシステム開発を行ってきた[1]が、予め設定した領域内にユーザがとどまっていないと認識が不可能、システム起動を第3者が行わなければならないという課題があった。そこで本研究では体動があり従来のカメラ取り込み画面から外れてしまうユーザへの対応および、システムの全自動化を行うことを目的とした。

2. 処理の流れ

本システムの処理の流れを図1に示す。フレーム間の明度ヒストグラム差分変化により対象の出現・着席を検出、システムを起動する。起動後は合成音声によりユーザに指示し、各口部パターンの処理領域内の正規化した明度ヒストグラムを標準データとして作成後、会話選択システムを開始する。標準データ作成や認識中の処理領域設定には動的輪郭モデルを用いて口唇輪郭を追跡し、毎フレームごとに自動設定する。ユーザの口部パターンを相互部分空間法により認識し、ユーザに認識結果の確認を行う。ユーザは口の形状を保つことで確認を行い、キャンセルする場合は口を開閉する。確認の結果、最終的に選択された結果を合成音声で出力する。

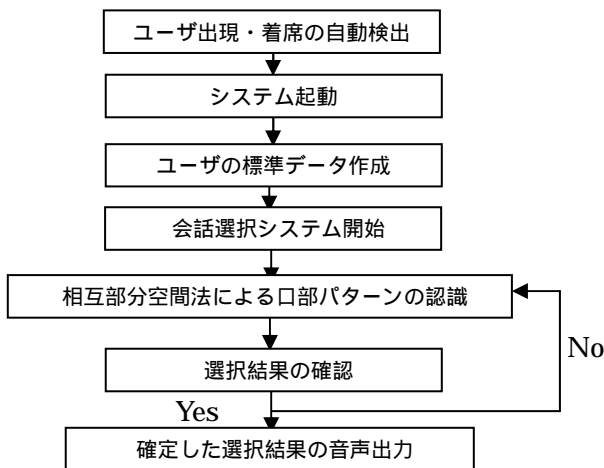


図1: 処理の流れ

2.1 動的輪郭モデルを用いた口唇領域の自動切り出し

昨年度までは入力画像のエッジ濃度より唇の両端を検出し、この2点間の距離に実験的に求めた定数をかけることで口部の設定を行っていたが、ユーザの個人性や体動に充分追従できていないという問題点があった。そこで、対象の輪郭形状や位置が変化しても追跡が可能である動的輪郭モデルを導入した。動的輪郭モデルとは、外部からの強制力 E_{con} 、線やエッジなどの画像特徴に引き寄せられる力 E_{image} 、滑らかさを保つ曲線内部の力 E_{int} のエネルギー和を最小化する問題を解くことで対象物の輪郭の位置を決定する手法である。[2]

$$E_{snake} \left(v(\vec{s}) \right) = \int_0^1 \left\{ E_{int} \left(v(\vec{s}) \right) + E_{image} \left(v(\vec{s}) \right) + E_{con} \left(v(\vec{s}) \right) \right\} ds$$

動的輪郭モデルで得られた輪郭線の外接長方形を処理領域とした。動的輪郭モデルを用いて追跡した輪郭線と設定された処理領域の例を図3に示す。口周辺に適切に設定できていることを確認できる。

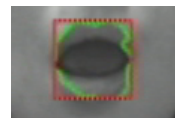


図2: 動的輪郭モデルによる口唇領域の切り出し

3. 実験結果

ユーザ検出中の明度ヒストグラムのフレーム間差分変化を図3に示す。ユーザが出現すると大きな変化が生じ、また着席を完了すると変化が減少することで着席を検出する。被験者15名(男性11名、女性4名)を対象にユーザ検出の実験を行ったところ、ユーザの出現・着席ともに検出100%であった。



図3: ユーザ検出

動的輪郭モデルの有効性を確認するため、脳性小児マヒによる体動の状態を模擬的に再現し口唇輪郭の追跡実験を行ったところ、適切に口唇を追跡し処理領域の設定が行われていることを確認した。

同被験者らを対象に動的輪郭モデルを用いて設定した処理領域での認識実験を行った認識結果を表1に示す。この結果より平均83.3%と実用上十分な認識率が得られたことを確認した。しかしながら、歯を見せるパターンについては動的輪郭モデルの追跡が不十分な点があり、改良の余地があると考えられる。

表1: 動的輪郭モデルによる処理領域の認識率

口をあける	歯を見せる	口をつぼめる	舌を出す	平均
90.0%	76.7%	88.7%	78.0%	83.3%

4. まとめ

ユーザ出現・着席の自動検出を行いシステムを起動するよう改良したことと、合成音声を用いてユーザへの指示や確認を行ったことで完全に自動で動作するシステムが実現できた。また、動的輪郭モデルを用いて毎フレームごとに処理領域を設定したことで、体動のあるユーザも認識が可能であり、平均83.3%と実用上十分な認識率が得られたことでシステムの適用範囲が大きく広がったと言える。

参考文献

- [1] 渡邊睦, 西奈津子, “顔画像処理による対話型コミュニケーションの研究”, 電気学会論文誌(部門誌 C)vol.124, No.3, pp.680-688 (2004.3)
- [2] M.Kass, A.Witkin, D.Terzopoulos, “Snakes: ActiveContour models,” Int. J. Computer Vision, Vol. 1, pp. 321-331, 1988