

Web カメラを用いた唇検出法の検討

二階 雅弘[†] 西村 広光[†]

[†] 神奈川工科大学大学院工学研究科情報工学専攻

1. はじめに

手話による会話は、手話内容の程度や強弱を表情で表現し、互いに表情を読み解くことによって成り立っている。我々がこれまでに提案してきた、カメラで取得した映像から手話を認識する手法^[1]に対し、表情認識で得られた情報を組み合わせる事によって、より高精度な手話認識を実現することが可能である。本研究では、表情認識の初期段階として、表情の中でも唇の形状に着目し、唇検出と手話認識への応用を検討した。

2. 撮影環境

本研究の最終目標は HMD などのウェアラブルデバイスでの手話認識であるため、撮影機材には Logicool社の C930e の Web カメラを用いた。Web カメラを唇の正面に設置するため、Web カメラを床から 140cm の高さ、床と水平になるように設置し、Web カメラから被写体までの距離は、会話時の距離と同じ 100cm とした。

3. 唇検出アルゴリズム

唇検出アルゴリズムを以下の(1)から(5)に示す。唇検出には HSV 色空間の Hue(以下、色相とする)を用いる。(1)から(3)のループ処理において、色相検出の閾値を徐々に変化させることで、唇と近似する形状になるまで動的に唇候補領域の探索を行う。

- (1) 抽出 1 回目は、10 人を対象に撮影した 200 枚の色相調査用データ群 A から得られた閾値をもとに色相より唇候補領域を抽出する。抽出 2 回目以降は、(3)の結果をもとに唇の候補領域ごとに閾値を変更し、新たに唇候補領域を色相より抽出する。
- (2) (1)で抽出した唇候補領域を唇候補領域の外接矩形内占有率(以下、占有率とする)、上下対称率、左右対称率より、唇領域と非唇領域に分類する。
- (3) (2)で非唇領域に分類された全ての領域に対し、面積が指定範囲外の領域は面積過度による終了とし、面積が指定範囲内の領域は、面積に応じて(1)で用いる色相検出の閾値の増減を決定する。閾値の増減を決定した非唇領域が 1 つ以上存在する場合には(1)の処理に戻る。
- (4) (1)から(3)までの処理によって動的に探索して得られた唇候補領域に対し、唇の形状特徴からスコアを算出する。
- (5) 算出したスコア上位 3 つの唇候補領域を唇領域として出力する。

4. 唇領域の形状特徴によるスコア化

色相のみを用いて唇候補領域を抽出すると、耳や首元が唇領域として検出される場合があった。そこで、抽出した唇候補領域の唇らしさをスコアとして算出することで、非唇領域が唇領域として検出結果として出力されることを回避した。唇領域の形状特徴の調査には、10 人を対象に撮影した 200 枚の調査用データ群 B から抽出した唇領域と非唇領域の占有率、縦横比、上下対称率、左右対称率を図 1, 2 のように確認した。この結果から唇領域の重心からの距離に応じてスコアを算出した。

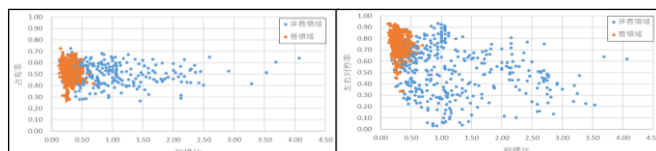


図1. 縦横比と占有率

図2. 縦横比と左右対称率

5. 唇検出アルゴリズムの評価実験

本研究で開発した唇検出アルゴリズムの唇検出精度の調査を目的として評価実験を行った。被験者 10 人を対象に撮影した 200 枚の評価実験用データ群 C から唇検出を行った。唇検出結果は主観評価によって正検出、過剰検出、誤検出、無検出に分類した。正検出、過剰検出、誤検出の例を図 3 から図 5 に示す。

唇検出の結果、正検出と過剰検出を含め、約 6 割の精度で唇検出が成功していたが、被験者によって誤検出や無検出が多い結果が見受けられた。



図3. 正検出

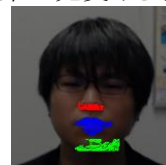


図4. 過剰検出



図5. 過剰検出

6. まとめ

本研究では、唇検出と手話認識への応用を目的とし、色相から動的に探索して抽出した唇候補領域と唇の形状特徴をもとに算出したスコアから、唇検出を行った。その結果、正検出と過剰検出を含め、約 6 割の精度で唇を検出していることが確認できた。

今後の課題として、検出結果に含まれる唇領域の絞り込み手法を模索していく。

参考文献

- [1] 小澤 辰典 ほか, "複数の弱式別統合による手話認識性能の向上", 信学ソ大, B-20-3, 2017