

口唇画像の特徴を用いた読唇処理による日本語母音の認識

情報科学科 吉永 壮汰

指導教員：何 立風

1 はじめに

近年では音声認識技術の実用化が進んでいる。しかし、音声認識技術は、喉の不調や聴覚に障害を抱える人が使用することは困難であり、公共の場など発声が躊躇われる場所での使用は好ましくない。そこで、人間が口唇の形状から発話内容を推定する”読唇”に注目し、コンピュータ上での実現を考えた。

本研究では、口唇画像から日本語母音および閉唇口形の合計 6 状態に関する認識認識を提案する。

2 コンピュータにおける読唇

”コンピュータにおける読唇”は口唇画像の特徴を基に、尤もらしい口形に分類することである。これは、パターン認識の一種であり、学習データ群から”クラス辞書”を作成することで未知の入力データについて認識を行う。”読唇”について、その認識結果を基に、処理精度を向上させるための研究がなされている。

本研究では、母音認識に関する先行研究 [1] との認識率の比較を行う。同様の認識手法を用いるが、異なる特徴量を用いることで、その有用性を検討する。

3 提案手法

1. 発話画像の収集

顔領域を含む母音発話映像を被験者から撮影し、任意のフレームを切り出すことで入力画像を収集する。

2. 画像の位置合わせ

取得した顔画像は被験者の癖によって傾きや位置が僅かに異なる。各画像に特徴点を 3 点設定し、アフィン変換を行うことで口唇位置の正規化を行う。

3. 口唇領域の抽出

Haar-like 特徴 [2] を用いた AdaBoost 法 [3] によって口唇領域を抽出する。口唇のみの抽出では図 1 に示す誤検出が発生する。そこで、検出精度の高い顔、鼻の順に検出し、口唇領域の探索範囲を限定することで、図 2 に示すように正確な抽出に成功した。

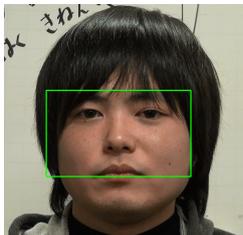


図 1 口唇領域の誤検出

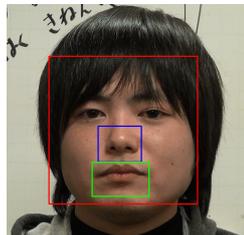


図 2 口唇領域の正しい検出

4. 特徴選択

顔画像認識に用いられる”固有顔”[4]と同様に、口唇画像を用いて”固有唇”を作成した。固有唇から元の画像を近似する際に得られる係数ベクトルを特徴量として使用する。

5. 認識手法

ニューラルネットワークを用いて認識を行う。入力は係数ベクトル、出力は母音と閉唇口形の 6 ユニットを設定。中間層はヒューリスティックに定め、出力関数としてシグモイド関数を用いる。

4 実験・考察

本研究は 4 人の被験者から画像の収集を行った。3 人の被験者から 270 枚を学習用、1 人の被験者から 270 枚を認識用として取得した。計 540 枚の画像を用いて固有唇を作成し、係数ベクトルをニューラルネットの入力とした。以下の表 1 に先行研究との比較を合わせた結果を示す。

表 1 母音認識結果

	先行研究	提案手法
A	80.0%	48.9%
I	50.0%	100.0%
U	88.0%	100.0%
E	72.0%	42.2%
O	68.0%	80.0%
X	76.0%	100.0%
平均	72.3%	78.5%

表中の (A,I,U,E,O,X) はそれぞれ (あ, い, う, え, お, 未発話) と対応している。

結果から、”I”,”U”,”X”において認識率が 100% となった。”O”では最低限の認識結果が得られていることがわかる。一方で、”A”と”E”は認識率が 50% 以下となった。これらは口形が似ているため、特徴量の類似が発生したと考える。学習における被験者数を増やし、様々な人物の形状に対応することが可能となれば、認識率の向上が期待される。

5 まとめ

本研究では、固有唇を元にした特徴量から日本語母音および閉唇口形の認識を試みた。ニューラルネットワークによる認識の結果、一部の認識率の改善が可能となれば非常に有用なシステムとして期待されることを示した。今後の課題として、学習データ数を増やし、幅広い入力に対応することが挙げられる。

参考文献

- [1] 中邨覚, 川村尚生, 菅原一孔:”動的輪郭モデルを利用した読唇母音認識システムの開発について”, 情報科学技術フォーラム一般講演論文集 5(3), pp.353-356, 2006-08-21.
- [2] Paul Viola, Michael Jones:”Rapid Object Detection using a Boosted Cascade of Simple Features”, Proc. of CVPR, pp.511-518, 2001.
- [3] 三田雄志:”AdaBoost の基本原理と顔検出への応用:CVIM 研究会チュートリアルシリーズ”, 情報処理学会研究報告.CVIM, [コンピュータビジョンとイメージメディア] 2007(42), pp.265-272, 2007-05-14.
- [4] Turk, M.A., Pentland, A.P.:”Face recognition using eigenfaces”, Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.586-591, 1991.