

畳み込みニューラルネットワークによる 視線方向の推定

張超凡[†] 阿部清彦[†]

[†] 東京電機大学大学院システムデザイン工学研究科

1.はじめに

視線入力装置は、筋ジストロフィーや筋萎縮性側索硬化症(ALS)、脊髄性筋萎縮症(SMA)等、知的な遅れはないが身体を思うように動かすことが難しい肢体不自由者にとって非常に有用なツールの一つである^[1]。視線入力システムは、肢体不自由者のコミュニケーション支援にとどまらず、ゲームプレイ時の視線方向測定など、様々な目的のために活用されている。

2.目的

本研究では、ニューラルネットワークを用い、1台の一般的なカメラから得た画像のみで視線方向を推定する新しい手法について提案する。本研究を始めるにあたって、先行研究^[2]で用いた畳み込みニューラルネットワークを利用した画像分類学習モデルの利用し、視線識別実験を行なった。

3.視線方向識別システム

3.1 視線方向識別のための CNN の構造

今回使用した CNN(Convolutional Neural Network)は、中間層が7層で構成されたものとなっており、入力した32x128ピクセルのRGB画像を、フィルタ数が32でフィルタサイズが3x3の畳み込み層に通した後、フィルタ内の最大値をそのフィルタの値として圧縮するMax-poolingを行う。Max-pooling層1層に通すことを2度行い、再度畳み込み層に通してMax-pooling層に通した後で、全結合層に通し出力するという流れである。図1に今回使用したCNNの模式図を示す。また、畳み込み層で使用した活性化関数はReLU、全結合層で使用した活性化関数はsoftmaxであり、畳み込み層3層ごとにdropoutをそれぞれ実施している。

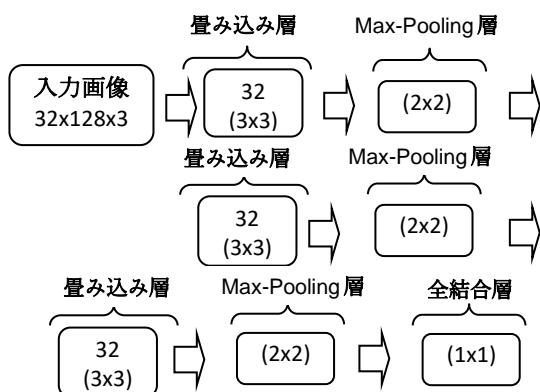


図1. 提案する CNN の構造

3.2 学習用画像の収集

学習モデル作成の流れとして、まず、被験者の顔画像をノートパソコンのカメラで撮影する。被験者がノートパソコンの前に座った状態で、webカメラを用いて上下左右正面左上右上の7方向をそれぞれ注視し、720p、30fpsの動画を600フレーム撮影した。撮影環境はノートパソコンの後ろに7方向に指標を配置する。この時、被験者の目から指示版の左右上下それぞれまでの視角は45°になるよう設定した。作成した画像からHaar-like特徴量を用いたカスケード分類器を用いて顔部分をトリミングし、その後128x128(画素)に縮小した。これを学習時には、あらかじめ指定した縦座標に基づき、目の周辺を32x128ピクセルに切り抜いて使用している。

4.評価実験

Webカメラで動画を撮影しながらフレームごとに視線方向を推定するプログラムを使い、評価実験を行なった。実験では被験者(1名)が視線を上下左右正面左上右上の7方向を注視し、各々の方向20回でテストを行う。その際の識別結果を求め、実験結果を表1に示す。各々の方向いずれも90%以上の高い識別率を示した。

表1. 識別結果

方向	上	下	左	右	正面	左上	右上
識別率 (%)	90	95	100	95	100	90	90

5.まとめ

本研究では、特殊なセンサーやカメラを使用しない、一般的なカメラのみで動作する安価で使用しやすい視線入力システムを提案することを目的として、CNNを用いて、被験者の顔画像を上下左右正面の7種類に識別することを試みた。実験の時、目とカメラの位置など実験環境に対して精度が変わるので、今後この問題についての改善を検討する。

参考文献

- [1] 塩塚敬介, 本吉大介, “重度肢体不自由教育における視線入力装置活用の現状と課題”, 教育情報研究, 2020年35巻2号 p.3-14
- [2] 菅谷裕一, 阿部清彦: “畳み込みニューラルネットワークによる視線方向推定”, 2019年電気学会電子・情報・システム部門大会講演論文集, pp.699-700