

浸透学習法を用いたマルチモーダル感情推定

佐保田 和宏[†] 柳元 美玖^{††} 長尾 智晴^{†††}
[†] 横浜国立大学 理工学部 ^{††} 横浜国立大学 大学院環境情報学府
^{†††} 横浜国立大学 大学院環境情報研究院

1. はじめに

近年、機械による人の感情推定の研究が盛んに行われている。このときに使用される情報は、取得や解析が容易な点から顔画像が主となっている。また技術の発展により、脳波や心拍といった生体信号を用いた推定の研究も行われている。生体信号は基本的に自分でコントロールすることが困難であるため、感情推定に有効な特徴量を持つことが考えられる。しかし、生体信号は取得できる場面が限られているため、実用面で問題がある。そこで本稿では浸透学習[1][2]と呼ばれる Neural Network(以下, NN)ベースの学習方法によって、生体信号から得られる認識に有効な情報を顔画像から得られる情報で補完することで、試験時に情報が欠落していても適用できるモデルを構築し、マルチモーダル感情推定を行う手法を提案する。

2. 提案手法

浸透学習では学習時、試験時ともに使える Main データと学習時のみに使える Aux データを入力として扱う。はじめに、事前学習としてこの二つのデータを入力した確率的勾配降下に基づく学習を行い、認識に有効な特徴量(以下, 浸透特徴)を中間層に形成する。次に、浸透学習として Aux データの入力を徐々に減らし、浸透特徴を Main データのみで再現するように学習する。このようにすることで試験時でも、Main データから Aux データの情報を再現し利用することができる。

本稿では Main データとして顔画像、Aux データとして脳波データを用いるモデルを提案する。入力データの形式が異なるため、それぞれの特徴量を抽出する。抽出方法として、脳波データに Convolutional Neural Network(以下, CNN)を用い、顔画像に Residual Network(以下, ResNet)を用いて抽出を行う。脳波データは、電極間の波形も考慮に入れるため、二次元の CNN を用いる。図 1 にマルチモーダル感情推定モデルを示す。

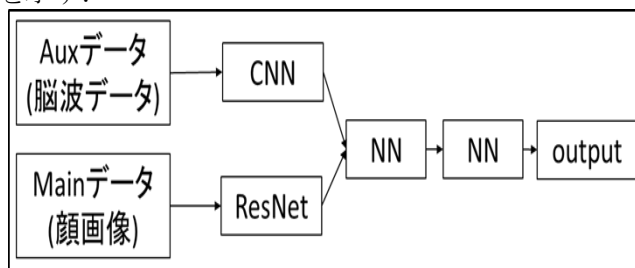


図1 マルチモーダル感情推定モデル

3. 実験方法

今回の実験では、DEAP データセット[3]という感情状態の分析のためのマルチモーダルデータセットを用いて、人の快・不快の感情を分類する実験を行った。このデータセットは音楽ビデオを視聴している被験者の正面顔映像や脳波データが記録されている。Main データとして被験者の正面顔映像から抽出した顔画像、Aux データとして脳波(EEG)データを使用する。今回の実験では快度を表す valence の値に注目した。valence は 1 から 9 までの値が付いており、今回はこの値が 4 以下でラベル付けされているデータを不快、6 以上でラベル付けされているデータを快と定義し、分類を行った。

4. 実験結果

顔画像のみによる分類結果と提案手法による分類結果については当日発表する。提案手法により、脳波による特徴を補完することができ、推定精度が向上するのではと考えられる。これによりマルチモーダル感情推定において浸透学習法による有用性が示されたと言える。

5. まとめ

本稿では浸透学習を用いたマルチモーダル感情推定の手法を提案した。提案手法により試験時に使用できない情報を補完するネットワークを構築することができた。今後は心拍や視線などといった他の生体信号を Aux データの入力として検討する予定である。

参考文献

- [1] 柳元美玖, 長尾智晴. (2017). 学習時のみに使用可能な情報を浸透させるニューラルネット. 情報処理学会研究 MPS 研究会 2017(6), 1-6.
- [2] 特許出願済:特願 2017-153613
- [3] Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... & Patras, I. (2012). Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, 3(1), 18-31.