

MediaPipe Face Mesh と SVM を用いた表情認識

1. はじめに

近年、画像処理の技術の向上によって画像や動画から顔の表情などの情報を抽出することが可能になっている。顔画像を用いた研究として、顔筋肉の動作から表情を認識する手法[1]や、表情から感情を推察する手法[2]などが開発されている。しかし、日本人の表情は Ekman[1]の提唱した基本6感情の表情と対応しない可能性が示唆されている[3]。そこで、本研究では日本人の表情に対応するため、MediaPipe Face Mesh で取得した顔の特徴点を取得し、SVM を用いて表情を分類する手法を提案する。

2. 提案手法

2.1 提案システムの概要

本研究では顔画像から感情を推察することを目的とする。Ekman[1]の基本6感情（笑顔、悲しみ、怒り、嫌悪、恐怖、驚き）に真顔を追加した7つの表情を分類する。表情データを撮影する際には、関連研究[3]と同様に6感情の画像条件の顔をまねするように指示している。MediaPipe Face Mesh によって顔の特徴量を取得し、SVM (Support Vector Machine) によって分類を行う。

2.2 MediaPipe Face Mesh を用いた特徴点抽出

顔の特徴点抽出にはMediaPipe Face Meshを用いる。Google が提供するオープンソースのフェイシャルランドマーク検出ライブラリであり、画像から468点の顔の3次元座標を取得することができる。図1はその例であり、非常に詳細な顔特徴点の座標を取得できていることが確認できる。

2.3 SVM を用いた表情分類

表情の判別にはSVMを用いる。SVMは教師あり学習を用いるパターン認識モデルの1つであり、非線形なRBFカーネル、ガンマ値を自動で設定するスケールモードを使用した。まず、MediaPipe Face Mesh で取得した特徴点に対し、 x, y, z 軸それぞれにおいて0から1の範囲に正規化を行い、468点の3次元画像座標を並べた1404次元の特徴ベクトルを作成する。それを用いてSVMにより7つの表情の分類を行う。

3. 評価実験

表情分類の評価実験を行った。分類する表情は真顔、笑顔、悲しみ、怒り、嫌悪、恐怖、驚きの7クラスである。被験者は3名とし、それぞれ各クラス100枚ずつの学習データと、同じく各クラス100枚ずつのテストデータを用意した。実験はそれぞれ本人のみの学習データ700枚を用いて学習を行

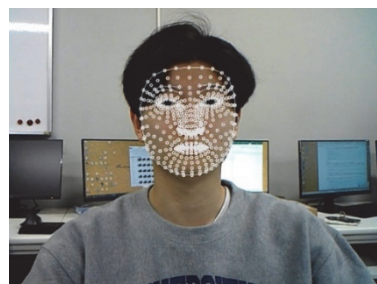


図1 MediaPipe Face Mesh による特徴点抽出

表1 個人と複数人での正答率

	被験者 A	被験者 B	被験者 C
個人モデル	96.0 %	87.9 %	89.7 %
複数人モデル	60.9 %	57.1 %	44.1 %

った個人モデルと、各クラス300枚、3名分の2100枚を用いた複数人モデルを作成し、それらの認識率を比較した。各人物のテストデータに対する認識結果を表1に示す。実験結果から本人のデータのみで学習を行った場合には認識精度が平均90%程度となり、高い精度で分類を行えたと考えられる。しかし、複数人モデルでは全員の認識精度が下がり、特に被験者Cの低下が著しかった。これは同じ表情でも人によって動く箇所が違ったり顔のパーツの個人差が原因だと考えられる。また、表情間でも精度の差が存在し、笑顔や驚きなどは共通して高い分類精度が見られたが、怒りや嫌悪、恐怖などは精度に差が生じた。これらは表情として大きく動かす点の有無が関係していると思われる。

4. おわりに

本研究ではMediaPipe Face Mesh と SVM を用いて表情を認識する手法を提案した。MediaPipe Face Mesh で画像から取得した468点の3次元座標をSVMで学習させ、7つの表情の分類を行った。評価実験では個人モデルでの分類は高い精度を確認できたが、複数人モデルでは精度が低下した。また、表情間で精度にばらつきがあった。今後の課題として複数人での精度の向上や、表情間の精度差の是正が挙げられる。

【参考文献】

- [1] P. Ekman: Facial action coding system (FACS), A human face, 2002
- [2] 斉藤功樹, 中川靖士: 日本人における基本感情と表情の関係分析, 日本認知科学学会第38回大会, 2021
- [3] W. Sato et al.: Facial expressions of basic emotions in Japanese laypeople, Frontiers in Psychology, vol. 10, 2019