

フェイスマスク着用時の表情認識システム構築に向けた表情変化動画中の有効特徴量分析

竹原 有花[†] 鎌木 崇史[†]

国際基督教大学 教養学部 自然科学部門[†]

1. はじめに

対人コミュニケーションにおいて、人がどのような感情を抱いているかを他者が判断するにあたり人の表情は nonverbal communication として非常に重要な情報とされる。広く知られている説である「メッセージの送り手がどちらとも取れるメッセージを送った場合、内容の解釈にあたりメッセージの受け手が声の調子や身体言語といったものを重視する」というメラビアンの法則などコミュニケーションにおける表情の重要性を主張するものは多く存在する[1]。この事からも表情を初めとした非言語コミュニケーションが非常に重要な役割を担っていることがよく分かる。それ故に表情認識に関する研究は様々な企業でも取り組まれており、今後も顧客の反応の観察等様々な場面での活用が想定される。しかし、昨今の新型コロナウイルスの感染拡大に伴いマスクの着用が一般化した今、顔認識・表情認識の運用にあたって様々な問題が生じている。マスクを着用すると顔の70%近くが覆われてしまい、表情を認識する上で重要な要素である唇、口角、顎、頬など鼻から下の動きの正確な計測が困難であり、従来の表情認識システムでは対応しきれない為である。実際の対人コミュニケーションにおける状況下でも、マスクをした状態では相手の表情が読み取りづらいと感じる人が3割以上いる、という結果が株式会社マンダム[2]の調査で明らかになっており、マスクが円滑な対人コミュニケーション阻害の一因になっていると言えるだろう。このような課題を解決する為マスク着用状態に対応した表情認識システムの構築に先立ち、本研究ではマスク着用時の表情認識において有効な特徴量を検討する。

2. 実験概要

本研究ではフェイスマスクを着用することが

Analysis of Effective Features in Wearing Face Mask Facial Expression Change Videos

[†] Yuka Takehara, Takashi Kaburagi, International Christian University, Dept. of Natural Science



図1. 概要図

表情認識において与える影響を検証する。データセットは The Extended Cohn-Kanade (CK+) を用いる[2]。これは年齢、性別、人種等様々な人物123人の無表情から特定表情への変化動画を30fpsで記録したものであり、表情は怒り・軽蔑・嫌悪・恐怖・喜び・悲しみ・驚きの7種類に分類される。本研究ではこのデータセットにフェイスマスクを自動付与し、OSSであるOpenFaceを用い抽出した特徴量で表情認識を行う。はじめにフェイスマスクの有無で表情認識の精度がどの程度変化するか検証する。さらにフェイスマスク着用時に表情認識を実施するうえで有効な特徴量を検証する。

3. マスク付与と特徴量抽出

3.1 手法

まずはデータセットの動画に対し疑似的なフェイスマスクを付与する処理を行う。これはPythonの画像処理ライブラリであるPILを用い、OpenCVを用いて検出した画像内の顔の位置に合わせて顎から鼻が隠れるようにフェイスマスク画像を付与する。

次にフェイスマスク着用時の特徴量を抽出する目的で、OpenFaceを用いる[3]。OpenFaceは画像や動画に対し顔の行動研究に必要な顔のランドマーク、顔の姿勢、Action Unit(以下AU)の強度等の特徴量を検出する事が可能である。表情認識では顔の筋肉の動きであるAUが用いられる場合が多く、本研究でも表情認識はOpenFaceで検出可能な17種類のAUを用い行う。疑似的フェイスマスク着用動画にOpenFaceを適用し、AU強度情報を時系列データとして取得する。

怒り	0.639	喜び	0.825
軽蔑	0.558	悲しみ	0.724
嫌悪	0.801	驚き	0.733
恐怖	0.694	全体	0.687
参考:フェイスマスク無のデータセット全体	1.000		

表 1. 表情別 Open Face 反応率

3.2 結果

画像総数 3804 枚のデータセットにフェイスマスク画像を付加した結果、3502 枚は顔の下半分に正しくフェイスマスク画像が付与出来た。一方、302 枚は顔以外の物体に反応し 2 枚以上画像が付加されていた。以下の実験では顔自体には影響が無い箇所に付加されている為フェイスマスク付与済みデータセットとしてそのまま用いた。

次に総フレーム数のうち Open Face が反応したフレーム数を算出した。フェイスマスク付与済みデータセットへの Open Face の反応率の結果を表 1 に示す。フェイスマスク無のデータセットに対する Open Face の反応率が 100%である一方で、フェイスマスクを付与したデータセットは 68.7%の反応率となっている。

4. 表情認識

4.1 提案手法

以上の前準備を踏まえ、フェイスマスク着用に対応した表情認識システムの構築にあたり有効といえる特徴量を検討していく。本研究では DNN の中でも時系列データの分析に長ける RNN の一種である LSTM を用い、Open Face で検出した 17 種類の AU の時系列データをもとに表情認識を行う。実装は Google Collaboratory 上で行い、機械学習フレームワークは Keras を用いた。精度の比較は accuracy の値を基準として行う。

フェイスマスク着用下での表情認識に有効な特徴量を検討するにあたり、17 種類の AU(01, 02, 04, 05, 06, 07, 09, 10, 12, 14, 15, 17, 20, 23, 25, 26, 45) の時系列データで構成した特徴量データセットから 1 つずつ特徴量を除き 16 種類の AU で表情認識を行う。乱数シードを固定しない事で精度にばらつきが出る為、30 回ずつ検証を行いその分布から AU がそれぞれどの程度認識精度に影響があるのかを考察する。

4.2 実験結果

はじめに、フェイスマスクの有無でどの程度認識精度が変化するか検証する。フェイスマスク無だと accuracy が 0.914 である一方で、フェイスマスク有りの場合は accuracy が 0.666 とやはりフェイスマスク有りで精度が著しく落ちる事が分かった。これを比較対象とし、フェイスマ

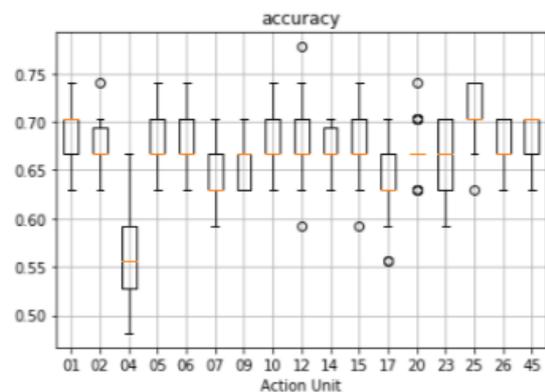


図 2. 取り除いた Action Unit ごとの精度比較

スク着用状態での表情認識に有効な特徴量を検証する。

前述の手法の通りに行った結果が図 2 である。accuracy が低い程除いた AU が認識にあたり重要な特徴量であり、高い程認識には比較的不必要な特徴量であると言える。図 2 によると顕著に精度が下がっているのは AU04 の眉を下げる、反対に精度が上がっているのは AU25 の顎を下げずに唇を開くとなった。つまりフェイスマスクを着用した状態での表情認識においては眉の上下が最も有効な特徴量であり、逆に唇の開閉は最も重要度が低いといえる。実際 AU25 を除いた場合は accuracy が 0.7037, 更に AU25 と中央値が等しい AU01(眉の内側を上げる), AU45(瞬き)を除いた 14 種類の AU を用いると accuracy は 0.7407 となり認識精度が上がった。このため、認識にあたり有効な特徴量の判断に今回の手法が有効であると考えられる。

まとめ

本稿ではフェイスマスク着用状態に対応した表情認識システムの構築に先立ち、フェイスマスク着用時の表情認識において有効な特徴量を検証した。今後は今回検証した AU 情報の他にランドマーク情報やピクセルごとの画素値等画像それ自体の特徴量を加えることでさらなる認識精度の向上を目指す。

参考文献

- [1]株式会社マンドム 『New Normal New Styling』プロジェクト調査報告第二弾” <https://www.mandom.co.jp/release/pdf/2020100501.pdf> (2022.05月12日閲覧).
- [2]Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., and Matthews, I. (2010). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, pages 94–101.
- [3] Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency, OpenFace 2.0: Facial Behavior Analysis Toolkit Tadas Baltrušaitis, *IEEE International Conference on Automatic Face and Gesture Recognition*, 2018

謝辞

本研究は JSPS 科研費 16K16392, 22K12276 の助成を受けたものです。