

カメラを用いた咀嚼センシングアプリの開発

怡土周平[†] 松井良太[†] 杉浦裕太[†]

慶應義塾大学[†]

1. はじめに

咀嚼運動は人間の身体に様々な良い影響を与える。その一方で、咀嚼の左右バランスが偏ることで、顔のゆがみや歯並びなどへの悪影響がある[1]。

咀嚼検出のためのアルゴリズムの評価[2]においては、実験協力者にイヤフォンを装着してもらい、その入力によって咀嚼を検知する手法をとった。また Zhang らは、骨振動センサを搭載した眼鏡型デバイスを実験協力者に装着し、センサ値をもとに咀嚼を検知した[3]。

これらの先行研究によって咀嚼を計測することは可能であるが、課題として、マイクや眼鏡などの周辺機器を装着する必要がある点が挙げられる。これらの機器は煩わしく、実際の食事の妨げになる可能性がある。

本研究の目的は、周辺機器の装着を必要としない、カメラを用いた咀嚼の計測である。我々が提案する手法は以下のとおりである。まず被験者の咀嚼運動をカメラで撮影する。これらの動画データから、画像処理によってパラメータを抽出することで咀嚼を計測する。

2. 実装

被験者の前方に設置されたカメラを用いて咀嚼運動を 30 秒間撮影した。この時、フレームレートは 30 [fps] で固定した。

カメラから得た動画に対し、Google の提供している画像処理系のフレームワークである、Mediapipe で特徴点を抽出した (図 1, 2)。Mediapipe は顔の特徴点計 468 点の x, y, z 座標を取得することができるが、今回は計算コストの削減のため、口回りの特徴点計 10 個の 2 次元座標 (x および y) を抽出した。これらの特徴量を、顔の回転方向や距離成分に関して正規化し、それぞれに対するロバスト性を向上させた。



図 1 右で噛んでいる場合



図 2 右で噛んでいる場合

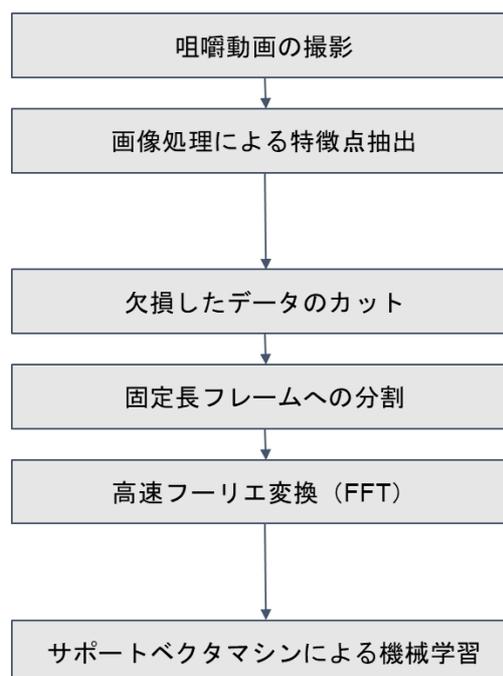


図 3 実装の流れ

また、時間軸が隣り合ったフレーム間の差分により求められる速度成分を、特徴量として追加した。この一連の処理によりデータの次元数は $20 \times 2 = 40$ 次元に増加した。

取得した生データは一般に時間長が固定されていないため、固定長フレームに分割した。ここで 32 フレームを一つのブロックとして分割した。取得した動画のフレームレートは 30 [fps] であるため、およそ 1 秒分のデータが一つのブロッ

Development of mastication sensing application using a camera

Shuhei Ito, Ryota Matsui, Yuta Sugiura, Keio University

クに含まれていることになる。取得した映像の長さは約 30 秒なので、生データの総フレーム数は約 900 であり、これを 30 個の固定長フレームに分割した。この処理の後に、高速フーリエ変換 (FFT) により各次元を周波数成分に変換した。なおこの際、トレンド除去やハン窓適用などの処理を事前に施した。サポートベクタマシン (SVM) による 3 クラス分類器を作成した。ここでいう 3 クラスとは、噛んでいない、左で噛んでいる、右で噛んでいる状態の計 3 つである。これらの実装の大まかな流れを図 3 に示した。

また、咀嚼バランスを出力する方法として、GUI によるフィードバックの方法を考案した。32 フレームごとに分類された識別結果を、左右のバランスとして表示する。左右のバランスが偏っている場合、ユーザにコメントとしてフィードバックする (図 4)。

3. 実験

6 人の実験協力者 (男 : 6 名, 女 : 0 名) に、市販のガムを口に含んでもらい、30 秒の咀嚼動画を撮影した。撮影においては、(1)何も噛んでいない、(2)左で噛んでいる、(3)右で噛んでいるの 3 状態の動画を 5 本ずつ撮影した (計 15 本)。これらの動画データに対し、前章で述べた処理を施したのち、交差検証に基づいて性能を評価した。

検証の結果、平均精度 92.1%での識別に成功した (図 5)。

4. 議論および今後の展望

本研究では、画像処理の処理において、MediaPipe を使用しているため、システムの性能が、MediaPipe の性能に依存してしまう点が制約の一つとして挙げられる。例えば、顔の前に食器があると、顔の特徴点が正しく取得できないといった問題である。また、今回の実験においては、協力者に噛む場所を指示したが、よりリアルな食事シーンに対するシステムの適用を考える必要がある。また、フィードバックの方法として GUI を作成したが、ユーザーにバランスの良い咀嚼を継続していくためのフィードバックの方法も検討していきたい。

5. おわりに

本研究では、カメラを用いて咀嚼を計測した。識別精度は 92.1%であった。今後は GUI 以外でのフィードバックの方法も考える必要がある。

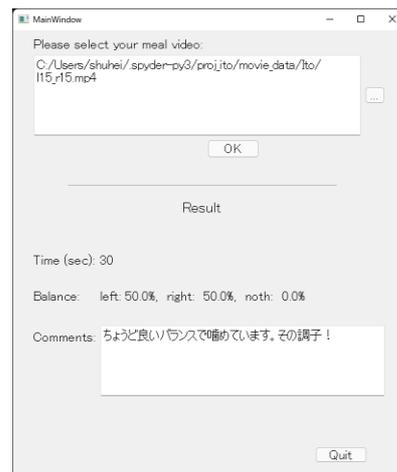


図 4 作成した GUI

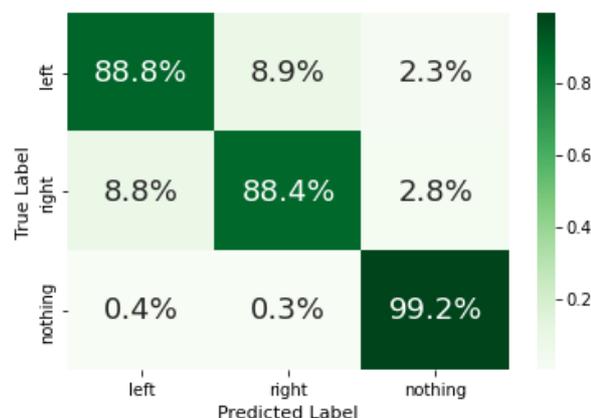


図 5 識別の結果得られた混同行列

謝辞

本研究は、JSPS 科研費 JP21H03485 の助成を受けたものである。

参考文献

- [1] 歯科衛生士 2020 年 9 月号, <https://www.quint-j.co.jp/web/keyword/keyword.php?no=40769> (参照 2022-01-06)
- [2] Sebastian Päßler, Wolf-Joachim Fischer, Evaluation of Algorithms for Chew Event Detection, BodyNets '12: Proceedings of the 7th International Conference on Body Area Networks, 20-26
- [3] Rui Zhang, Oliver Amft, Regular-look eyeglasses can monitor chewing, UbiComp '16: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, 389-392