

# 咀嚼音に基づく咀嚼能力自動評価システム

中村亮裕<sup>†</sup> 津賀一弘<sup>‡</sup> 西村雅史<sup>†</sup>静岡大学大学院総合科学技術研究科<sup>†</sup> 広島大学医系科学研究科(歯)<sup>‡</sup>

## 1. はじめに

口腔機能の維持・向上の観点から、咀嚼能力は咀嚼機能を客観的に評価するための重要な指標である。咀嚼能力の評価方法として、グミゼリー咀嚼時のグルコース溶出量の測定による方法が提案されており、定量的かつ客観的に評価する上で有効であることが報告されている[1]。しかしながら、これらの測定は専門家が衛生状態に配慮しつつ専用の機材を用いて行う必要がある。

口腔内の行動を測定する方法として、簡便に収集可能な音情報を用いる手段が存在する。我々は食事の際に生じる生体音に注目しており、大量の咀嚼・嚥下音を学習させて摂食嚥下行動の識別モデルを構築した[2]。その結果、咀嚼位置(前・左・右)と嚥下について安定した検出性能を得られることを確認した。

ここでは、先に提案した摂食嚥下行動の識別モデルを事前学習モデルとして活用し、グミゼリーの咀嚼音から咀嚼機能を直接推定する方法を提案する。

## 2. グミゼリーを用いた咀嚼能力評価

グミゼリーを用いた咀嚼能力評価手順[1]を説明する。まず、約 2g のグルコース含有グミ(グルコラム:GC 社)を 20 秒間咀嚼する。次に 10ml の水を口に含み、咀嚼したグミと水を、ろ過用メッシュを乗せたカップの上から吐き出す。最後に、咀嚼能力検査システム(グルコセンサーGS-II®:GC 社)を用いて、ろ液中のグルコース溶出量を計測することで、咀嚼能力を評価する。

## 3. 提案手法

### 3.1. 咀嚼・嚥下音を用いた摂食嚥下行動の自動識別モデル

今回事前学習モデルとして用いた摂食嚥下行動の自動識別モデルの概要を説明する。耳下装着の左右マイクから咀嚼・嚥下音を収録し、各イベントを前咀嚼、左咀嚼、右咀嚼、嚥下、その

Automatic Evaluation System of Chewing Ability Based on Chewing Sound

Akihiro Nakamura<sup>†</sup>, Kazuhiro Tsuga<sup>‡</sup>, Masafumi Nishimura<sup>†</sup>

<sup>†</sup> Shizuoka University

<sup>‡</sup> Hiroshima University

他の 5 クラスに分類する。その他には無音区間やノイズが含まれる。

最初に耳下マイクを用いて摂食嚥下時に生じる生体音(咀嚼音、嚥下音)の収録(サンプリング周波数 22,050Hz, 量子化 16bit)を行う。同時に学習モデルの作成のため各イベントに対し正解弱ラベルの付与を行う。弱ラベルは正確な時間情報を持たないラベルである。ラベルの付与は、開発したアプリケーション[2]によって被験者が収録と同時に自己申告で行う。

次に収録した生体音から特徴抽出を行う。特徴量は、フレームシフト 40ms, 窓幅 80ms で抽出した。信号強調のため左右マイクの信号和をとった上で 39 次元の MFCC を求め、これを前後 7 点シフトさせて相互相関係数を求めたものと連結した。

最後にこれらの特徴量を CTC-Transformer[3]に入力して、咀嚼(前・左・右)と嚥下の検出を行う。Attention により文脈を考慮した検出が有効に働くことを確認しており[2]、今回はこれと CTC を併用したモデルを構築した。

### 3.2. 咀嚼音を用いた咀嚼機能の自動推定モデル

グミゼリーを 20 秒間咀嚼した咀嚼音は、咀嚼能力の低下がみられる場合、左右咀嚼が弱まるなど、咀嚼能力の違いによって事前学習した Encoder 出力が変化することが考えられる。図 1 に、グミ咀嚼時に Encoder からフレーム単位で出力される認識対象クラスごとの対数尤度を示した。グミを強く噛み続けた場合、弱く噛み続けた場合と比較して、咀嚼回数が増加し、左右咀嚼が高い尤度で推定されていることが実際にも確認できる。

そこで、グミ咀嚼音の Encoder 出力を用いて、グルコース溶出量を推定するモデルを構築する。モデルの構造を図 2 に示した。

まず、耳下マイクから得られる咀嚼音を 3.1 で事前学習した摂食嚥下行動の自動識別システムに入力し、左咀嚼、右咀嚼、Blank の Encoder 出力をフレーム単位でとる。このとき、スムージングのため、5 フレーム間隔でイベントごとの平均を求める。これらの出力を 3 層 200 次元の全結合層(FC)に入力し、グミゼリーから溶出したグルコース量を推定する。

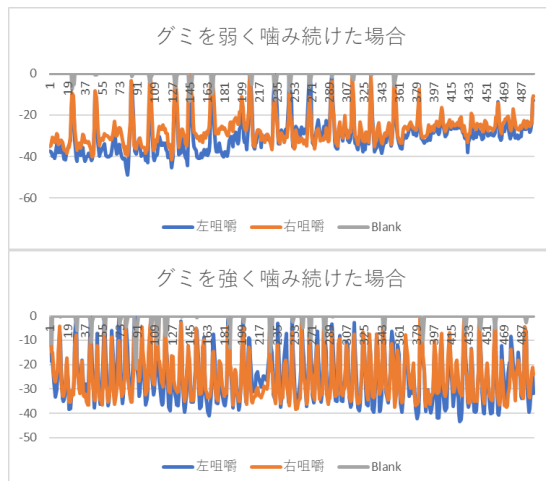


図 1. グミ咀嚼時の認識対象クラスごとの対数尤度

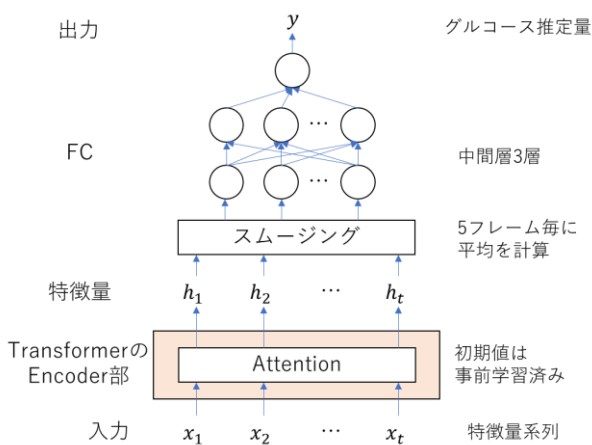


図 2. 咀嚼機能の自動推定モデル

## 4. 実験

### 4.1. データセット

事前学習モデルの構築のため、クラッカー(リッツ), キャベツ(千切り), チューインガムの生体音(咀嚼音・嚥下音)を若年者(20代男女)30名から収録した。同時に被験者の自己申告により弱ラベルの付与を行った。収集した音イベント数は咀嚼と嚥下で約30,000個となった。学習時にはN-gram based augment [2]とSpec augmentにより100倍のデータ拡張を行っている。

グルコース溶出量を推定するモデルの学習・評価のため、強く咀嚼した場合(パターン1)、通常の咀嚼をした場合(パターン2)、弱く咀嚼した場合(パターン3)の3パターンのグミゼリーの咀嚼音を若年者(20代男女)6名から計54回分収録した。収録した咀嚼音からEncoder出力をとった特徴量を入力データ、収録時に計測したグルコース値を正解データとして、FC部の再学習を行う。このとき、話者オープンとするため、話者単位で6分割し交差検証を行う。また、高齢者に対する性能評価のため、高齢者男女5

名から計10回分のグミゼリーの咀嚼音を収録した。

### 4.2. 実験結果・考察

図3にグルコース溶出量の正解値と推定値の分布を示した。若年者の平均絶対誤差率は12.8%、高齢者は7.2%であった。いずれの食べ方を行っても、グミゼリーの咀嚼音から咀嚼能力を高い精度で自動推定できることを確認した。自動推定モデルは、若年者のみの咀嚼音を学習して構築しているが、咀嚼能力が異なる高齢者に対してもグルコース溶出量を安定して推定できることを確認した。

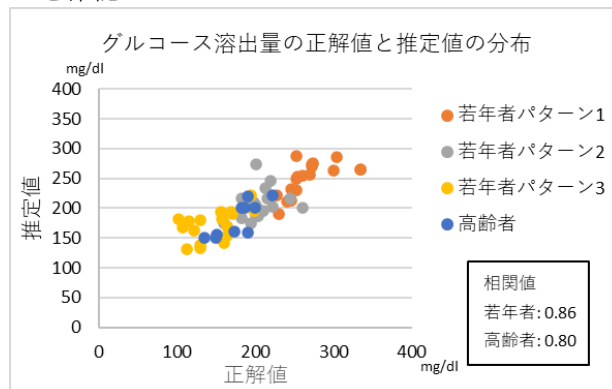


図 3. グルコース溶出量の正解値と推定値の分布

## 5. おわりに

摂食嚥下行動の事前学習モデルのEncoder出力を活用して、咀嚼機能を直接推定するモデルを構築した。その結果、現在、グミゼリー咀嚼時のグルコース溶出量として測定されている咀嚼能力を、咀嚼音の情報だけから推定できる見通しを得ることができた。

今後は検出精度のさらなる向上のため、データ拡張や使用するモデルの検討を行う予定である。

## 謝辞

本研究の一部はJSPS 科研費18H03260, 21K18305の助成を受けたものである。

## 参考文献

- [1] Akira Tanaka et al. “グミゼリー咀嚼時のグルコースの溶出量の分析による運動機能および咀嚼筋筋活動の定量的評価,” 日本補綴歯科学会雑誌, 38(6), pp.1281-1294, 1994.
- [2] Akihiro Nakamura et al. “Automatic Detection of Chewing and Swallowing,” Sensors 2021, Vol.21, Issue 10, 3378, 2021.
- [3] Shigeki Karita et al. “Improving transformer-based end-to-end speech recognition with connectionist temporal classification and language model integration,” Interspeech 2019, pp.1408-1412, 2019.