

# 自然な食事環境に対応した骨伝導音を用いた 咀嚼・嚥下・発話分類手法の提案

近藤 匠海<sup>1</sup> 蒲地 遥<sup>1</sup> 横窪 安奈<sup>1</sup> ロペズ ギヨーム<sup>1</sup>

**概要:** 食事中の咀嚼の回数が少ないと肥満につながる。また、食事中の会話は健康に関連があることより、食事中に会話することが望ましい。他にも、従来の咀嚼の測定装置は大きく持ち運びができないため、日常で使用することは難しい。そこで本研究では、自然な食事環境下での食事行動の定量化を実現することを目的とし、自然な食事環境に対応した咀嚼・嚥下・発話の高精度分類を目標とする。システムは骨伝導マイクロフォンとフィードバックに使用するスマートフォンから構成される。ウェアラブルデバイスを用いた食事行動の定量化の研究は今までも行われているが、実験的な環境でのテストしか行われていない。この研究では、日常的な食事環境での食事音声データを収集し、分類手法を評価、咀嚼・嚥下・発話の分類最適化の手法を提案する。

## Classification of Chewing, Swallowing and Speaking Activities in Daily Meal Environment using Bone-Conduction Microphone

TAKUMI KONDO<sup>1</sup> HARUKA KAMACHI<sup>1</sup>  
ANNA YOKOKUBO<sup>1</sup> GUILLAUME LOPEZ<sup>1</sup>

### 1. はじめに

咀嚼とは、口に取り込んだ食べ物を噛み砕くことである。また、日常生活の中では、食物を嚥下に適した性状に調整するために噛み砕き、唾液と混ぜる口の働きと定義されている[1]。そして、食べる速さが速いほど BMI が高い傾向にあり、早食いと肥満は強い関係にあることが明らかになっている[2]。他にも食事の咀嚼回数が少ないことにより肥満につながることも明らかになった[3]。そのため、肥満を予防するには食事の際にゆっくりとよく噛むように意識することが重要であるといえる。これに加え、食事の会話は健康との関連が見られることがわかっている[4]。よって、食事中に会話をするのが望ましい。

近年、市販されているウェアラブルデバイスはカロリー出力に関連した人間の活動レベルをモニタリングすることが可能である。しかし、自由な食事環境での食事行動を自動的に検出する市販デバイスは存在しない[5]。食事行動の中でも咀嚼や発話を高精度に識別可能になることにより、リアルタイムで咀嚼回数や食事の会話時間を食事者に提示することで、咀嚼回数増加、発話意識向上を実現できると考える。よって、本研究では、自然な食事環境下での食事行動の定量化を実現することを目的とし、自然な食事環境に対応した咀嚼・嚥下・発話の高精度分類を目標とする。本研究での自然な食事環境とは、日常生活における食事環境と定義する。

### 2. 関連研究

#### 2.1 音声をを用いない食事行動解析・認識手法

Zhang らは、特に食べ物を咀嚼することをセンシングするために、食事モニタリングのためのスマートグラスを提案した[6]。提案されたスマートグラスは、3D プリントで製造され、マイクロコントローラ、EMG 電極を収納している。電極配置を分析し、側頭が電極の配置に最適であると確認した。咀嚼の検出は、適合率と再現率が 80%に達し、5 種類の食品の分類精度は 63%から 84%の範囲であった。

Chun らは食事検出のためのネックレス型デバイスを提案した[7]。近接センサで顎骨の動きを捉え、食事と非食事活動を区別し、実験環境では適合率 91.2%、再現率 92.6%と良い結果を示したが、自然な環境では適合率 78.2%、再現率 72.5%と実験環境と比べて 10%以上劣る結果となった。

#### 2.2 音声をを用いた食事行動解析・認識手法

Amft らはマイクロフォンを用いた咀嚼音の解析を行った[8]。耳の内側にマイクロフォンを配置することによって、質の高い咀嚼音を取得できることがわかり、99%の精度で食事を認識することができ、80%以上の精度で 4 種類の食べ物を分類することができた。

安藤らは、集音マイクと咽喉マイクの 2 チャンネルの音声から会話、食事の検出を行った[9]。GMM を用いて咀嚼、嚥下、発話の識別を行い、その識別結果を Long Short-Term Memory(LSTM) を用いることにより、食事と会話の検出を行った。既存手法と比べて食事と会話の検出精度が大幅に

<sup>1</sup> 青山学院大学  
Aoyama Gakuin University

上がったが、食事の検出精度が55%、会話の検出精度が77%と食事の検出精度が会話と比べて20%以上低い。その原因としてGMMを用いた咀嚼、嚥下の識別結果が低いことが挙げられた。

### 2.3 骨伝導音を用いた食事行動解析・認識手法

酒造らは、骨伝導マイクロフォンを利用したICレコーダーを用いた食習慣分析を行っている[10]。解析する音声は、食事時の音声から食べ物を食べている部分を抜き出し、その音声の値に対して絶対値を取り、ローパスフィルタをかけたものを4つの状態（硬い食べ物を食べている、柔らかい食べ物を食べている、水を飲んでいる、喋っている）に分類し、各状態に対して80%以上の分類をすることができた。しかし、解析のためにPC、サンプリングレート44.1kHzの骨伝導マイクロフォンを利用したため、手軽に扱えない。日常生活で扱うことを考慮するのであれば、Bluetooth接続した骨伝導マイクロフォンを用いる必要がある。Bluetooth接続が可能な機種はサンプリングレートが8kHzの物が多いので、それに合わせたアルゴリズムが必要である。

三井らは、食行動を改善するために骨伝導マイクロフォンを用いて咀嚼回数と発話状態をリアルタイムで認識ユーザーにフィードバックするシステムを提案した[11, 12]。異なる食感であるキャベツとおにぎりをを用いて、リアルタイムでの咀嚼判定精度を約96%で判定でき、発話の時間測定精度も約96%という結果ができた。しかし、提案したシステムを使いたくないという否定的な意見や、どの食材にも対応できないことが問題点として挙げられる。接触型マイクロフォンは通常のマイクロフォンと比較し体の表面から直接内部振動を補足し、周囲のノイズに影響されない。また咀嚼検出が80%の精度を超えているので、咀嚼検出に適したマイクロフォンだといえる。しかし、実験環境で限定した食べ物をを用いた咀嚼検出を行っているため、自然な食事環境ではない。

### 2.4 まとめ

音を用いない食事検出を行うアプローチは、市販されていて一般人でも容易に手に入るデバイスを用いず、専門知識がある人間にしか製造できないようなデバイスを利用している。

Amftらの研究によって耳の内側にマイクロフォンを配置することにより、質の高い咀嚼音が取得できるということが判明したので、耳の内側にマイクロフォンを配置することが理想であると考えられる。しかし、先行研究では、実験環境で限定した食べ物をを用いた咀嚼検出を行っているため、自然な食事環境ではない。また、骨伝導マイクロフォンを用いて咀嚼、嚥下、発話の食事詳細行動の識別を行った研究はない。

以上から本研究では、食品を限定しない自然な食事環境での咀嚼・嚥下・発話の分類を行いたいため、接触型で耳

の内側に配置でき、容易に手に入る市販の骨伝導マイクロフォンを利用する。

## 3. データセット作成と分類手法の検討

本研究ではリアルタイムでの分類は行わないため、咀嚼、嚥下、発話に該当する音声区間を手動で抽出し、分類に用いる。

### 3.1 自然な食事環境下での食事音声データ収集

本研究では自然な食事環境下での咀嚼・嚥下・発話の分類を行うために、自然な食事環境下での食事音声データの収集を行った。

食事音声データの収集にはスマートフォンとBluetooth通信を行う骨伝導マイクロフォンを使用した。スマートフォンはGoogle社製のGoogle Pixel 3、Bluetooth通信を行う骨伝導マイクロフォンはMotorola社製のMOTOROLA Finiti HZ800 Bluetooth Headsetを用いた。骨伝導マイクロフォンのBluetooth規格は2.1であった。Androidアプリケーションソフトウェアにより食事音声データを収集した。

データ収集では、19±5歳の男性、女性6名を被験者とした。図3.1が示すように、被験者の片耳にBluetooth通信を行う骨伝導マイクロフォンを装着して行った。また、音声データ収集後に、咀嚼、嚥下、発話に該当する音声区間をラベリングするため、ラベリング作業を補助するために音声データと一緒に動画を撮影した。動画を撮影する際、被験者の口と喉が映るように撮影を行った。データ収集は、実験室環境ではなく食堂や一般家庭の食卓、研究室など自由な食事環境で行い、被験者には日常生活と同じように食事をするよう求めた。また、食事の開始から収集した場合や、途中から収集した場合があり、収集したデータ時間にはばつきがある。



図 3.1. データ収集の様子

### 3.2 音声データのラベリング

収集したデータの整理は、咀嚼、嚥下、発話の分類を行う際に用いる真値のデータセットを作成することを目的とした。真値のデータセットを作成するために、骨伝導マイクروفオンを用いて収集した音声データを咀嚼、嚥下、発話に該当する区間の音声をラベリングした。

骨伝導マイクروفオンにより収集した音声データをラベリングする際に、音声のみでは咀嚼、嚥下、発話のラベリングを行うことが困難であるため、音声データと、音声データ収録と同時に撮影した動画を同期させてラベリングを行った。撮影した動画は同期させることにより、映像音声から骨伝導マイクروفオンにより収集された音声データの音声に変換した。

音声データのラベリングは、音声にラベルを付与できる音声分析用ソフトウェアの Praat を用いて行った[13]。ラベルの種類は、「咀嚼・嚥下・発話」の3種類に設定した。

咀嚼、嚥下、発話のラベルを付与された音声区間のみを抽出しデータセットを作成した。表 3.1 に示すように、咀嚼、嚥下、発話のラベルデータ数は咀嚼が他に比べて圧倒的に多く不均衡であるということがわかる。

表 3.1. ラベルデータの詳細

被験者	食事時間 [min] (記録回数)	咀嚼	嚥下	発話	合計記録 時間 [min]
1	10:41(5)	487	55	149	41:59
2	3:01(1)	164	10	30	7:25
3	5:43(1)	496	35	36	5:43
4	2:02(1)	61	13	53	8:39
5	1:27(1)	102	2	8	1:27
6	3:03(1)	202	8	50	8:21
合計	25:57	1512	123	326	73:34

### 3.3 特徴量抽出

抽出した特徴量について表 3.2 に示す。表 3.2 が示すように、生データのカテゴリは、生データのデータ時間と振幅のピーク数と零交差数を用いて合計 3 個抽出した。これらを用いた理由は、咀嚼、嚥下と比べ発話の時間が長く、振幅数が多いと考えたからである。

パワーのカテゴリは、パワーの最大値を抽出した。パワーに関しても、発話が他のデータと比べパワーの値が大きいことと、嚥下が他のデータと比べパワーの値が小さいことから抽出した。

表 3.2. 抽出された 75 個の特徴

特徴カテゴリ	説明	特徴数
生データ	データの長さ	3
	ピーク数	
	零交差数	
パワー	最大値	1
Amplitude difference accumulation	全データの合計	2
	最大値	
Short term energy	全データの合計	2
	最大値	
移動中央値	零交差数	1
パワースペクトル密度(PSD)	特定の周波数範囲の合計	
	中央周波数	
	特定の周波数範囲のバンドパワー	23
	特定の周波数範囲の最大値になる	
	周波数	
一番と二番目に大きいピーク値の周波数		
相互相関	生データ同士の相互相関の最大値	4
	PSD 同士の相互相関の最大値	
MFCCs	メル周波数ケプストラム係数	39

次に、音声信号を特徴化させるために Amplitude Difference Accumulation (ADA) をかけた。処理したい信号を  $x$ 、 $N$  を各フレームのサンプル数、 $n$  番目の ADA として式 3.1 に示す[14]。

$$ADA_n = \sum_{m=(n-1)*N+1}^{n*N} |x(m) - x(m-1)| \quad (3.1)$$

同じく、音声信号を特徴化させるために Short Term Energy (STE) をかけた。処理したい信号を  $s$ 、サンプル数のシフトを  $n$ 、窓関数を  $w(n)$ 、として式 3.2 に示す。

$$e(n) = \sum_{m=-\infty}^{\infty} (s(m) * w(n-m))^2 \quad (3.2)$$

Amplitude Difference Accumulation と Short Term Energy カテゴリは、ADA と STE によって変換されたデータのデータ値の合計、最大値をそれぞれ用い、2 個ずつ特徴量として抽出した。

移動中央値カテゴリは、移動平均をかけた後の零交差数を用いて、1 個の特徴量を抽出した。これも生データの特徴化させ、咀嚼、嚥下、発話の振幅数に違いが出たため利用した。

また、生データのパワースペクトル密度も特徴として利用した。パワースペクトル密度は、パワースペクトルを平滑化するウェルチ法により変換した。パワースペクトル密度を用いた理由は、音声認識に頻繁に用いられているため

である。本研究では生データのパワースペクトル密度のデータ値の合計、中央周波数、バンドパワー、パワースペクトル密度のピーク値の中で一番目と二番目に大きい周波数を特徴として用い、5 個抽出した。また、特定の周波数範囲のパワースペクトル密度のデータ値の合計、バンドパワー、中央周波数を用いて 18 個抽出し、以上パワースペクトルのカテゴリでは合計 23 個の特徴量を抽出した。

分類精度向上のために、咀嚼、嚥下の生データからそれぞれ一つデータを選択し、そのデータとすべての生データ、パワースペクトル密度を相互相関にかけた。それぞれの相関結果の最大値を特徴量として利用した。よって、合計 4 個の特徴量を抽出した。

本研究では、音声認識で頻繁に利用されるメル周波数ケプストラム係数 (MFCC) も特徴として利用した。先行研究で用いた 39 次の MFCC を特徴量として抽出した[9]。

以上より合計 75 個の特徴量を分類のために抽出した。

### 3.4 不均衡データの均衡化

ラベリング作業により作成されたデータセットは、表 3.1 で示すように、咀嚼のデータ数は多いが、咀嚼以外のデータは咀嚼と比べて非常に少ない。よって、不均衡なデータセットを均衡にさせるため、SMOTE (Synthetic Minority Oversampling TEchnique) を用いた[15]。SMOTE は Chawla らによって提案されたマイノリティクラスのオーバーサンプリングとマジョリティクラスのアンダーサンプリングを組み合わせた手法である。SMOTE を用いた結果、表 3.3 で示すように、咀嚼以外のデータも咀嚼のデータ数と遜色ないデータ数になった。また、均衡化されたデータセットを 4:1 で訓練データとテストデータに分割した。

表 3.3. SMOTE によりサンプリングされたラベルデータの詳細

ラベル名	ラベルデータの合計	訓練データ	テストデータ
咀嚼	1512	1209	303
嚥下	1512	1209	303
発話	1512	1210	302
合計	4536	3628	908

### 3.5 分類器毎の精度検証

本研究では、骨伝導音データを用いて咀嚼・嚥下・発話の分類を行うために、機械学習を用いる。機械学習の手法は、教師あり学習と教師なし学習の 2 種類ある。既に咀嚼、嚥下、発話とラベリングされたデータを用いるので、既知の入出力データを用いてモデルを訓練し、出力を予測できる教師あり学習を用いる。さらに、教師あり学習は分類や回帰の手法を用いてモデルを作成するが、本研究では咀嚼、嚥下、発話の分類を行うので分類モデルを使用する。

一般的な教師あり学習の分類器は、サポートベクタマシン (SVM)、決定木、KNN、アンサンブル分類器などがあるが、分類を行うにあたって分類器を選定する必要がある。

どの分類器が最適なモデルであるかを調べるために Matlab にあるアプリケーションの「分類学習器」を利用する。分類学習器は、自動化された学習を実行して、最適な分類モデルのタイプを探することができる。

分類学習器を用いた分類器選択の手順を以下に述べる。まず初めに、アプリケーションを実行するには既知の入出力データセットとデータセットに対するラベルを分類学習器に与える必要がある。よって、SMOTE により均衡化された咀嚼・嚥下・発話の特徴量データセットとこのデータセットに対する咀嚼、嚥下、発話のラベルを与える。また、検証方法を指定する必要がある。五分割交差検証を指定した。次に、与えたデータに対して、決定木、線形 SVM、ガウス SVM、最近傍分類器、バギング決定木の分類モデルに自動学習させた。分類学習器は、モデルの自動学習後に各モデルの指定した検証方法でのスコアが出力される。本研究では、五分割交差検証を指定したので五分割交差検証の平均精度が出力された。最後に、分類学習器の出力結果により咀嚼、嚥下、発話の分類に用いるモデルを選定した。出力された各モデルの平均精度を表 3.4 に示す。

表 3.4. 各モデルの交差検証精度

分類モデル		精度[%]
決定木	複雑な木	83.9
	中程度の決定木	74.8
	粗い木	67.8
サポートベクタマシン	線形 SVM	84.2
	細かいガウス SVM	86.8
	中程度のガウス SVM	97.6
最近傍分類器	粗いガウス SVM	82.4
	細かい KNN	90.3
	中程度の KNN	82.2
	粗い KNN	64.5
	コサイン KNN	88.5
	3 次 KNN	81.0
アンサンブル分類器	重み付き KNN	84.7
	ブースティング決定木	81.7
	バギング決定木	95.5

表 3.4 で示すように、中程度のガウス SVM が一番平均精度が高いことがわかる。よって、自然な食事環境下での骨伝導音を用いた咀嚼・嚥下・発話の分類のために、ガウシアンカーネル (rbf カーネル) を用いた SVM を選定した。

## 4. 分類性能評価

自然な食事環境下での骨伝導音データを用いた咀嚼・嚥下・発話の分類モデルの最適化と分類性能評価について述べる。評価の目的は、選択された分類器のパラメータを高精度な分類が実現するように調整することである。また、将来リアルタイムでの分類を行う際に、特徴選択を行うことにより分類に関連しない特徴を削減することによって計算量を削減することである。

### 4.1 特徴選択

本研究で用いる学習モデルは、rbf カーネルを用いた SVM である。SVM を利用する際に、特徴量をスケール変換する必要がある。本研究では、StandardScaler を使い、個々の特徴量の平均が 0 で分散が 1 になるように変換した。より高精度な分類を実現するためにパラメータの調整を行う必要がある。rbf カーネルのパラメータは、正則化パラメータ  $C$  とガウシアンカーネルの幅の逆数を表す  $\gamma$  である。 $C$  と  $\gamma$  は両方ともモデルの複雑さを制御するパラメータで、大きくするとより複雑なモデルになり、2 つのパラメータの設定は強く相関する。よって、 $C$  と  $\gamma$  は同時に調整しなければならない[16]。本研究ではパラメータの調整にグリッドサーチと呼ばれる手法を用いる。グリッドサーチを用いて、合計 36 通りの  $C$  と  $\gamma$  の組合せで 5 分割交差検証を行った。グリッドサーチの出力結果では、75 個の特徴量を用いた SVM の最適なパラメータは  $C$  が 10、 $\gamma$  が 0.01 となった。

咀嚼・嚥下・発話の分類のための学習モデルに特徴量を 75 個使用した。それらの特徴の中には、この分類とは関係の薄い特徴があるのではないかと考えた。よって、使用する特徴数を削減するために、特徴選択を行った。特徴選択には Sequential Feature Selector を用いる。Sequential Feature Selector には 4 つのアルゴリズムがあり、我々は Sequential Forward Floating Selection を使用した。Sequential Forward Floating Selection は、ラッパー法であり、最初は特徴が 1 つも選択されていない状態から始まり、最もモデルを改善する特徴を追加し、削除すべき特徴があるか探す工程を指定した特徴数になるまで繰り返す。本研究では、Sequential Forward Floating Selection で指定する特徴数を 30 に設定して実行した。30 に設定したのは、計算量削減のために特徴数を半分以下にするためである。削減実行結果を図 4.1 に示す。

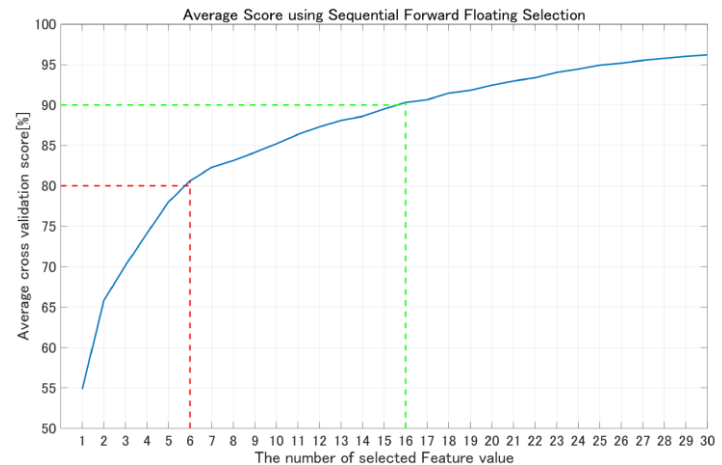


図 4.1. Sequential Forward Floating Selection を用いた特徴数による精度推移

先行研究でリアルタイムでの咀嚼検出精度が 80%以上だと、咀嚼回数のフィードバックシステムに対してユーザは違和感を感じないという結果が出た[11]。また、90%以上あるとフィードバックを行う必要のないヘルスケアや医療の分野、自動食事モニタリング実現に貢献すると考える。図 4.1 で示すように、特徴数が 6 のときに交差検証スコアが 80%を超え、16 のときに 90%を超えることが分かる。また、特徴数が 6 個、16 個のときに選択された特徴について表 4.2、表 4.1 に示す。

表 4.1. 特徴選択された 16 個の特徴

特徴カテゴリ	説明	特徴数
生データ	データの長さ ピーク数	2
Amplitude difference accumulation	全データの合計	1
Short term energy	最大値	1
パワースペクトル密度	特定の周波数範囲の合計 特定の周波数範囲の最大値になる 周波数	2
相互相関	生データ同士の相互相関の最大値	1
MFCCs	メル周波数ケプストラム係数	9

表 4.2. 特徴選択された 6 個の特徴

特徴カテゴリ	説明	特徴数
生データ	データの長さ ピーク数	2
Short term energy	最大値	1
パワースペクトル密度	特定の周波数範囲の最大値になる 周波数	1
相互相関	生データ同士の相互相関の最大値	1
MFCCs	メル周波数ケプストラム係数	1

#### 4.2 特徴選択された特徴のみを用いた SVM のパラメータ調整

特徴選択された特徴のみを用いた SVM のパラメータ調整を行う。今回もグリッドサーチを用いて、rbf カーネルを用いた SVM のパラメータ調整を行う。特徴数による性能比較を行うために、特徴選択での交差検証スコアが約 80% の 6 個の特徴、同じく約 90% の 16 個の特徴、特徴選択された 30 個の特徴を用いてパラメータ調整をそれぞれ行う。合計 36 通りの C と gamma の組み合わせで 5 分割交差検証を行った。30 個の特徴を用いたときの結果を図 4.2 に示す。

特徴数が 30 個の場合、C が 10、gamma が 0.1 のときが最適なパラメータであると示された。

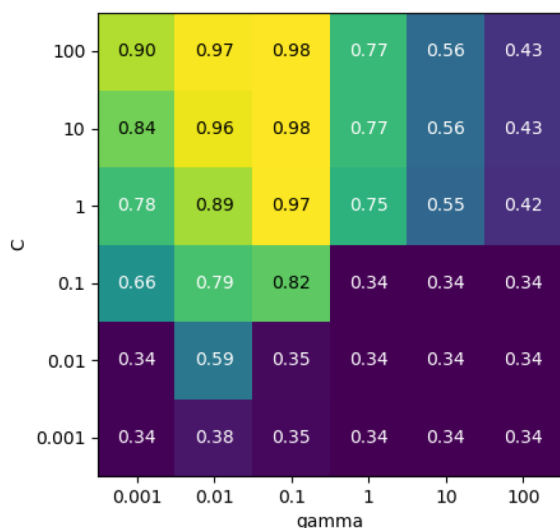


図 4.2. 30 個の特徴を用いたときの、SVM の C と gamma の影響による交差検証精度

#### 4.3 テストデータによる SVM の汎化性能結果

テストデータを用いて、最適なパラメータに調整した SVM の汎化性能結果を出した。汎化性能とは、モデルが未知のデータに対して正確に予測する能力のことである。訓練データセットを用いた交差検証では精度が高かったが、

未知のテストデータセットを用いてのモデルの予測精度は高いとは限らない。特徴選択前の 75 個の特徴を用いた SVM と特徴選択された特徴のみを用いた SVM、この各モデルの汎化性能結果を出した。また、特徴選択された特徴のみを用いた SVM では、各モデルの上位 5 つのパラメータの組み合わせの汎化性能結果を出した。結果はそれぞれのクラスの F1 値を出した。特徴数ごとの汎化性能評価結果を表 4.3 に示す。

表 4.3 特徴数ごとの汎化性能評価結果

特徴数	F1 値			
	6	16	30	75
咀嚼	0.84	0.92	0.98	0.98
嚥下	0.88	0.95	0.99	0.98
発話	0.94	0.97	1.00	0.99

汎化性能評価結果より、16 個と 30 個の特徴を用いた場合は、F1 値が 90% 以上と非常に良い結果であった。6 個の場合も F1 値が 90% は超えなかったが、80% を超え、リアルタイムでの咀嚼・嚥下・発話の分類をフィードバックを行うには問題のない結果であるといえる。

## 5. まとめ

本研究では、自然な食事環境に対応した骨伝導音を用いた咀嚼、嚥下、発話の分類手法を提案した。rbf カーネルを用いた SVM により、咀嚼、嚥下、発話の分類を行った。75 個あった特徴量を特徴選択により削減し、特徴選択の際に出た交差検証スコアが 80%、90% の 6 個、16 個と特徴選択を行う際に設定した 30 個と 75 個の汎化性能を比較した。結果としては、特徴数が 6 個のときは F1 値が 84%、88%、90% とリアルタイムでのフィードバックを行うには問題のない結果となった。16 個と 30 個は F1 値が 90% を超え非常に良い結果となった。

今後の展望としては、本研究で用いた分類モデルを利用し、骨伝導マイクロフォン、スマートフォンを用いてリアルタイムでの咀嚼、嚥下、発話の分類を目指す。

## 参考文献

- [1] 日本咀嚼学会. 日本咀嚼学会からの発信. <http://sosyaku.umin.jp/info/file/info01.pdf>
- [2] 安藤雄一, 花田信弘, 柳澤繁孝. 「ゆっくりとよく噛んで食べること」は肥満予防につながるか? ヘルスサイエンスヘルスケア, Vol. 8, No. 2, pp. 51-63, 2008.
- [3] 岩崎正則, 葭原明弘, 宮崎秀夫. 成人期および高齢期における咀嚼回数と体格の関連. 口腔衛生学会雑誌, Vol. 61, No. 5, pp. 563-572, 2011.
- [4] 岸田典子, 上村芳枝. 学童の食事における会話の有無と健康及び食生活との関連. 栄養学雑誌, Vol. 51, No. 1, pp. 23-30, 1993.
- [5] Shengjie Bi, Tao Wang, Nicole Tobias, Josephine Nordrum, Shang

- Wang, George Halvorsen, Sougata Sen, Ronald Peterson, Kofi Odame, Kelly Caine, et al. Auracle: Detecting eating episodes with an earmounted sensor. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol.2, No. 3, p. 92, 2018.
- [6] Rui Zhang, Severin Bernhart, and Oliver Amft. Diet eyeglasses: Recognising foodchewing using emg and smart eyeglasses. In *Wearable and Implantable Body Sensor Networks (BSN), 2016 IEEE 13th International Conference on*, pp.7-12. IEEE,2016.
- [7] Keum San Chun, Sarnab Bhattacharya, and Edison Thomaz. Detecting eating episodes by tracking jawbone movements with a non-contact wearable sensor. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 2, No. 1, p. 4, 2018.
- [8] Oliver Amft, Mathias Stäger, Paul Lukowicz, and Gerhard Tröster. Analysis of chewing sounds for dietary monitoring. In *International Conference on Ubiquitous Computing*, pp. 56-72. Springer, 2005.
- [9] 安藤純平, 斎藤隆仁, 川崎仁嗣, 片桐雅二, 池田大造, 峰野博史, 西村雅史ほか. 咽喉音を利用した会話・摂食行動の認識. マルチメディア, 分散協調とモバイルシンポジウム 2017 論文集, Vol. 2017, pp. 116-123, 2017.
- [10] Masaki Shuzo, Shintaro Komori, Tomoko Takashima, Guillaume Lopez, Seiji Tatsuta, Shintaro Yanagimoto, Shin'ichi Waris-awa, Jean-Jacques Delaunay, and Ichiro Yamada. Wearable eat-ing habit sensing system using internal body sound. *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, Vol. 4, No. 1, pp.158-166, 2010.
- [11] 三井秀人, 中野和哉, 磯山直也ほか. 咀嚼回数向上支援システム. マルチメディア, 分散協調とモバイルシンポジウム 2017 論文集, Vol. 2017, pp. 1733-1737, 2017.
- [12] 三井秀人, 小原文, 横濱安奈ほか. 骨伝導音を用いたリアルタイム咀嚼・発話判定精度向上手法の提案. マルチメディア, 分散協調とモバイルシンポジウム 2018 論文集, Vol. 2018, pp. 562-566, 2018.
- [13] Praat. <http://www.fon.hum.uva.nl/praat/>.
- [14] Hao Zhang, Guillaume Lopez, Masaki Shuzo, Jean-Jacques Delaunay, and Ichiro Yamada. Mastication counting method robust to food type and individual. In *HEALTHINF*, 2012.
- [15] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, Vol. 16, pp. 321-357, 2002.
- [16] Sarah Guido Andress C. Muller. Python ではじめる機械学習. O'Reily, 2017.