

嚥下音による飲水量モニタリング

2015年3月

中藤 寛己

嚥下音による飲水量モニタリング

中藤 寛己

概要

食物を摂取する際には、喉頭において嚥下音が発生する。嚥下音は体の外側から非侵襲的に計測することができるため、嚥下機能の評価方法として応用が期待されている。嚥下音を解析することで発生機序や音響特性等の知見を得る研究が行われており、誤嚥を検出するシステムの構築も試みられている。しかし、嚥下音を解析することで食物を定量的に計測する研究は少ない。嚥下量の計測が可能となれば、食生活のモニタリングによって病気の予防等へ応用されることが期待される。

本研究では、嚥下音に対して信号処理を行うことで、被験者がどの程度の水分を摂取したかを定量的に計測するシステムの構築を目的とする。まず、本システムでは甲状軟骨の表面にマイクロフォンを接触させ、水を摂取した際の嚥下音を計測する。計測した嚥下音に対して、平滑化、線形予測分析、フーリエ変換、ウェーブレット変換の四つの信号処理を施し、得られた結果から一度の嚥下での摂取水分量を反映し得る 160 個の特徴量を抽出する。次に、主成分分析によって特徴量の低次元化を行う。さらに、摂取した水分量の真値が既知である嚥下音信号の特徴量を学習用データとして用い、サポートベクタマシンによって摂取水分量を推定するための識別器を決定する。最後に、摂取した水分量が未知の嚥下音を、決定した識別器によって分類していくことで摂取した総水分量の推定を行う。

提案手法の性能を確認する実験を行った。まず、被験者 3 名に対して容量 5 mL、10 mL、15 mL の水の嚥下音を各 50 回計測した。摂取水分量を推定するための識別器を一つ抜き交差法により決定し、未知データに対する水分量の推定を行った結果、5 mL と 10 mL、15 mL の嚥下音を、各被験者でそれぞれ 70.0 %、66.7 %、96.7 % の精度で推定することができた。次に、50 mL と 100 mL の水を複数回の嚥下にて摂取した際の嚥下音を各 10 回計測した。抽出した各嚥下音を決定した識別器によって 5 mL と 10 mL、15 mL へ分類し、摂取水分総量を推定した結果、本システムによって摂取水分総量を推定できることが示唆された。

キーワード

ヘルスケア、生体情報処理、嚥下音、摂取水分量、主成分分析、サポートベクタマシン

Monitoring of Ingested Water from Swallowing Sound

Hiroki Nakafuji

Abstract

When human beings ingest something, swallowing sounds occur at the larynx. Since swallowing sounds can be measured noninvasively, the sounds can be used to examine swallowing ability. Researchers have been trying to obtain knowledges about what features swallowing sounds have. In addition, researchers are intended to make systems to detect aspirations. However, there are few researches which measure how much food or drink is swallowed by monitoring eating habits.

The purpose of this research is to estimate how much water is swallowed by the analysis of swallowing sounds. First in the proposed system, swallowing sounds are measured on the skin surface of the thyroid cartilage. Second, signal processing methods such as smoothing and wavelet transform are applied to the sounds. As the results, 160 feature amounts, which may reflect how much water is swallowed, are extracted, and the number of amounts is reduced by principal component analysis. Finally, the discriminator is determined by support vector machines from learning data which true value is known in advance, and unknown data are classified using a discriminator.

In these experiments, swallowing sounds were measured when water of the volume of 5 mL, 10 mL, 15 mL was swallowed, respectively 50 times from 3 subjects. The discriminator determined by using the proposed system was evaluated by leave-one-out method, and three kinds of swallowing sounds can be estimated with the probability of 70 % , 66.7 % , 96.7 % for the respective subjects. In addition, continuous swallowing sounds were measured. The volume of water as for respective sounds were estimated by the discriminator, it was concluded that this system can estimate total swallowed water.

Key words

Health care, Biomedical information processing, Swallowing sound, Volume of swallowed water, Principal component analysis, Support vector machine

目次

第1章	はじめに	1
第2章	嚙下に関する解析	3
2.1	嚙下機能	3
2.2	嚙下機能の解析手法	4
2.3	嚙下音の特徴	10
第3章	摂取水分量の推定システム	12
3.1	システムの概要	12
3.2	音声信号の計測と処理	13
3.3	音声特徴量の解析	23
第4章	嚙下音の計測と解析結果	31
4.1	嚙下音の計測	31
4.2	音声信号処理による特徴量の抽出	34
4.3	抽出した特徴量の解析結果	41
4.4	連続的嚙下音の総摂取水分量の推定	45
4.5	考察	48
第5章	おわりに	50
	謝辞	51
	業績	52
	参考文献	53

目 次

1.1	日常での生活習慣のモニタリング	2
2.1	一嚥下の過程	4
2.2	VF 検査による嚥下解析	5
2.3	超音波画像による嚥下解析	6
2.4	インピーダンス咽頭図による嚥下解析	7
2.5	嚥下加速度による解析	8
2.6	嚥下音による解析	9
2.7	嚥下音を構成する要素	10
2.8	嚥下音信号	11
3.1	処理の流れ	12
3.2	嚥下音の計測位置	13
3.3	信号の標本化	14
3.4	信号の量子化	14
3.5	音声信号	15
3.6	音声信号の平滑化	16
3.7	平滑化結果からの特徴量の抽出	17
3.8	フーリエ変換結果からの特徴量の抽出	19
3.9	フーリエ変換による時間周波数解析	20
3.10	マザーウェーブレット関数	21
3.11	ウェーブレット変換による時間周波数解析	21
3.12	ウェーブレット変換結果からの特徴量の抽出	22
3.13	多クラス SVM の例	29
4.1	被験者実験の風景	31

4.2	摂取する水分量	32
4.3	使用したマイクロフォン	32
4.4	計測した嚙下音信号の例	33
4.5	整流した信号	34
4.6	平滑化した信号 ($l_1 = 300$)	35
4.7	平滑化した信号 ($l_2 = 600$)	36
4.8	次数による FPE の推移の例	37
4.9	フーリエ解析結果	38
4.10	ガボールウェーブレット関数	39
4.11	ウェーブレット変換結果	40
4.12	RBF カーネルを用いた SVM による分類結果	43
4.13	多項式カーネルを用いた SVM による分類結果	44
4.14	50 mL と 100 mL の嚙下音信号	46
4.15	被験者 3 名における摂取水分量の推定結果	47
4.16	異なる 10 名の被験者における摂取水分量の推定結果	48

表 目 次

4.1	マイクロフォンの仕様	32
4.2	各主成分における寄与の大きい特徴量	41
4.3	線形カーネルにおける分類結果	42
4.4	被験者 A の未知データの分類結果	45
4.5	被験者 B の未知データの分類結果	45
4.6	被験者 C の未知データの分類結果	45

第1章 はじめに

古来より、人類にとって水は不可欠なものであった。古代のエジプト文明はナイル川に、メソポタミア文明はチグリス・ユーフラテス川に沿って誕生するなど、文明は水と共に発達してきた。日常生活においても、炊事、洗濯、水浴など、生活を支える様々な活動には水が用いられている [1]。しかし、ヒトにとって水は、飲料として最も重要な物質である。ヒトは様々な物質によって構成されているが、中でも水は生体の質量の約 60% を占めており、ヒトの生命活動を維持するためには必要不可欠な物質である [2]。

ヒトは生命活動の維持のために、体内のガス交換を行う呼吸や、食物を口から胃まで輸送するための動作である嚥下を通じて、水に限らず空気や食物も取り入れる。したがって、呼吸音や咀嚼音等、様々な生体音を口腔付近で発する [3][4]。特に、ヒトが食物を摂取する際に、喉頭において嚥下音が発せられる [5]。老化や後遺症などの影響で嚥下機能が低下し、嚥下の際に食物が誤って気道に入る誤嚥を原因として、肺炎を引き起こす等の症状は嚥下障害と呼ばれる [6]。嚥下音は体の外側から非侵襲的に計測することが可能であるため、嚥下障害の検査方法へ応用が期待されている [7]。嚥下音の解析を行うために、嚥下音の発生機序や音響的特徴に関して多くの研究が行われており [8]、嚥下音の音響的特徴を利用して誤嚥の検出も試みられている [9]。また、嚥下音は周囲の雑音や生体自身から発生する雑音の影響を大きく受けるため、嚥下音のみを正確に検出する研究も行われている [10]。さらに、嚥下音は簡便に計測が可能であることから、図 1.1 に示すように日常における生活習慣、特に食生活のモニタリングの一手法として応用が期待されている [11]-[14]。食生活のモニタリングという観点においては、どの程度の量を摂取したかが重要な指標の一つである。しかし、嚥下回数を計測する [15]、嚥下した食塊の種類を推定する [16] 等の研究は行われているが、嚥下した食塊を定量的に計測する研究は、侵襲的な手法を用いた研究に限られ [17][18]、嚥下音を用いた非侵襲的な手法は確立していない。

本研究では、嚥下音を解析することで、摂取水分量を定量的に計測するシステムの構築を目指す。摂取した水分量の計測が可能となれば、人工透析を受ける人や痛風を患う人など、

一日に摂取できる水分量が制限される患者に対し、摂取水分量のモニタリングを行うことで水分補給への補助を行うことが可能となる。本論文では、計測した嚥下音に対する音声信号処理結果から、一度の嚥下での摂取水分量を反映する特徴量を抽出し、5 mL の分解能で摂取水分量を推定する手法を提案する。次に、提案した手法を用いて嚥下音に対して摂取水分量の推定を試みた結果について述べる。さらに、一定量の水を連続的に嚥下した際の嚥下音を計測し、摂取した水分総量を推定した結果についても述べる。

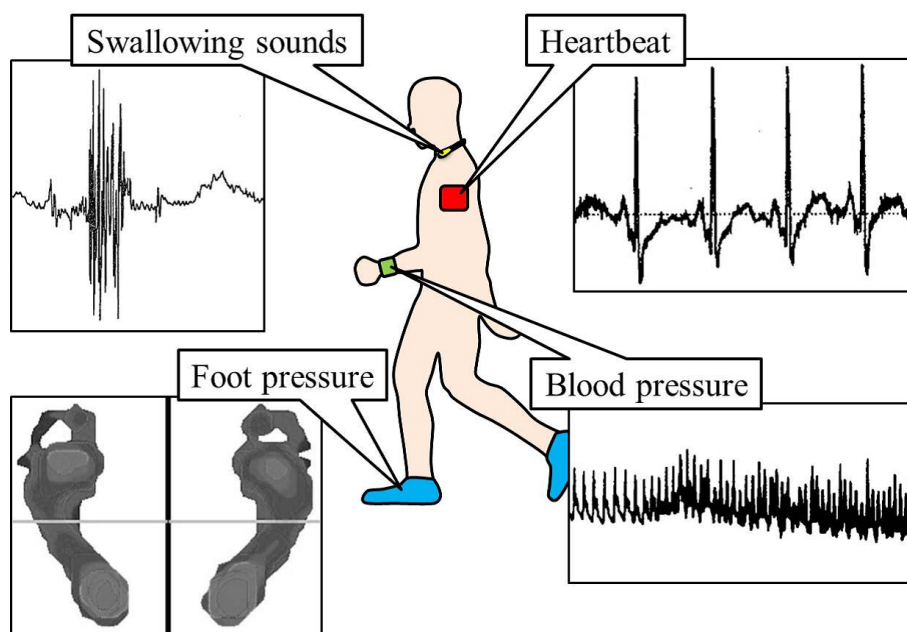


図 1.1: 日常での生活習慣のモニタリング [11]-[14]

第2章 嚥下に関する解析

嚥下は様々な器官が関わることで行うことが可能となる。本章では、嚥下が行われる機序について述べ、嚥下造影や超音波等の嚥下機能を解析するための手法に関して説明する。さらに、嚥下音が擁する特徴について述べる。

2.1 嚥下機能

嚥下は舌 (Tongue) や咽頭 (Pharynx)、喉頭 (Larynx) 等の器官が複雑に連携して運動することで行うことができる [19]。図 2.1 に示すように一度の嚥下は以下の三つの過程に分けられる [20]。

嚥下第 1 期 食塊 (Bolus) を形成し、食物や飲料を舌によって咽頭へ送り込む過程

嚥下第 2 期 咽頭に入ってから食道 (Esophagus) に入るまでの過程

嚥下第 3 期 食道の蠕動運動によって胃へと送り込まれる過程

嚥下音は、嚥下第 2 期に発生することが知られている。

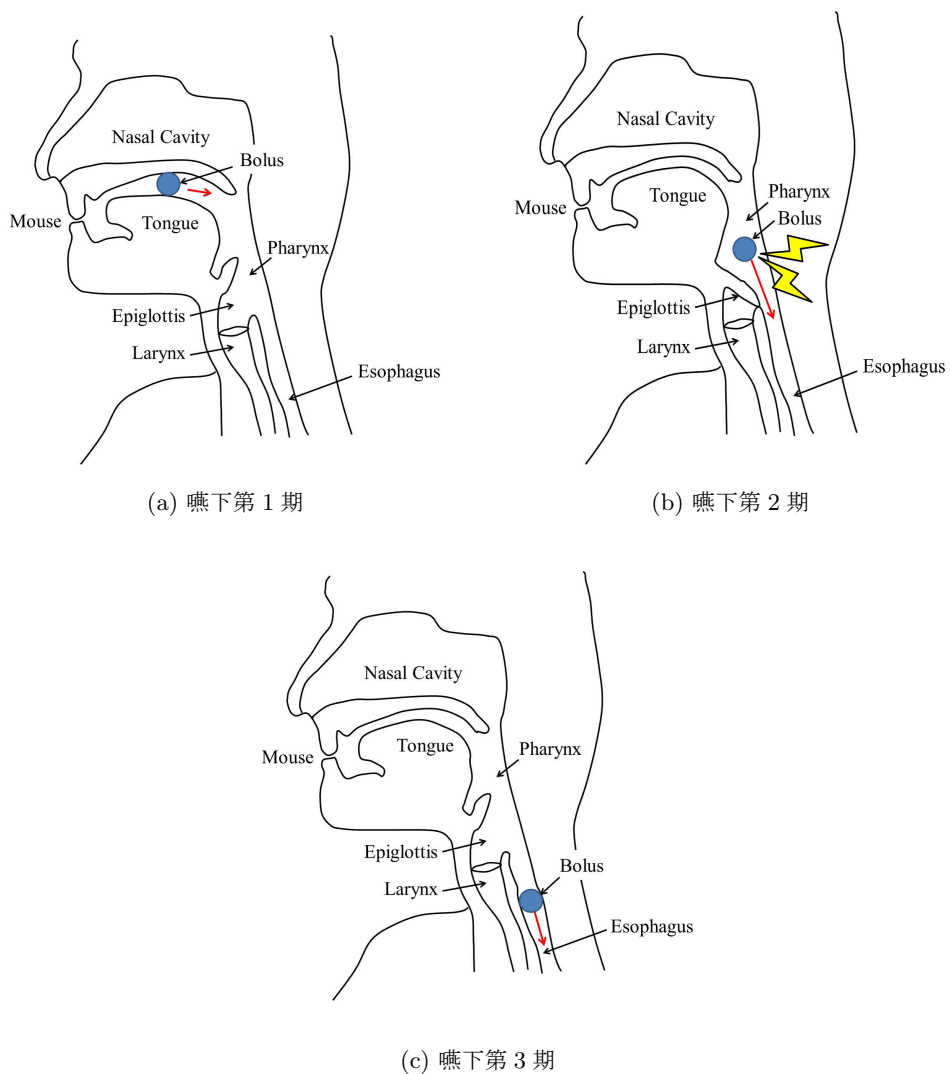


図 2.1: 一嚥下の過程

2.2 嚥下機能の解析手法

嚥下機能の解析手法として、嚥下造影や嚥下時の喉頭運動を用いた手法が挙げられる。以下では、それぞれの手法と特徴について述べる。

2.2.1 嚥下造影による解析

硫酸バリウムなどの造影剤を含む模擬食品を嚥下する様子を、X線透視装置を用いて観察することを嚥下造影と呼ぶ。また、嚥下造影で得られる画像を記録して解析する検査手法は嚥

下造影検査 (Videofluoroscopic Examination of Swallowing: VF 検査) と呼ばれる [20][21]。VF 検査の概要を図 2.2(a) に、食道を通過する食塊の X 線透画像の例を図 2.2(b) に示す [22]。

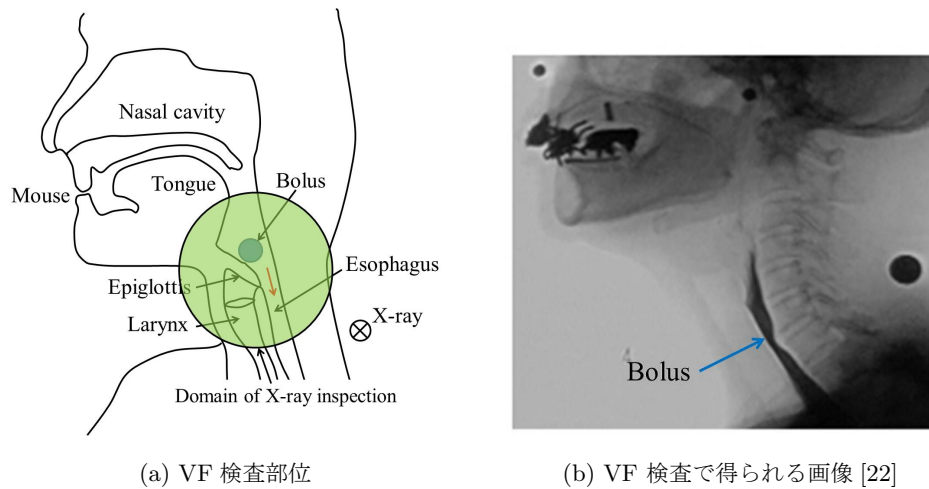


図 2.2: VF 検査による嚥下解析

VF 検査の主な目的としては以下の二点が挙げられる [23]。

- 嚥下時における喉頭、咽頭の形態的異常、機能的異常、誤嚥、残留などを明らかにする
- 食品や体位、摂食方法などを調整することで安全に嚥下し、誤嚥や咽頭残留を減少させる方法を発見する

一般に、VF 検査は嚥下・摂食障害が疑われる患者に行われる検査であるが、嚥下の機構の解明のためにも用いられている。VF 検査は、摂食・嚥下動作の全過程を画像として捕らえることができるだけでなく、咽頭・喉頭部や食道、頸椎の解剖学的な異常も観察することができる信頼性の高い検査である。しかし、VF 検査法は専用の設備内でしか行えず、さらに、X 線被曝の問題もある。従って、嚥下機能の回復、悪化に応じて検査を頻繁に行うことや、摂食中等に長時間の評価を行うことは極めて困難である [24]。

2.2.2 超音波画像による解析

図 2.3(a) に示すように、頸部に超音波診断装置のプローブを接触させ、咽頭での食塊の通過を画像として捉えることで嚥下機能を評価することが可能である [25]。図 2.3(b) にカラー Doppler 法で撮影された超音波画像の例を示す [26]。ただし、TR は気管、CA は頸動脈を表す。また、食塊の種類と咽頭での流速を比較することで、高齢者や嚥下障害者に誤嚥の危険性の少ない食物の決定等にも用いられている [27]。しかし、超音波画像はプローブを接触させる位置や力によって得られる画像の質に大きく差があり、日常生活において簡便に用いることは困難である。

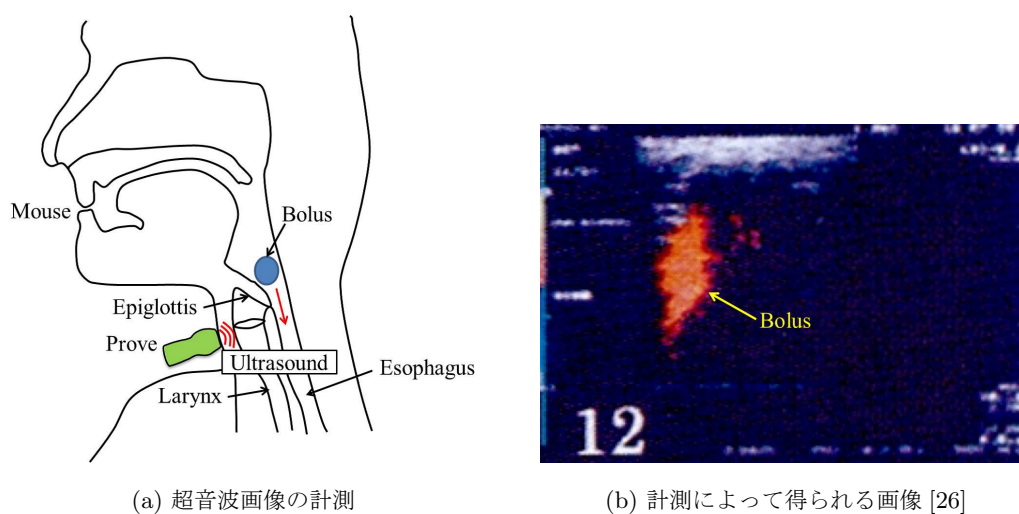


図 2.3: 超音波画像による嚥下解析

2.2.3 インピーダンス咽頭図による解析

図 2.4(a) に示すように頸部に電極を貼付し、嚥下活動中の頸部電気インピーダンスを測定する手法をインピーダンス咽頭図 (Impedance Pharyngography: IPG) と呼ぶ [28]。図 2.4(b) に IPG によって得られる電気インピーダンスの例を示す。IPG によって食塊の咽頭通過時間等を計測し、嚥下動態の解析を行うことができる。IPG は非侵襲的に計測することが可能であるが、電極の貼付位置が変化することによって、波形が変化したり、振幅が減少してしまう。したがって、安定的に計測を行うためには、熟練者の存在が必要不可欠であるという欠点が挙げられる [29][30]。

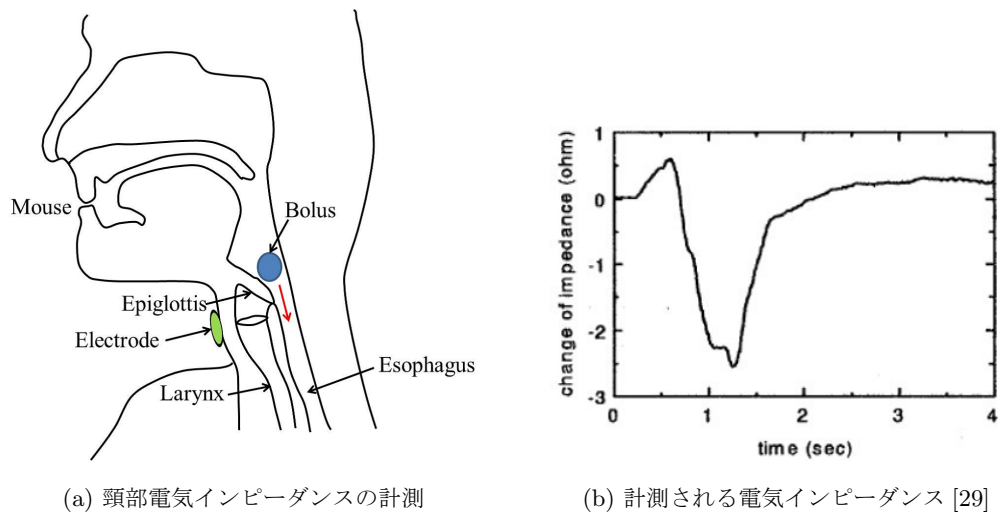


図 2.4: インピーダンス咽頭図による嚥下解析

2.2.4 嚥下時の喉頭運動に伴う振動波形による解析

嚥下を行う際に喉頭の上運動に伴って発生する加速度を嚥下加速度と呼ぶ。図 2.5(a) に示すように喉頭付近の体表面に装着した加速度センサによって、嚥下の際に発生する喉頭の上運動の加速度を計測することで、嚥下動態の解析を行うことができる。図 2.5(b) は計測によって得られる嚥下加速度の例である [31]。嚥下加速度は生体の外側から非侵襲的に計測できるという利点がある [32]。しかし、咳や発声、頭部の動き等による影響を受け、嚥下加速度が正確に計測できない場合がある [33]。さらに、喉頭の運動を計測するため、喉頭に加速度センサを装着しなければ計測することができない。

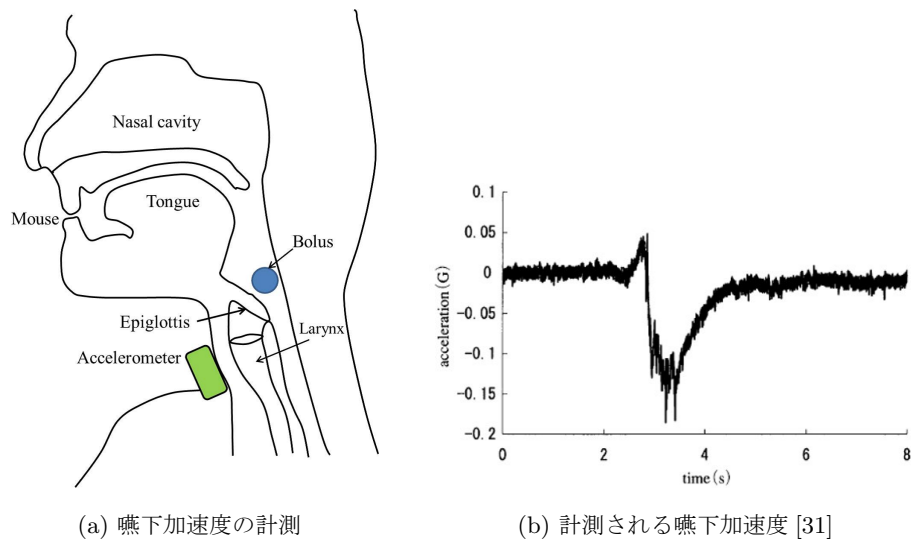


図 2.5: 嚥下加速度による解析

また、嚥下を行う際に、喉頭の上下運動に伴って発生する音を嚥下音と呼び、図 2.6(a) に示すように喉頭付近の体表面に装着したマイクロフォンによって計測し、嚥下動態の解析を行う手法も挙げられる [34]。図 2.6(b) は計測によって得られる嚥下音信号の例である。嚥下音も嚥下加速度と同様に、生体の外側から非侵襲的に計測を行うことが可能である [35]。また、嚥下音は喉頭付近から離れた部分でも計測することが可能である。したがって、加速度センサを用いる手法よりも、計測位置の拘束条件に捕われずに計測ができるという利点が挙げられる。

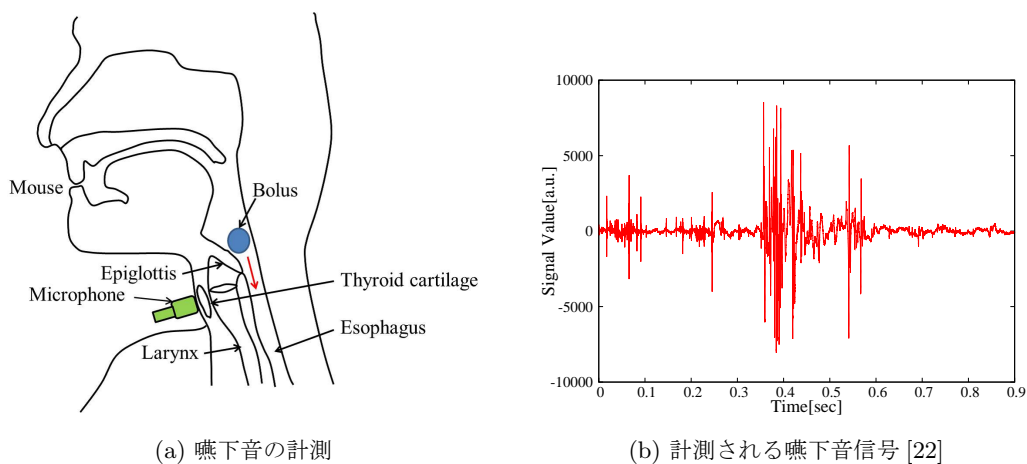


図 2.6: 嚥下音による解析

日常生活における摂取水分量を推定するためには、長時間の計測が必要であり、被験者に対して負担の少ない手法が望ましい。したがって、本研究では被験者への負担が小さく、長時間の計測が可能な嚥下音を用いる。

2.3 嚥下音の特徴

中山らはVF検査と嚥下音の計測を同期させて行うことで、嚥下音が図2.7に示すような嚥下第2期に発生し、以下の三つの要素で構成されることを示した [37]。三つの要素をそれぞれI音、II音、III音と呼ぶ。

I音 喉頭挙上と咽頭への食塊流入によって発生する音

II音 下咽頭への食塊流入と食道入口部の開大によって発生する音

III音 嚥下終了後の喉頭の下降によって発生する音

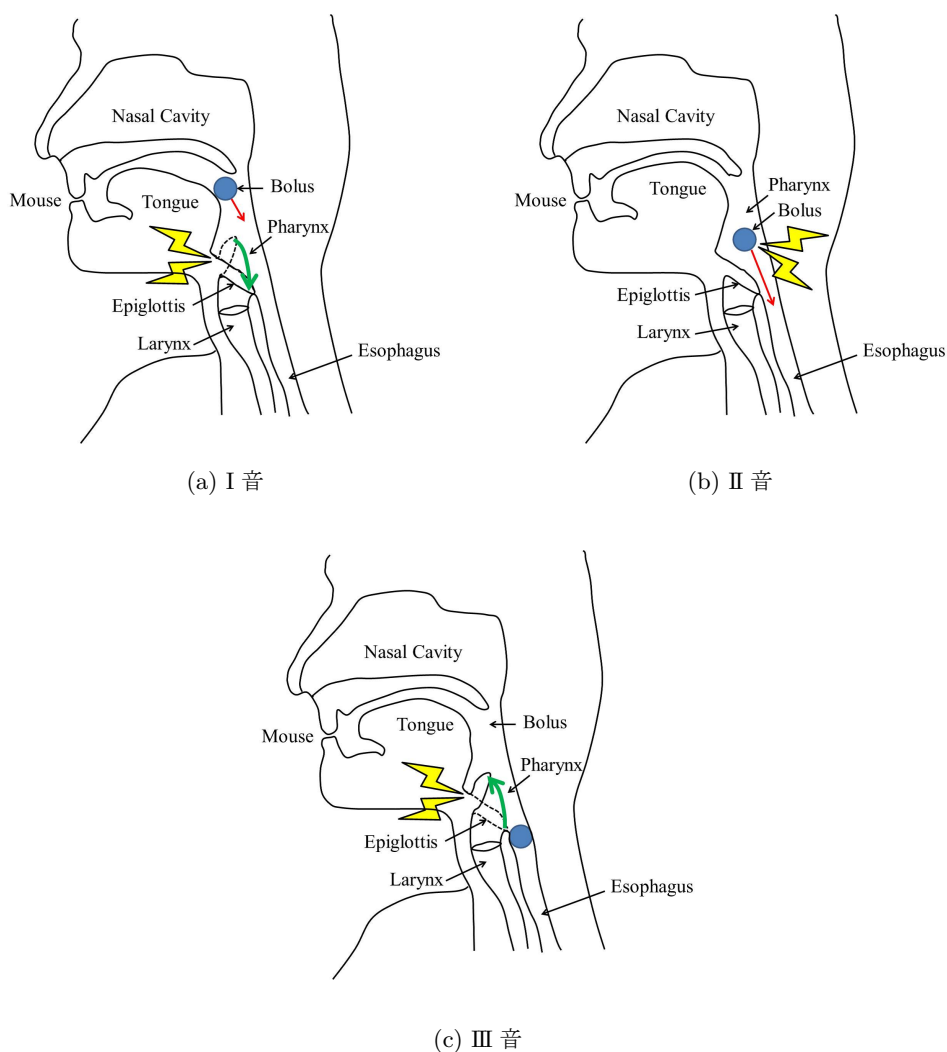


図 2.7: 嚥下音を構成する要素

嚙下音信号の例を図 2.8 に示す。嚙下の際に発生する「ごくっ」という音が II 音に相当し、II 音の前後に発生する「ずっ」という音が I 音と III 音に相当する。また、嚙下音の主要な音響周波数成分は主に 3000 Hz 以下に存在することが知られている [38][39]。

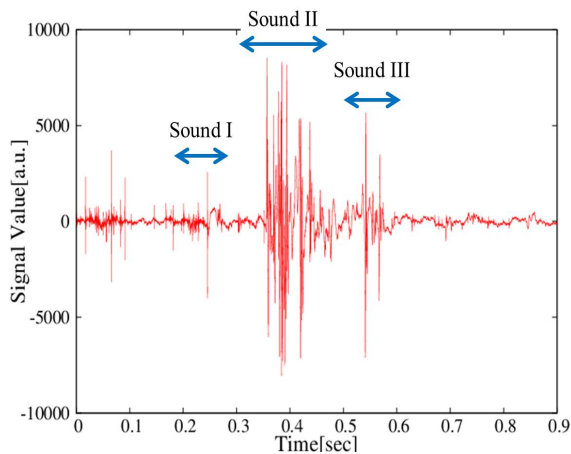


図 2.8: 嚙下音信号

本研究では、摂取水分量を反映し得る特徴量を抽出するために、嚙下音信号に対して音声信号処理を施す。処理結果より得られた特徴量を用いて、摂取した水分量を推定することを目指す。

第3章 摂取水分量の推定システム

摂取水分量の推定は、嚥下音が保持する、一度の嚥下での摂取水分量を反映する特徴量を用いて行う。本章では、嚥下音を用いて摂取水分量を推定するためのシステムの構成について述べ、音声信号処理やパターン認識等の具体的な処理方法について述べる。

3.1 システムの概要

本システムにおける処理の流れを図 3.1 に示す。

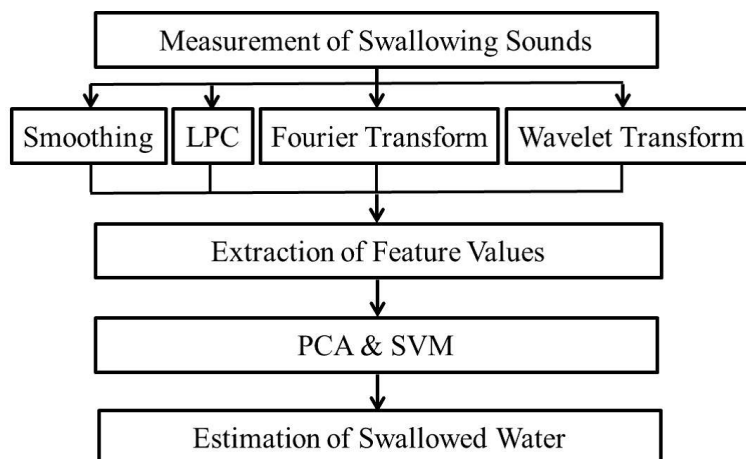


図 3.1: 処理の流れ

嚥下音は図 3.2 に示すように、甲状軟骨 (Thyroid Cartilage) の表面にマイクロフォンを接触させることで計測し、A/D 変換することで PC に取り込む。計測した嚥下音に対して、平滑化 (Smoothing), 線形予測分析 (Linear Prediction Coding: LPC), フーリエ変換 (Fourier Transform), ウェーブレット変換 (Wavelet Transform) を施し、得られた結果から一度の嚥下での摂取水分量を反映する特徴量を抽出する。次に、主成分分析 (Principal Component Analysis: PCA) によって特徴量の低次元化を行う。さらに、摂取した水分量の真値が既知である嚥下音信号の特徴量を学習用データとして使い、サポートベクタマシン (Support Vector Machine: SVM) によって摂取水分量を推定するための識別器を決定する。最後に、摂取し

た水分量が未知の嚥下音を、決定した識別器によって分類していくことで摂取した総水分量の推定を行う。

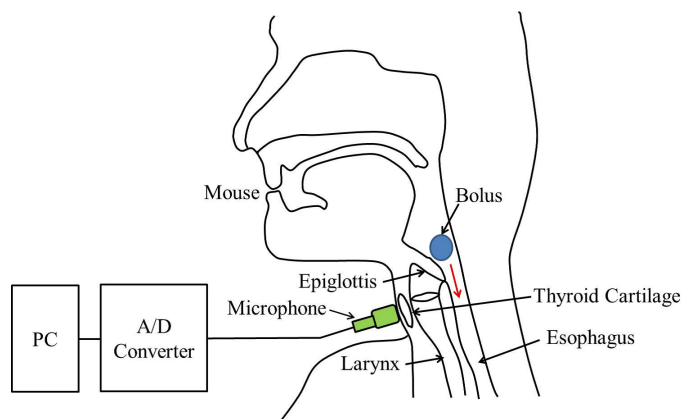


図 3.2: 嚥下音の計測位置

3.2 音声信号の計測と処理

嚥下音を用いて摂取水分量を推定するためには、嚥下音信号から摂取水分量を反映する特徴量を決定する必要がある。本節では、音声信号の計測について説明し、摂取水分量を反映する嚥下音の特徴量を抽出するための音声信号の処理方法と抽出する特徴量について述べる。

3.2.1 音声信号の計測

音声は連続的な信号である。しかし、音声信号の処理を行うコンピュータは、離散的な信号のみしか扱うことができない。したがって、コンピュータでの処理を実現するために連続音声信号 $x(t)$ を離散信号 x_i に変換する必要がある。

まず、一定の時間ごとに連続信号の値を取得する標本化を行う。標本化を行う時間間隔を標本化周期と呼ぶ。また、標本化周波数は標本化周期の逆数である。図 3.3 は標本化周期が 0.2 sec、すなわち、標本化周波数が 5 Hz の例である。標本化周期が短い、すなわち、標本化周波数が高いほど時間分解能が高い。また、標本化周波数が、入力信号の周波数の 2 倍以上でなければ、波形を正確に復元できないという性質があり、標本化定理と呼ばれる。

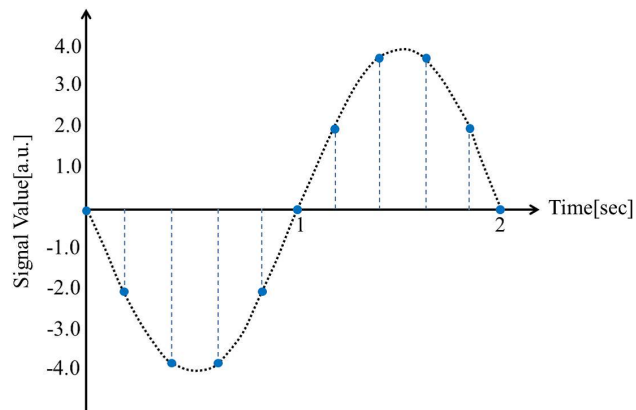


図 3.3: 信号の標本化

次に、離散信号の値を複数の段階に分割し、それぞれに2進数のデジタル値を配分する量子化を行う。配分する2進数の1桁を1ビット、データ値を表現するビットの数を量子化ビット数と呼ぶ。図3.4は量子化ビット数を3とし、データ値を8分割して離散化を行った例である。量子化ビット数を多くすることで、詳細なデータ値の変化を表現することができる。離散信号に変換した後、コンピュータに入力した音声信号に対して、信号処理を施し解析を行っていく。

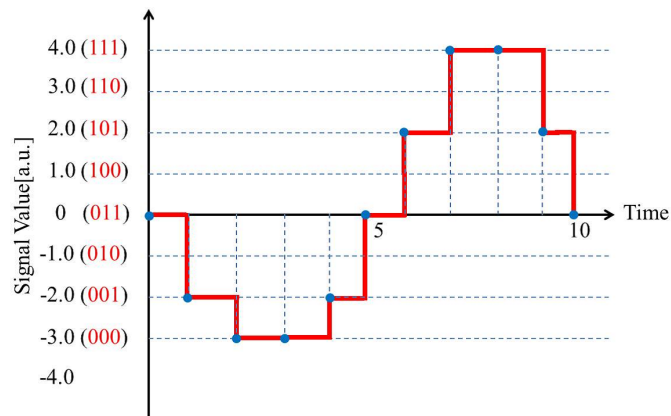


図 3.4: 信号の量子化

3.2.2 音声信号の処理

本項では、計測した音声信号に対して、特徴量を抽出するために施す平滑化、線形予測分析、フーリエ変換、ウェーブレット変換について述べる。

信号の整流

図 3.5(a) は音声信号の例である。音声信号が N 個のデータ $\{x_0, x_1, \dots, x_{N-2}, x_{N-1}\}$ で構成されるとすると、信号を整流するために式 (3.1) のように x_i の絶対値 y_i を取得し、図 3.5(b) のような信号を得る。ただし、 $i = 0, 1, \dots, N-2, N-1$ である。

$$y_i = |x_i| \quad (3.1)$$

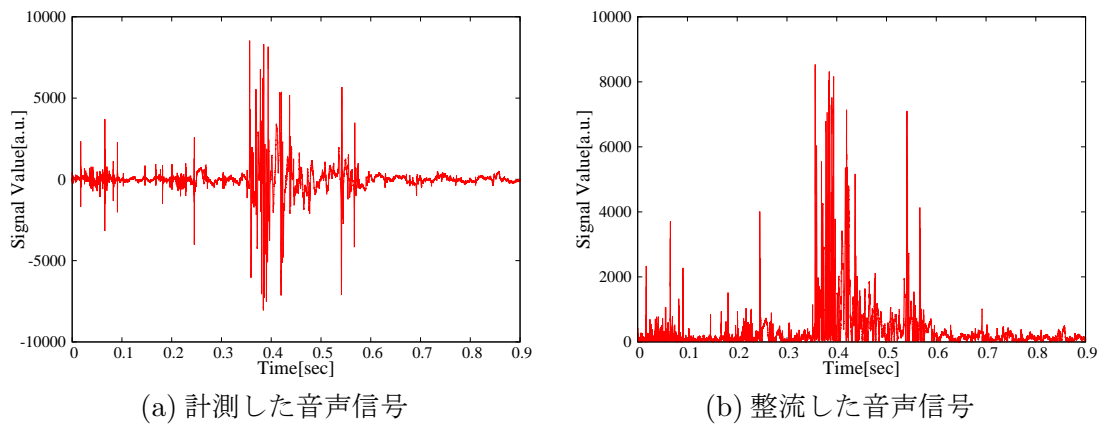


図 3.5: 音声信号

単純移動平均による平滑化処理

平滑化処理とは、信号の中で近傍の値から乖離している値を、除去する、他の値に近づけるといった処理を行うことで、高周波成分を除去して信号の概形を滑らかな状態にすることである。単純移動平均は、 N 個の信号点 $\{y_0, y_1, \dots, y_{N-2}, y_{N-1}\}$ に対して各データの直近の $l+1$ 個のデータの単純な平均を新たなデータとして用いることで平滑化を行う。単純移動平均を施した後の信号の値 z_i は式 (3.2) で表される。ただし、 l は偶数である。オフラインで処理を行う場合、対象とする時間の前後のデータを用いることで平滑化処理前と位相変化のない結果を得ることができる。

$$z_i = \frac{1}{l+1} \sum_{i'=-\frac{l}{2}}^{\frac{l}{2}} y_{i+i'} \quad (3.2)$$

図 3.6(a) に示される信号に対して単純移動平均によって平滑化処理を行うことで、図 3.6(b) に示すように雑音の影響を受けた信号を滑らかにすることが可能である。信号を滑らかにすることで、I 音、II 音、III 音の関係を把握することができる。

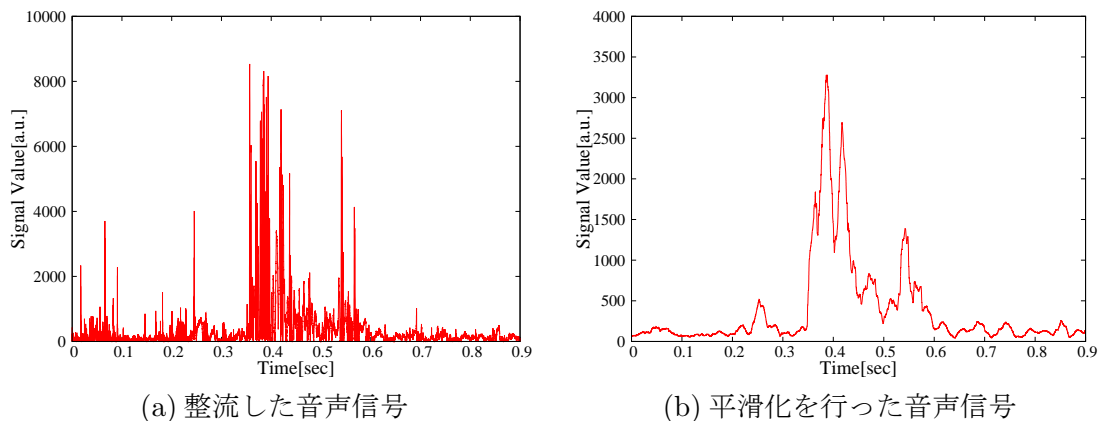
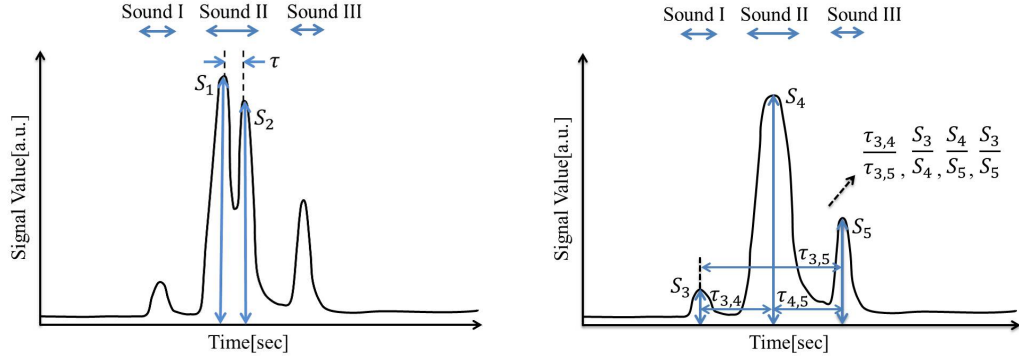


図 3.6: 音声信号の平滑化

定数 l を l_1 として平滑化を施し、高周波成分を除去して信号を滑らかにすることで、図 3.7(a) に示すように嚥下を行った際の「ごくっ」という音声に相当する二つのピークを II 音において検出できる。嚥下量に応じて二つのピークの振幅や間隔に変化が起こり得るため、二つのピーク値の大きさ S_1 、 S_2 と二つのピークの間隔 τ を特徴量として抽出する。また、 l の値を l_1 よりもさらに大きい l_2 として平滑化を行うことで、II 音の二つのピークを滑らかにし、図 3.7(b) に示すように I 音、II 音、III 音のピークを同時に検出することができる。嚥下量に応じて、より強い力で喉頭を上下させる、嚥下時間が長くなる等の影響によって、三つの音のピーク値が大きくなる、時間間隔が長くなる等の変化が起こり得る。したがって、以下に示すような 10 個の特徴量を抽出する。

- I 音、II 音、III 音の振幅値 (S_3, S_4, S_5)
- I 音と II 音、II 音と III 音、I 音と III 音の振幅値の比 ($S_{3,4}, S_{3,5}, S_{4,5}$)
- I 音と II 音、II 音と III 音、I 音と III 音の間隔 ($\tau_{3,4}, \tau_{4,5}, \tau_{3,5}$)
- I 音と III 音の間隔に対する I 音と II 音の間隔の割合 ($\tau_{3,4}/\tau_{3,5}$)



(a) $l = l_1$ として平滑化を行った場合

(b) $l = l_2$ として平滑化を行った場合

図 3.7: 平滑化結果からの特徴量の抽出

線形予測分析

嚙下音信号が N 個の信号点から構成されるとする。線形予測分析では、 n 番目の信号点の値 x_n を式 (3.3) のように直前の μ 個の信号点の値 $x_{n-1}, x_{n-2}, \dots, x_{n-\mu-1}, x_{n-\mu}$ の線形和で予測することを考える [40]。ただし、 e_n は予測誤差、 $d_1, d_2, \dots, d_{\mu-1}, d_\mu$ は線形予測係数である。

$$x_n = -(d_1 x_{n-1} + d_2 x_{n-2} + \dots + d_{\mu-1} x_{n-\mu-1} + d_\mu x_{n-\mu}) + e_n \quad (3.3)$$

$d_0 = 1$ とすると、予測誤差 e_n は式 (3.4)、予測誤差の平均値 \hat{e}_n は式 (3.5) で表される。

$$e_n = \sum_{i=0}^{\mu} d_i x_{n-i} \quad (3.4)$$

$$\hat{e}_n = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{i=0}^{\mu} d_i x_{n-i} \quad (3.5)$$

予測誤差の二乗平均 \hat{e}_n^2 が最小となるように線形予測係数を決定するためには、 \hat{e}_n^2 を d_i ($i = 0, 1, \dots, \mu - 1, \mu$) で一階偏微分した結果が 0、二階偏微分した結果が正となる必要がある。予測誤差の二乗平均 \hat{e}_n^2 を d_i で偏微分すると式 (3.6) のように表すことができる。

$$\begin{aligned} \frac{\partial \hat{e}_n^2}{\partial d_i} &= 2 \hat{e}_n \frac{\partial \hat{e}_n}{\partial d_i} \\ &= 2 \left(\frac{1}{N} \sum_{n=0}^{N-1} \sum_{i'=0}^{\mu} d_{i'} x_{n-i'} \right) x_{n-i} \end{aligned} \quad (3.6)$$

必要条件として偏微分した値は 0 であるため、式 (3.6) は式 (3.7) のように表される。

$$\frac{1}{N} \sum_{n=0}^{N-1} \sum_{i'=0}^{\mu} d_{i'} x_{n-i'} x_{n-i} = 0 \quad (3.7)$$

式 (3.7) は式 (3.8) で定義される自己相関関数 $r_{i'}$ を用いて式 (3.10) のように表される。

$$r_{i'} = \frac{1}{N} \sum_{n=0}^{N-1} x_n x_{n+i'} \quad (3.8)$$

$$\sum_{i'=0}^{\mu} d_{i'} r_{|i-i'|} = 0 \quad (3.9)$$

$$\sum_{i'=1}^{\mu} d_{i'} r_{|i-i'|} = -r_i \quad (3.10)$$

式 (3.10) で表される μ 個の連立方程式は、Levinson-Durbin 法によって解を導出することが可能である [41]。

嚙下量に応じて咽頭を通過する水分量が増えるため、嚙下音が発生する際のモデルに相違が見られる可能性が考えられる。したがって、嚙下音信号全体に対して線形予測分析を行い、結果として得られる μ 個の線形予測係数を特徴量として抽出する。さらに、咽頭内を水が通過する際に発生する II 音のみに対しても線形予測分析を行い、同様に μ 個の線形予測係数を特徴量として抽出する。

フーリエ変換

フーリエ変換は信号を周波数の異なる三角関数の足し合わせで表現することで、信号全体の中に各周波数がどの程度含まれているかを表す周波数解析手法である。連続信号 $x(t)$ に対する連続フーリエ変換を $X(\omega)$ とすると、 $X(\omega)$ は式 (3.11) のように表される。ただし、 j は虚数単位であり、 ω は周波数 f を用いて $\omega = 2\pi f$ で表される。

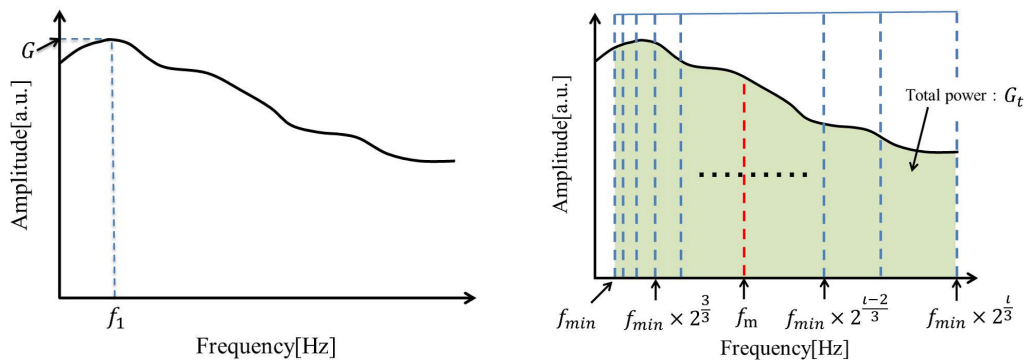
$$X(\omega) = \int_{-\infty}^{\infty} x(t) \exp(-j\omega t) dt \quad (3.11)$$

離散フーリエ変換は音声信号や画像などの離散信号に対して周波数解析を行うための手法である。離散信号のデータ数を N 、信号を x_i ($i = 0, 1, \dots, N-2, N-1$) とすると、離散フーリエ変換 X_k は式 (3.12) のように表される。ただし、 $k = 0, 1, \dots, N-2, N-1$ である。また、離散フーリエ変換では、信号の標本化周波数を f_s とすると、 f_s の半分以下の周波数成分しか表現することができない。したがって、標本化周波数を、抽出する周波数成分の 2 倍以上に設定する必要がある。

$$X_k = \sum_{i=0}^{N-1} x_i \exp\left(-j \frac{2\pi i k}{N}\right) \quad (3.12)$$

嚙下量が変化することで、嚙下音信号の持つ周波数成分に影響がある可能性が考えられる。したがって、図 3.8(a) に示すように最も大きい周波数スペクトルの値 G と対応する周波数 f_1 を特徴量として抽出する。さらに、耳で知覚する嚙下音は常に一定ではないことが確認できるため、嚙下量に応じた変化が生じ得ることも考えられる。聴覚的な変化を捉えるために、フーリエ変換の結果を用いて $\frac{1}{3}$ オクターブ分析を行う。ある周波数とその 2 倍の周波数までの帯域を 1 オクターブと呼び、 $\frac{1}{3}$ オクターブは 1 オクターブをさらに三つの帯域に等比的に分割したものである。 $\frac{1}{3}$ オクターブ分析では、ある周波数帯を $\frac{1}{3}$ オクターブ毎の帯域に分割し、周波数帯全体のスペクトル和に対する各帯域内のスペクトル和の割合を算出する。図 3.8(b) に示すように、周波数帯の下限の周波数を f_{min} として、 l 個の帯域に分割し、緑色で示される周波数帯域全体のスペクトル和 G_t と、 G_t に対する青い点線で区切られた各帯域内のスペクトル和の割合 $h_1 \sim h_l$ を特徴量として抽出する。したがって、 $h_1 \sim h_l$ は G_t を 1 として正規化されることとなる。さらに、周波数帯域全体における重心周波数 f_m も特徴量として抽出する。重心周波数は、周波数 f におけるスペクトル値を G 、周波数帯の上限と下限の周波数をそれぞれ f_{max} 、 f_{min} とすると式 (3.13) のように表される。

$$f_m = \frac{\sum_{f=f_{min}}^{f_{max}} fG}{\sum_{f=f_{min}}^{f_{max}} f} \quad (3.13)$$



(a) 周波数スペクトルからの特徴量抽出

(b) $\frac{1}{3}$ オクターブ分析による特徴量の抽出

図 3.8: フーリエ変換結果からの特徴量の抽出

ウェーブレット変換による時間周波数解析

周波数解析では、信号全体に含まれる周波数成分のみしか表現することができない。一方で時間周波数解析では、信号に含まれる各周波数成分が時間と共にどのように変化していくかを解析する。代表的な手法として、短時間フーリエ変換とウェーブレット変換が挙げられる。短時間フーリエ変換では、信号に対して窓関数を掛けた後にフーリエ変換を施すことで、特定の時間における信号に対して周波数解析を行う。周波数を解析するためには1周期以上の信号を観察する必要がある。短時間フーリエを用いた時間周波数解析では、図 3.9 に示すように解析する時間の間隔が一定に固定され、高い周波数の解析の際に必要な以上の時間幅を持つ信号を解析してしまうため、非効率的である。

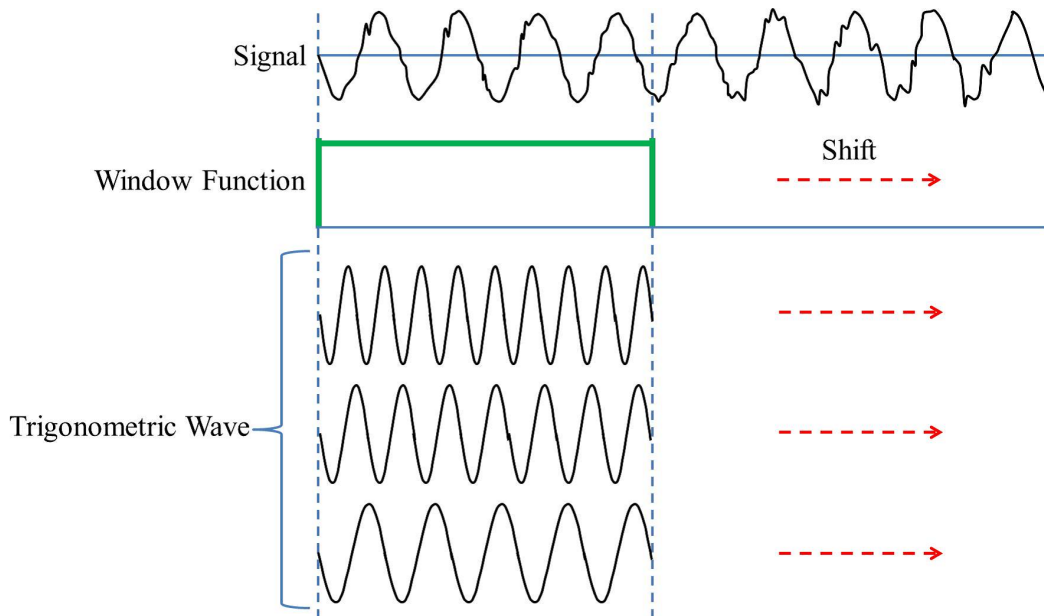


図 3.9: フーリエ変換による時間周波数解析

フーリエ変換に対して、ウェーブレット変換では図 3.10 に示すようなマザーウェーブレットと呼ばれる小波を用い、マザーウェーブレットの時間方向の縮尺を変化させることにより、様々な周波数を持つウェーブレット関数を作り出す。マザーウェーブレットは複素関数で表され、実部と虚部を持つ。さらに、周波数に応じて時間間隔も変化させる。図 3.11 に示すように、それぞれのウェーブレット関数を信号の時間方向に移動させながら、信号とウェーブレット関数の畳み込み積分を算出することで時間周波数解析を行っていく [42]。ウェーブレット変換では、高い周波数では周波数分解能は低く、時間分解能は高い。一方で、低い周波数では周波数分解能が高く、時間分解能が低い。

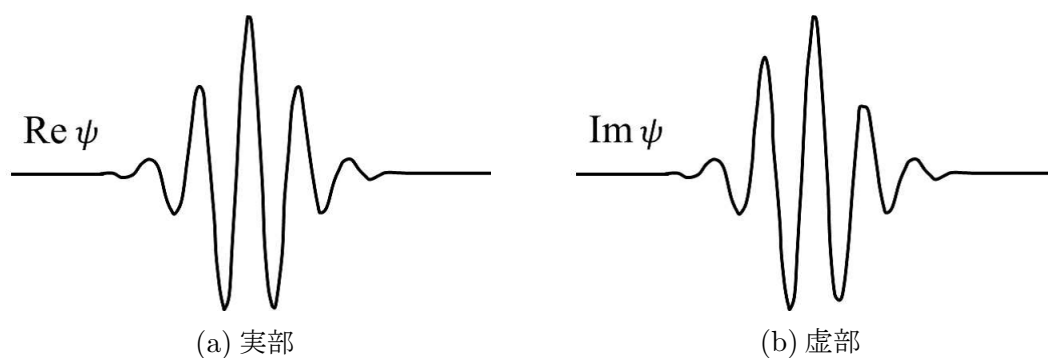


図 3.10: マザーウェーブレット関数

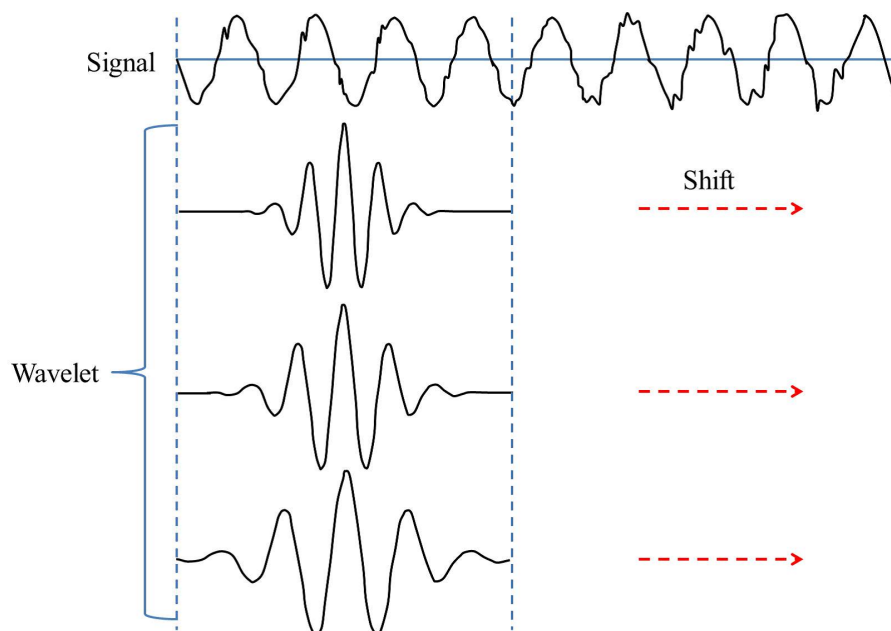


図 3.11: ウェーブレット変換による時間周波数解析

変換される信号を $x(t)$ 、マザーウェーブレット関数を $\psi(t)$ とする。 $\psi(t)$ を複素関数とすると、ウェーブレット変換 $W(b, a)$ は式 (3.14) のように、 $x(t)$ と $\psi(t)$ の複素共役 $\overline{\psi(t)}$ の積を時間方向に積分することで表現される。ただし、 a, b は定数であり、ウェーブレット関数 $\psi(t)$ において a が周波数成分に相当し、 b が時間方向への移動量に相当する。

$$W(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (3.14)$$

コンピュータで処理するのは離散信号であるため、式 (3.14) を離散化する必要がある。データが N 個の音声信号において、ウェーブレット変換 $W(b, a)$ は式 (3.15) で表される。

$$W(b, a) = \frac{1}{\sqrt{a}} \sum_{i=0}^{N-1} x_i \overline{\psi\left(\frac{i-b}{a}\right)} \quad (3.15)$$

また、時間周波数平面上に各成分の分布を表示させるために、スカログラム $P(b, a)$ を用いる。 $P(b, a)$ は式 (3.16) のように定義される。

$$P(b, a) = |W(b, a)|^2 \quad (3.16)$$

摂取する水分量が増えることは、咽頭内を水が通過する際に発生する音声である II 音における周波数成分に対して特に影響を与えることが考えられる。したがって、図 3.12 に示すように、II 音におけるスカログラムの最大値 P_{max} と最大値を示す時の周波数成分 a_{max} を特徴量として抽出する。

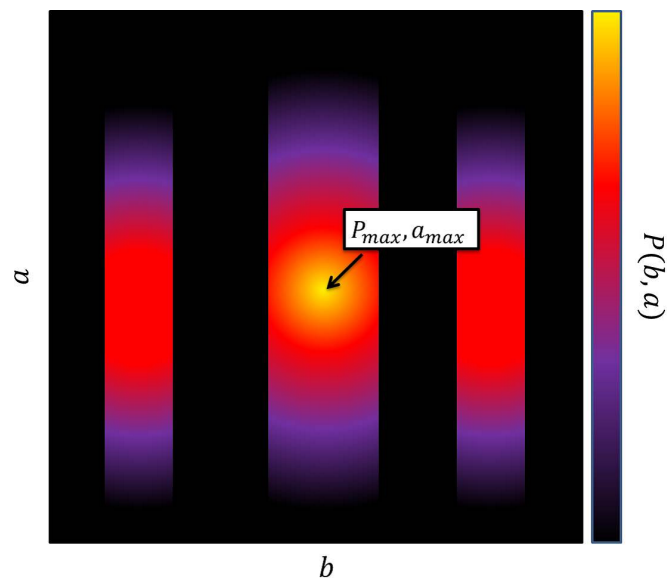


図 3.12: ウェーブレット変換結果からの特徴量の抽出

3.3 音声特徴量の解析

摂取水分量を推定するためには、音声信号から抽出した特徴量を用いて識別器を構成する必要がある。本節では、特徴量を低次元化することができる主成分分析と、特徴量を用いて識別器を構成するサポートベクタマシン (Support Vector Machine : SVM) について述べる。

3.3.1 主成分分析

主成分分析は、 p 個の変数からなるデータ $x_0, x_1, \dots, x_{p-2}, x_{p-1}$ が持つ情報を、式 (3.17) のように互いに独立な M 個の主成分 z_m ($m = 0, 1, \dots, M-2, M-1$) を用いた、より少数の変数で近似的に表現する手法である [43]。ただし、 $v_{m0}, v_{m1}, \dots, v_{mp-2}, v_{mp-1}$ は定数であり、 $v_{m0}^2 + v_{m1}^2 + \dots + v_{mp-2}^2 + v_{mp-1}^2 = 1$ を満たす。

$$z_m = v_{m0}x_0 + v_{m1}x_1 + \dots + v_{mp-2}x_{p-2} + v_{mp-1}x_{p-1} \quad (3.17)$$

少数の変数で表現する際に、元データの情報の損失を最小限に抑える必要がある。主成分 z_m の分散を最大にすることで、元のデータの情報量の損失を最小にすることが可能である。したがって、 z_m の分散を σ_m^2 とすると、式 (3.18) が成り立つように主成分 z_m を並び替える。 z_0 を第 1 主成分と呼び、以下同様に z_1 を第 2 主成分、 \dots 、 z_{M-1} を第 M 主成分と呼ぶ。

$$\sigma_0^2 \geq \sigma_1^2 \geq \dots \geq \sigma_{M-1}^2 \quad (3.18)$$

N 個の p 次元データ $\mathbf{x}_i = (x_{i0} \ x_{i1} \ \dots \ x_{ip-2} \ x_{ip-1})$ ($i = 0, \dots, N-1$) を考える。 p 個の変数の単位が異なる場合、 x_{iq} の平均値 \hat{x}_{iq} 、標準偏差 ϕ_{iq} ($0 \leq q \leq p-1$) を用いて式 (3.19) のように正規化を行う。

$$x'_{iq} = \frac{x_{iq} - \hat{x}_{iq}}{\phi_{iq}} \quad (3.19)$$

第 $m+1$ 主成分の係数ベクトルを $\mathbf{v}_m = (v_{m0} \ v_{m1} \ \dots \ v_{mp-2} \ v_{mp-1})^T$ とする。 T は転置を表す。式 (3.20) で表される \mathbf{x}'_i と \mathbf{v}_m の積 Υ_{im} を第 $m+1$ 主成分における i 番目のデータの主成分得点と呼ぶ。

$$\Upsilon_{im} = \mathbf{x}'_i \mathbf{v}_m \quad (3.20)$$

式 (3.21) に表すようにデータ \mathbf{x}'_i 全体を表す行列を \mathbf{X} とすると、 \mathbf{x}'_i の共分散行列 \mathbf{V} は式

(3.22) に示すように表される。

$$\mathbf{X} = \begin{bmatrix} x'_{00} & x'_{01} & \cdots & x'_{0N-1} \\ x'_{10} & x'_{11} & \cdots & x'_{1N-1} \\ \vdots & \vdots & \ddots & \vdots \\ x'_{N-10} & x'_{N-11} & \cdots & x'_{N-1N-1} \end{bmatrix} \quad (3.21)$$

$$\mathbf{V} = \frac{1}{N-1} \mathbf{X}^T \mathbf{X} \quad (3.22)$$

共分散行列の固有値を大きい順に並び替え、値の大きい M 個の固有値を $\lambda_0, \lambda_1, \dots, \lambda_{M-2}, \lambda_{M-1}$ ($\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{M-2} \geq \lambda_{M-1}$) とすると、式 (3.23) に示すように第 $m+1$ 主成分の分散は λ_m と等しくなる [44]。さらに、 λ_0 に対する固有ベクトルが \mathbf{v}_0 、 λ_1 に対する固有ベクトルが \mathbf{v}_1 、 \dots 、 λ_{M-2} に対する固有ベクトルが \mathbf{v}_{M-2} 、 λ_{M-1} に対する固有ベクトルが \mathbf{v}_{M-1} とそれぞれ一致する。

$$\sigma_m^2 = \lambda_m \quad (m = 0, 1, \dots, M-2, M-1) \quad (3.23)$$

また、第 $m+1$ 主成分にどの程度 of 元データ情報量を含んでいるかを示す指標として、式 (3.24) で表される β_m を第 $m+1$ 主成分の寄与率と呼ぶ。

$$\beta_m = \frac{\lambda_m}{\sum_{k=0}^{M-1} \lambda_k} \quad (m = 0, 1, \dots, M-2, M-1) \quad (3.24)$$

さらに、式 (3.25) で表される $\beta_{1,m}$ を第 1 主成分から第 $m+1$ 主成分までの累積寄与率と呼ぶ。

$$\beta_{1,m} = \frac{\sum_{k=0}^m \lambda_k}{\sum_{k=0}^{M-1} \lambda_k} \quad (m = 0, 1, \dots, M-2, M-1) \quad (3.25)$$

主成分分析によって多次元データを 2 次元や 3 次元等の低次元データに縮約することができる。

3.3.2 多クラス SVM

本項では、パターン認識の手法の一つであり、教師データを用いて未知データを各クラスに分類する多クラス SVM について述べる。

2 クラスハードマージン SVM

M 個の m 次元教師データ \mathbf{x}_i ($i = 0, 1, \dots, M-2, M-1$) がクラス 0、1 のいずれかに属するとして、クラスラベル y_i をクラス 0 のとき $y_i = 1$ 、クラス 1 のとき $y_i = -1$ とする。各データが線形分離可能とすると、式 (3.26) のような決定関数を決定することができる。ただし、 \mathbf{w} は m 次元係数ベクトル、 χ はバイアス項を表す [45]。

$$D(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + \chi \quad (3.26)$$

線形分離可能であるための条件として、式 (3.27) を満たすとする。

$$y_i(\mathbf{w}^T \mathbf{x}_i + \chi) \geq 1 \quad (3.27)$$

式 (3.26) の表す分離超平面と最も近い教師データとの距離をマージンと呼ぶ。SVM では、マージンが最大となるような最適超平面を決定する。教師データ \mathbf{x}_i と超平面との距離 u_i は式 (3.28) で表される。

$$u_i = \frac{|D(\mathbf{x}_i)|}{\|\mathbf{w}\|} \quad (3.28)$$

したがって、 u_{\max} をマージンとすると全ての教師データは式 (3.29) を満たす必要がある。ただし、 $\|\mathbf{w}\|$ は \mathbf{w} のノルムを表す。

$$\frac{y_i D(\mathbf{x}_i)}{\|\mathbf{w}\|} \geq u_{\max} \quad (3.29)$$

式 (3.27) と式 (3.29) より、式 (3.30) が導かれる。

$$u_{\max} \|\mathbf{w}\| = 1 \quad (3.30)$$

式 (3.30) より、マージン u_{\max} が大きくなるのは $\|\mathbf{w}\|$ が小さくなることであることがわかる。したがって、各クラスが最も大きく分離される決定関数は式 (3.31)、式 (3.32) で表される \mathbf{w} と χ に関する評価関数 $Q(\mathbf{w}, \chi)$ の最小化問題を解くことで算出される。ただし、 \min

は評価関数を最小にするパラメータを抽出することを表す。

$$\min Q(\mathbf{w}, \chi) = \frac{1}{2} \|\mathbf{w}\|^2 \quad (3.31)$$

$$y_i(\mathbf{w}^T \mathbf{x}_i + \chi) \geq 1 \quad (i = 0, 1, \dots, M-2, M-1) \quad (3.32)$$

最小化問題は変数の数が教師データの数と等しい双対問題に変換して解かれる。式 (3.31)、式 (3.32) は α_i を非負のラグランジュ乗数、 $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{M-2}, \alpha_{M-1})^T$ として式 (3.33) のように表される [46]。

$$\min Q(\mathbf{w}, \chi, \boldsymbol{\alpha}) = \frac{1}{2} \mathbf{w} \mathbf{w}^T - \sum_{i=0}^{M-1} \alpha_i (y_i(\mathbf{w}^T \mathbf{x}_i + \chi) - 1) \quad (3.33)$$

式 (3.33) の最適解を与える条件は式 (3.34)~ 式 (3.37) で表され、KKT(Karush Kuhn Tucker) 条件と呼ばれる。

$$\frac{\partial Q(\mathbf{w}, \chi, \boldsymbol{\alpha})}{\partial \mathbf{w}} = 0 \quad (3.34)$$

$$\frac{\partial Q(\mathbf{w}, \chi, \boldsymbol{\alpha})}{\partial \chi} = 0 \quad (3.35)$$

$$\alpha_i \{y_i(\mathbf{w}^T \mathbf{x}_i + \chi) - 1\} = 0 \quad (i = 0, 1, \dots, M-2, M-1) \quad (3.36)$$

$$\alpha_i \geq 0 \quad (i = 0, 1, \dots, M-2, M-1) \quad (3.37)$$

式 (3.36) より、 $\alpha_i = 0$ であるか、または $\alpha_i \neq 0$ かつ $y_i(\mathbf{w}^T \mathbf{x}_i + \chi) = 1$ を満たす必要がある。 $\alpha_i \neq 0$ となるデータ \mathbf{x}_i をサポートベクタと呼ぶ。また、式 (3.34)、式 (3.35) より式 (3.38)、式 (3.39) が成り立つ。

$$\mathbf{w} = \sum_{i=0}^{M-1} \alpha_i y_i \mathbf{x}_i \quad (3.38)$$

$$\sum_{i=0}^{M-1} \alpha_i y_i = 0 \quad (3.39)$$

式 (3.33)、式 (3.38)、式 (3.39) より式 (3.40)、式 (3.41) で表される双対問題が得られる。ただし、max は評価関数を最大にするパラメータを抽出することを表す。

$$\max Q(\boldsymbol{\alpha}) = \sum_{i=0}^{M-1} \alpha_i - \frac{1}{2} \sum_{i,k=0}^{M-1} \alpha_i \alpha_k y_i y_k \mathbf{x}_i^T \mathbf{x}_k \quad (3.40)$$

$$\sum_{i=0}^{M-1} \alpha_i y_i = 0, \quad \alpha_i \geq 0 \quad (i = 0, 1, \dots, M-2, M-1) \quad (3.41)$$

式 (3.42)、式 (3.43) のように定式化された双対問題を 2 クラスハードマージン SVM と呼ぶ。決定関数 $D(\mathbf{x})$ は式 (3.42)、式 (3.43) で表される。データ \mathbf{x} は $D(\mathbf{x})$ が正ならばクラス 0 に分類され、負ならばクラス 1 に分類される。

$$D(\mathbf{x}) = \sum_{i=0}^{M-1} \alpha_i y_i \mathbf{x}_i^T \mathbf{x} + \chi \quad (3.42)$$

$$\chi = y_i - \mathbf{w}_i^T \mathbf{x}_i \quad (3.43)$$

2 クラスソフトマージン SVM

2 クラスハードマージン SVM では線形分離可能であることを仮定している。2 クラスソフトマージン SVM では、線形分離が可能でない場合にも適用できるように拡張を行う。まず、式 (3.27) に非負の変数 ξ_i を導入する [47]。

$$y_i(\mathbf{w}^T \mathbf{x}_i + \chi) \geq 1 - \xi_i \quad (i = 0, 1, \dots, M-2, M-1) \quad (3.44)$$

次に、最小化問題を式 (3.45)、式 (3.46) のように考える。ただし、 C はマージンパラメータと呼ばれる定数である。一般的に、 C の値を大きくすることで教師データの認識率が上がるが、大きくするにしたがって汎化能力は低下する。したがって、教師データ認識率と汎化能力の双方が高くなるように設定する必要がある。

$$\min Q(\mathbf{w}, \chi) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=0}^{M-1} \xi_i \quad (3.45)$$

$$y_i(\mathbf{w}^T \mathbf{x}_i + \chi) \geq 1 - \xi_i \quad (i = 0, 1, \dots, M-2, M-1) \quad (3.46)$$

2 クラスハードマージン SVM の場合と同様にラグランジュ乗数を用いた式変形を行うことで、式 (3.45) の最適解を得るための、式 (3.47)~式 (3.50) で表される KKT 条件が得られる。

$$\mathbf{w} = \sum_{i=0}^{M-1} \alpha_i y_i \mathbf{x}_i \quad (3.47)$$

$$\sum_{i=0}^{M-1} \alpha_i y_i = 0 \quad (3.48)$$

$$\alpha_i \{y_i(\mathbf{w}^T \mathbf{x}_i + \chi) - 1 + \xi_i\} = 0 \quad (i = 0, 1, \dots, M-2, M-1) \quad (3.49)$$

$$(C - \alpha_i) \xi_i = 0 \quad (i = 0, 1, \dots, M-2, M-1) \quad (3.50)$$

さらに、最小化問題を双対問題に変換すると式 (3.51)、式 (3.52) で表される。

$$\max Q(\boldsymbol{\alpha}) = \sum_{i=0}^{M-1} \alpha_i - \frac{1}{2} \sum_{i,k=0}^{M-1} \alpha_i \alpha_k y_i y_k \mathbf{x}_i^T \mathbf{x}_k \quad (3.51)$$

$$\sum_{i=0}^{M-1} \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C \quad (i = 0, 1, \dots, M-2, M-1) \quad (3.52)$$

決定関数は 2 クラスハードマージン SVM と同様に式 (3.42)、式 (3.43) で表される。

カーネルトリック

SVM で用いる教師データと未知データはそれぞれ m 個の変数で構成されるため、各データは m 次元の入力空間上で表現される。2 クラスハードマージン SVM や 2 クラスソフトマージン SVM において、各クラスが複数のデータの集まりで構成されるような場合には、入力空間で分離することができず汎化能力が低くなってしまふ。汎化能力を高めるために、入力空間よりも高次元の特徴空間へと写像し、特徴空間上で最適な超平面を決定することによって線形分離する手法が提案されている。 m 次元の入力空間を ζ 次元の特徴空間へ写像する関数を $\rho(\mathbf{x})$ とすると決定関数 $D(\mathbf{x})$ は式 (3.53) に示すように表される。

$$D(\mathbf{x}) = \mathbf{W}^T \rho(\mathbf{x}) + \chi \quad (3.53)$$

しかし、 $\rho(\mathbf{x})$ を用いる場合、教師データの学習や未知データの識別の際に特徴空間での値を計算しなければならず、膨大な時間が必要となる場合がある。カーネルトリックでは、式 (3.54) に示す関数を導入する。ただし、カーネル関数 $K(\mathbf{x}_i, \mathbf{x}_k)$ は式 (3.55) に示す条件を満たす必要がある。

$$K(\mathbf{x}_i, \mathbf{x}_k) = \rho^T(\mathbf{x}_i) \rho(\mathbf{x}_k) \quad (3.54)$$

$$\sum_{i,k=0}^{M-1} \epsilon_i \epsilon_k K(\mathbf{x}_i, \mathbf{x}_k) \leq 0 \quad (3.55)$$

カーネルを用いた場合、SVM における双対問題は式 (3.56)、式 (3.57) のように表される。

$$\max Q(\boldsymbol{\alpha}) = \sum_{i=0}^{M-1} \alpha_i - \frac{1}{2} \sum_{i,k=0}^{M-1} \alpha_i \alpha_k y_i y_k K(\mathbf{x}_i, \mathbf{x}_k) \quad (3.56)$$

$$\sum_{i=0}^{M-1} \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C \quad (i = 1, 2, \dots, M-2, M-1) \quad (3.57)$$

また、KKT 条件は式 (3.58)、式 (3.59) のように表される。ただし、 $i = 0, 1, \dots, M - 2, M - 1$ である。

$$\alpha_i \left\{ y_i \left(\sum_{k=0}^{M-1} \alpha_k y_k K(\mathbf{x}_i, \mathbf{x}_k) + \chi \right) - 1 + \xi_i \right\} = 0 \quad (3.58)$$

$$(C - \alpha_i) \xi_i = 0 \quad (3.59)$$

カーネルを用いた時の識別のための決定関数は、式 (3.60) のように表すことができる。

$$D(\mathbf{x}) = \sum_{i=0}^{M-1} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + \chi \quad (3.60)$$

式 (3.56)~式 (3.60) で示されるように、 $\rho(\mathbf{x})$ での値を計算することなく入力空間における \mathbf{x}_i と \mathbf{x}_k の内積のみで表すことができるため、特徴空間の変数を用いる必要がない。カーネルトリックによって特徴空間の変数を用いることなく、入力空間の変数のみを用いて計算を行う手法をカーネル法と呼ぶ [48]。

多クラス SVM への拡張

2 クラスから多クラスへ拡張する場合、すべてのクラス対に対して決定関数を算出する。 Θ クラス分類の場合、クラス対 θ_0, θ_1 に対する決定関数は式 (3.61) のように表される。ただし、 $\theta_0, \theta_1 = 0, 1, \dots, \Theta - 1, \theta_0 \neq \theta_1$ である。

$$D_{\theta_0 \theta_1}(\mathbf{x}) = \mathbf{w}_{\theta_0 \theta_1}^T \phi(\mathbf{x}) + b_{\theta_0 \theta_1} \quad (3.61)$$

2 変数 3 クラスの場合の例を図 3.13 に示す。

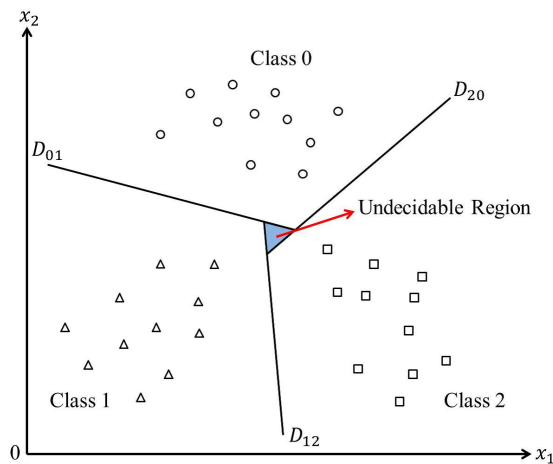


図 3.13: 多クラス SVM の例

図 3.13 に示される三つの決定関数で囲まれた部分では各クラスへ明確に分類することができない。したがって、分離不可能な領域を解消するために、式 (3.62)、式 (3.63) に示すようにクラス対 θ_0 、 θ_1 に対するメンバシップ関数 $\kappa_{\theta_0\theta_1}(\mathbf{x})$ 、クラス θ_0 に対するメンバシップ関数 $\kappa_{\theta_0}(\mathbf{x})$ を導入する。算出された $\kappa_{\theta_0}(\mathbf{x})$ が最大となるクラスへ未知データを分類する。

$$\kappa_{\theta_0\theta_1}(\mathbf{x}) = \begin{cases} 1 & (D_{\theta_0\theta_1}(\mathbf{x}) \geq 1) \\ D_{\theta_0\theta_1}(\mathbf{x}) & (D_{\theta_0\theta_1}(\mathbf{x}) \leq 1) \end{cases} \quad (3.62)$$

$$\kappa_{\theta_0}(\mathbf{x}) = \min_{\substack{\theta_1 \neq \theta_0, \\ \theta_1 = 1, 2, \dots, \Theta}} \kappa_{\theta_0\theta_1}(\mathbf{x}) \quad (3.63)$$

最適なカーネル関数とパラメータの決定

SVM によって学習データから決定関数を算出する際、カーネル関数やマージンパラメータ C 、カーネル関数に含まれる定数パラメータの値に応じて未知データに対する汎用性は大きく変化する。したがって、最も汎用性が高くなるようにカーネル関数とパラメータを設定する必要がある。

最適なカーネル関数とパラメータを決定するための手法として、本研究では一つ抜き交差法を用いる。まず、一つ抜き交差法では学習データからある一つのデータを取り除き、残りのデータを用いて最適超平面を決定する。次に、取り除いたデータを未知データとして扱い、決定された最適超平面を用いてデータが正しいクラスに分類されるかを判定する。以上の判定操作を学習データの回数繰り返し行い、全学習データ数に対する、正しく分類されたデータ数の割合を正答率として算出する。一つ抜き交差法をカーネル関数やパラメータを変化させて適用していき、正答率が最も高くなる場合を最適なカーネル関数とパラメータとして選択する。学習に用いなかったデータで汎用性を評価するため、汎用性が最も高いと予想されるカーネル関数とパラメータを決定することができる [49]。選択したカーネル関数とパラメータを設定し、全学習データを用いて未知データを分類するための決定関数を算出する。

第4章 嚥下音の計測と解析結果

計測した嚥下音信号において、平滑化、線形予測分析、フーリエ変換、ウェーブレット変換結果から特徴量の抽出し、主成分分析と SVM によって摂取水分量推定の評価を行った。本章では、嚥下音の計測方法と、解析結果について述べる。

4.1 嚥下音の計測

一度に嚥下が容易に行える、容量 5 mL、10 mL、15 mL の水を嚥下した際の嚥下音を被験者 3 名 (被験者 A、B、C) に対してそれぞれ 50 回計測した。マイクロフォン (15~20000 Hz において -63~-60 dBV/Pa で收音可能) によって計測が可能な音声信号の周波数成分を十分検出できるように、標本化周波数は 44100 Hz とした。また、嚥下音信号の振幅値を特徴量として抽出するため、振幅のより細かい変化を表現するために量子化ビット数は 16 とした。ただし、今回の計測ではマイクロフォンの位置は厳密には考慮していない。図 4.1 に実験の様子を、図 4.2 に各摂取嚥下量を示す。また、図 4.3 と表 4.1 に使用したマイクロフォンの外観と仕様を示す。

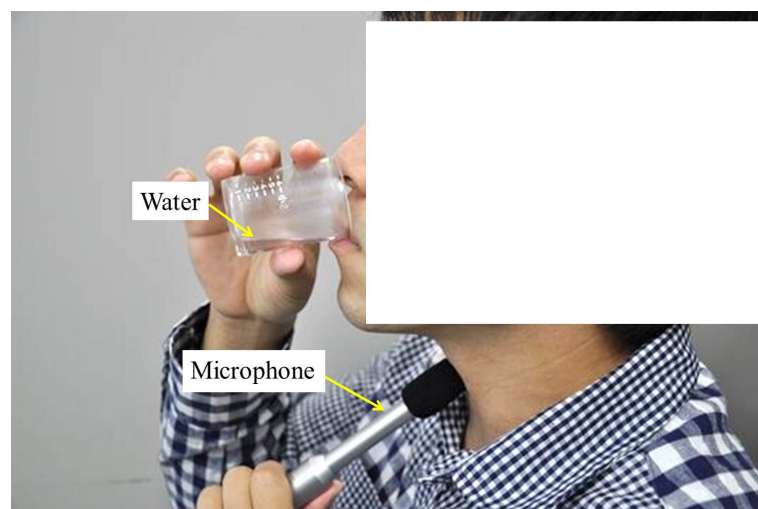


図 4.1: 被験者実験の風景



(a) 5 mL の水

(b) 10 mL の水

(c) 15 mL の水

図 4.2: 摂取する水分量



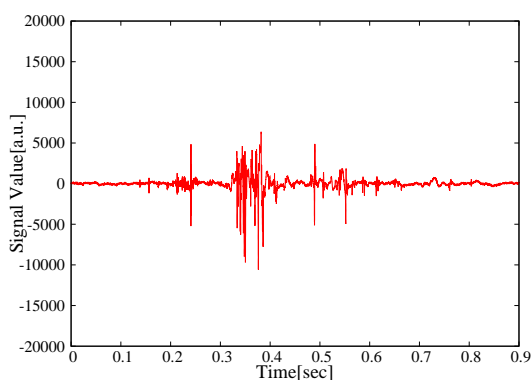
図 4.3: 使用したマイクロフォン

表 4.1: マイクロフォンの仕様

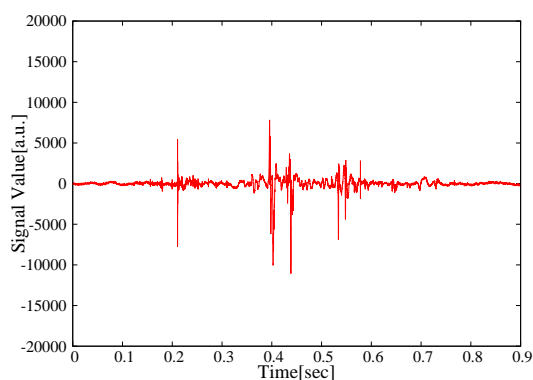
製造会社	型番	計測方式
BEHRINGER	ECM8000	コンデンサ型
感度	周波数特性	指向特性
-60 dBV/Pa(1000 Hz)	-63~-60 dBV/Pa between 15~20000 Hz	無指向性

コンデンサ型マイクロフォンはコンデンサを構成する振動板に電圧を印加し、音声振動によってコンデンサの振動板間隔が変動することで生じる静電容量の変化を、電気信号に変換する方式である。

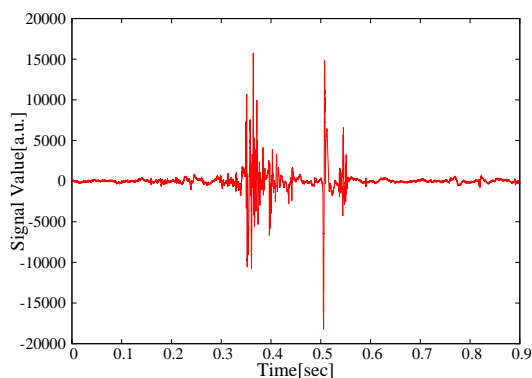
水を 5 mL、10 mL、15 mL 嚥下した際に得られた波形の例を図 4.4 に示す。嚥下音の持続時間は 1 sec 未満であるため、各嚥下音のデータ長は 40000 とした。また、嚥下音信号の始点と終点はそれぞれ手動で決定を行う。



(a) 5 mL の嚥下音信号



(b) 10 mL の嚥下音信号



(c) 15 mL の嚥下音信号

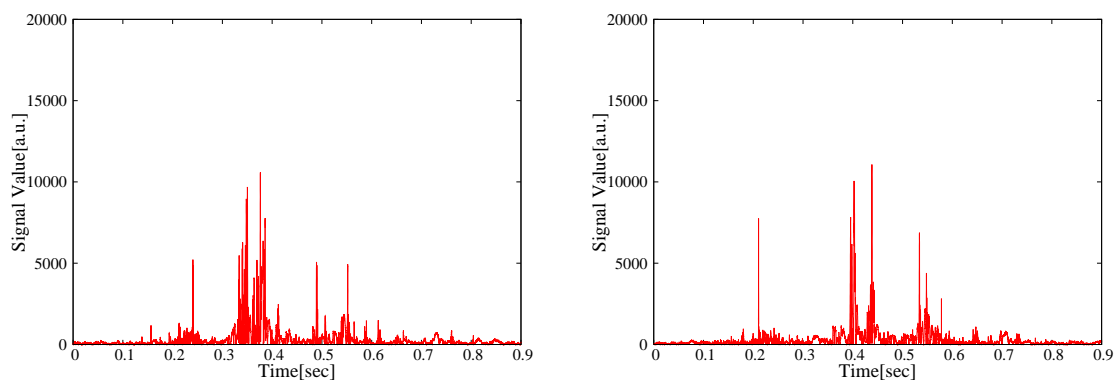
図 4.4: 計測した嚥下音信号の例

4.2 音声信号処理による特徴量の抽出

本節では、平滑化、線形予測分析、フーリエ変換、ウェーブレット変換による、嚙下音信号からの特徴量の抽出について述べる。

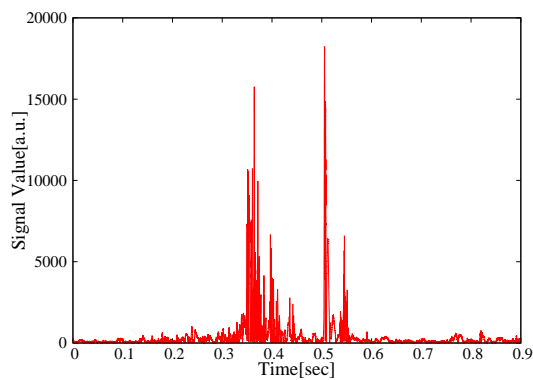
4.2.1 平滑化による特徴量の抽出

まず、計測した嚙下音信号を整流した。図 4.5 に整流した波形の例を示す。



(a) 整流した 5 mL の嚙下音信号

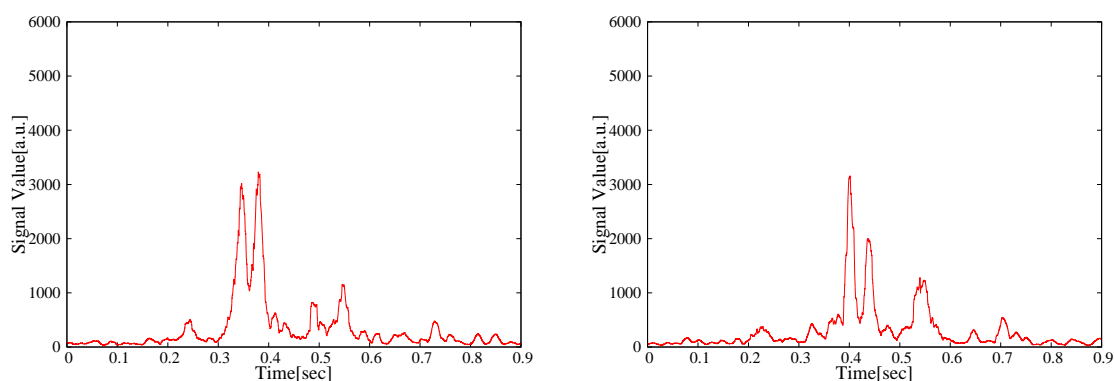
(b) 整流した 10 mL の嚙下音信号



(c) 整流した 15 mL の嚙下音信号

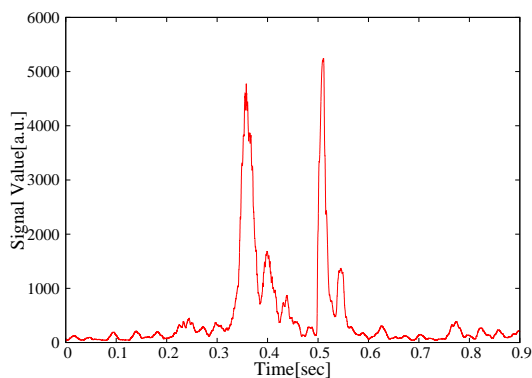
図 4.5: 整流した信号

続いて、式 (3.2) において、整流した嚙下音信号に対して平滑化を行った。 l を大きくすると II 音のピークが一つになってしまうため、 $l_1 = 300$ とした。各データは前後 6.8 msec のデータとの平均値となり、147.0 Hz 以上の周波数成分を除去することとなる。整流した嚙下音信号において、II 音の二つのピーク間隔の最小値が 11.0 msec (90.9 Hz) であるため、 $l_1 = 300$ よりも大きくすることで二つのピークが平均化されてしまうと考えられる。図 4.6 に平滑化を施した波形の例を示す。平滑化の結果から、図 3.7(a) で示される、II 音における二つのピーク値の大きさ S_1 、 S_2 と二つのピークの間隔 τ を特徴量として抽出した。



(a) 平滑化を施した 5 mL の嚙下音信号

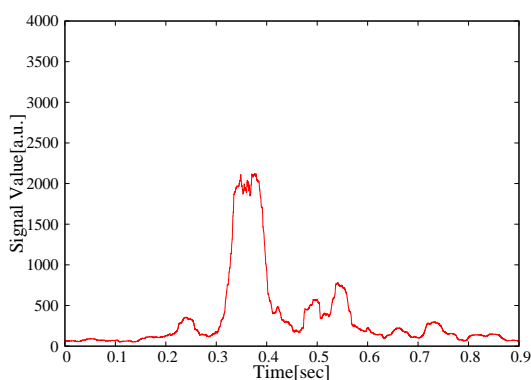
(b) 平滑化を施した 10 mL の嚙下音信号



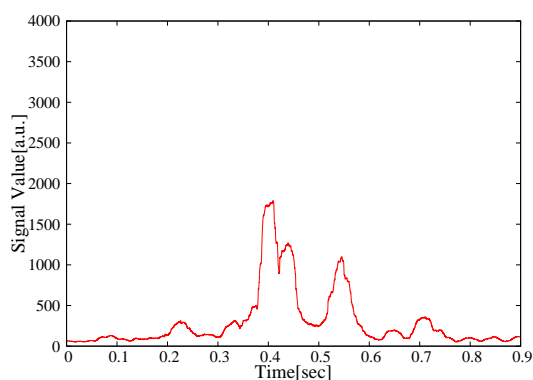
(c) 平滑化を施した 15 mL の嚙下音信号

図 4.6: 平滑化した信号 ($l_1 = 300$)

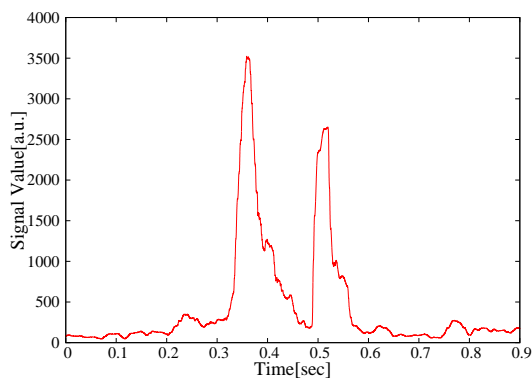
さらに、 l を一定以上にするとI音の識別が不可能となるデータが存在したため、 $l_2 = 600$ として同様に平滑化を行った。I音の振幅値 S_3 の最小値は120であり、 $l_2 = 600$ より大きくすることで、振幅が0に近いデータとさらに平均化されて識別が不可能になると考えられる。図4.7に平滑化を施した波形の例を示す。平滑化の結果から、図3.7(b)で示される、I音、II音、III音の振幅値(S_3, S_4, S_5)、I音とII音、II音とIII音、I音とIII音の振幅値の比($S_{3,4}, S_{3,5}, S_{4,5}$)、I音とII音、II音とIII音、I音とIII音の間隔($\tau_{3,4}, \tau_{4,5}, \tau_{3,5}$)、I音とIII音の間隔に対するI音とII音の間隔の割合($\tau_{3,4}/\tau_{3,5}$)をそれぞれ特徴量として抽出した。



(a) 平滑化を施した 5 mL の嚔下音信号



(b) 平滑化を施した 10 mL の嚔下音信号



(c) 平滑化を施した 15 mL の嚔下音信号

図 4.7: 平滑化した信号 ($l_2 = 600$)

4.2.2 線形予測分析による特徴量の抽出

計測した嚙下音信号に対して、線形予測分析を行った。直前の μ 個の信号点の値を用いて線形和を行う場合を、次数 μ の線形予測分析と呼ぶ。まず、最適な次数を決定する必要がある。次数を決定するための指標として、残差誤差を用いた最終予測誤差規範 (Final Prediction Error : FPE) が挙げられる [50]。FPE は式 (4.1) にて表される指標である。 N はデータ点数、 μ は次数、 \hat{e}^2 は平均残差誤差を表す。

$$\text{FPE}(\mu) = \frac{N + \mu}{N - \mu} \hat{e}^2 \quad (4.1)$$

次数 μ を 2 から 80 まで変化させて線形予測分析を行った場合の FPE の推移の例を図 4.8 に示す。次数が 60 前後で FPE が最小となった。次数が 60 の場合、直前の 1.36 msec のデータを用いて線形予測を行う。嚙下音信号には 1 msec 未満の間隔でデータの値の符号が変化する部分もあり、用いるデータを 1.36 msec よりも多くすることで予測係数に誤差が生じてしまうと考えられる。したがって、式 (3.3) における線形予測分析の次数を 60 とし、信号全体を線形予測した時の線形予測係数 $d_1 \sim d_{60}$ と、II 音部分を線形予測した時の線形予測係数 $d'_1 \sim d'_{60}$ を特徴量として抽出した。

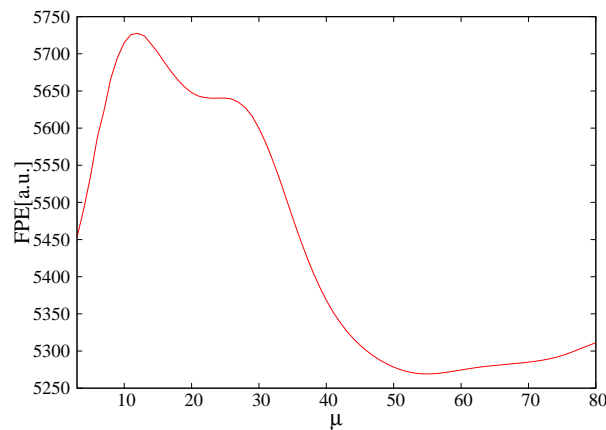


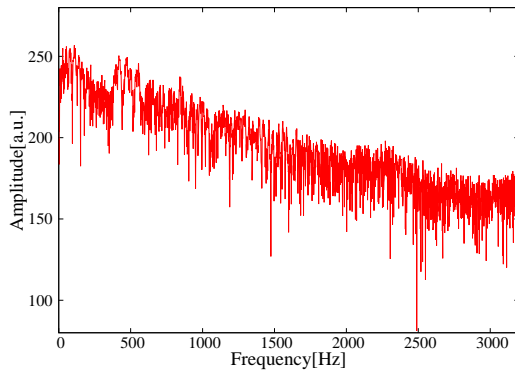
図 4.8: 次数による FPE の推移の例

4.2.3 離散フーリエ変換による特徴量の抽出

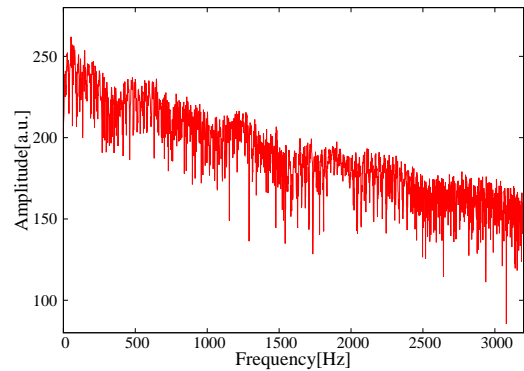
嚙下音信号の始点と終点の近傍では振幅値が0に近いので、式(4.2)で表される矩形窓関数 u_i と、計測した嚙下音 x_i の積 $u_i x_i (i = 0, 1, \dots, 39999)$ に対して、離散フーリエ変換を施し周波数解析を行った。フーリエ変換を高速に計算するために、 $0 \leq i \leq 32767$ で矩形窓関数 $u_i = 1$ とした。

$$u_i = \begin{cases} 1 & (i = 0, 1, \dots, 32766, 32767) \\ 0 & (\text{otherwise}) \end{cases} \quad (4.2)$$

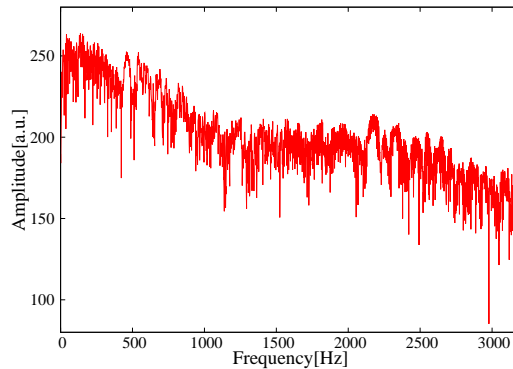
図4.4(a)~(c)に示す嚙下音に対して離散フーリエ変換を施した結果の例を図4.9に示す。



(a) 5 mL 嚙下音のフーリエ解析結果



(b) 10 mL 嚙下音のフーリエ解析結果



(c) 15 mL 嚙下音のフーリエ解析結果

図 4.9: フーリエ解析結果

まず、図3.8(a)で示される、最も強い周波数スペクトルの値 G と対応する周波数 f_1 を特徴量として抽出した。次に、マイクロフォンの周波数特性が $-3 \sim 0$ dB between 15~20000 Hz であることと、嚙下音信号が 3000Hz 以下にスペクトルを持つことから、25~3200 Hz を $\frac{1}{3}$

オクターブ毎で幅の異なる 21 個の帯域に分割して $\frac{1}{3}$ オクターブ分析を行った。結果から、図 3.8(b) で示される、25~3200 Hz のスペクトル和 G_t と G_t に対する各帯域内のスペクトル和の割合 $h_1 \sim h_{21}$ を特徴量として抽出した。したがって、 $h_1 \sim h_{21}$ は G_t を 1 として正規化されることとなる。さらに、25~3200 Hz における重心周波数 f_m を特徴量として抽出した。

4.2.4 ウェーブレット変換による特徴量の抽出

計測した嚙下音に対して、ウェーブレット変換を施し、時間周波数解析を行った。マザーウェーブレット関数として、信号の局所的な周波数成分を検出するために、周波数成分の分解能が高いガボールウェーブレット関数を用いた。ガボールウェーブレット関数は式 (4.3) で表され、図 4.10 に示すような概形を持つ。

$$\psi(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{t^2}{2\sigma^2}\right) \exp(j\omega_0 t) \quad (4.3)$$

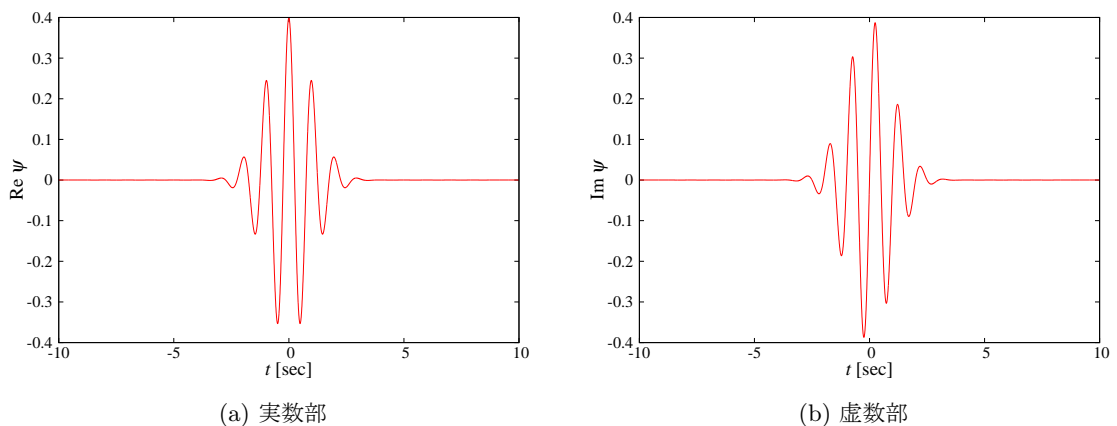
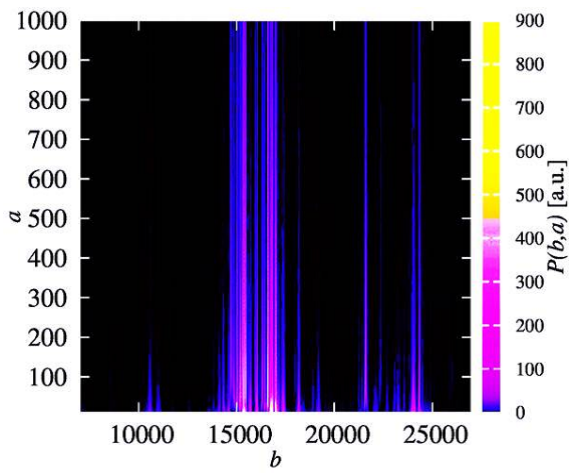
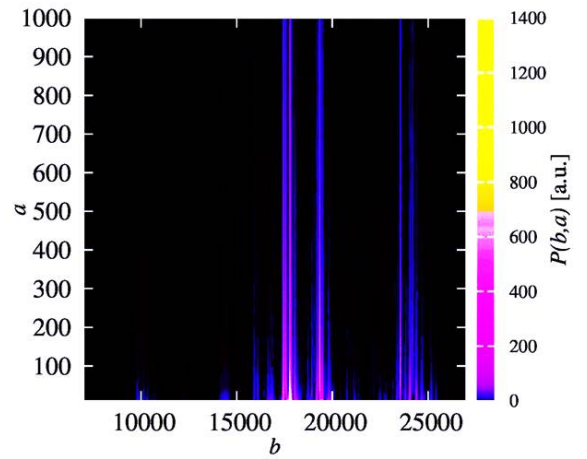


図 4.10: ガボールウェーブレット関数

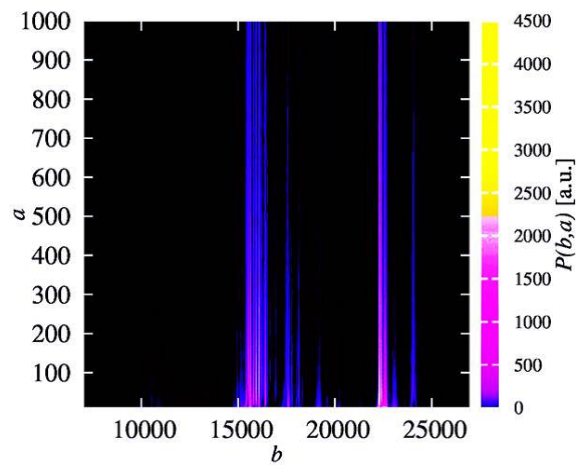
ガボールウェーブレット関数は減衰を表す指数関数と三角関数の積で表される。ウェーブレット関数の周波数は式 (3.15) に示される a の逆数で表される。図 4.11 にガボールウェーブレット変換した結果のスカログラムの例を示す。結果から、図 3.12 で示されるスカログラムの最大値 P_{max} と最大値を示す時の周波数 a_{max} を特徴量として抽出した。



(a) 5 mL 嚙下音のウェーブレット変換結果



(b) 10 mL 嚙下音のウェーブレット変換結果



(c) 15 mL 嚙下音のウェーブレット変換結果

図 4.11: ウェーブレット変換結果

4.3 抽出した特徴量の解析結果

本節では、抽出した特徴量に対して主成分分析とサポートベクタマシンを用いて解析を行った結果について述べる。

4.3.1 主成分分析による解析結果

嚙下音信号に対して平滑化、線形予測分析、フーリエ変換、ウェーブレット変換を施して160個の特徴量を抽出した。まず、特徴量を低次元化するために、抽出した160個の特徴量を用いて主成分分析を行った。解析は各被験者のみと全被験者を合わせた場合でそれぞれ行った。結果として、第1主成分から第15主成分までの累積寄与率が被験者Aでは90.1%、被験者Bでは91.6%、被験者Cでは90.5%、全被験者では91.1%となった。表4.2に、被験者Aの場合の各主成分における寄与の大きい5個の特徴量を例として示す。赤字で示される線形予測係数、青字で示される時系列信号での振幅値や時間間隔、緑字で示される $\frac{1}{3}$ オクターブ分析結果が多く現れた。得られた結果に基づき、160個の特徴量を、各被験者において累積寄与率が90%以上となる第1主成分から第15主成分までの15個の特徴量に縮約した。

表 4.2: 各主成分における寄与の大きい特徴量

寄与率の順位	1	2	3	4	5
第1主成分	d_{43}	d_{44}	d_{45}	d_{42}	d_{46}
第2主成分	d'_{19}	d'_{20}	d'_{18}	d_{22}	d'_{21}
第3主成分	d'_{12}	d'_{13}	d'_{11}	d'_{14}	d'_{10}
第4主成分	d_1	d_2	d'_{59}	d_{59}	d_{60}
第5主成分	h_{17}	f_m	h_{20}	h_{18}	h_{21}
第6主成分	h_{19}	d'_1	d'_2	G	d'_3
第7主成分	d_{59}	d'_{59}	h_{20}	d'_4	d'_3
第8主成分	S_5/S_3	h_{21}	S_5	f_1	S_4/S_3
第9主成分	$\tau_{3,4}$	G_t	S_5	P_{max}	$\tau_{3,5}$
第10主成分	S_4/S_3	S_4	S_5/S_4	S_3	$\tau_{3,4}$
第11主成分	S_5/S_3	S_2	S_5	S_3	$\tau_{4,5}$
第12主成分	$\tau_{3,5}$	$\tau_{4,5}$	S_3	S_5/S_4	$\tau_{3,4}/\tau_{3,5}$
第13主成分	h_{21}	$\tau_{3,4}/\tau_{3,5}$	h_9	h_{16}	d_{20}
第14主成分	S_3	τ	$\tau_{3,4}$	$\tau_{3,4}/\tau_{3,5}$	h_{21}
第15主成分	$\tau_{3,4}/\tau_{3,5}$	$\tau_{3,4}$	τ	S_4/S_3	S_3

4.3.2 SVMによる解析結果

SVMによって5 mLと10 mL、15 mLを分類するための識別器の決定を試みた。決定した識別器を一つ抜き交差法によって評価することで、最適なカーネル関数とマージンパラメータを決定した。SVMで用いるカーネルとして、線形カーネル、ラジアル基底関数 (Radial Basis Function: RBF) カーネル、多項式カーネルを用いた。線形カーネル、RBFカーネル、多項式カーネルはそれぞれ式(4.4)~式(4.6)で表される。ただし、 γ と η は定数である。線形カーネルを用いた場合はソフトマージンSVMと等価である。RBFカーネルでは、入力空間を教師データ数の部分領域に分割して各クラスを分類することができる。多項式カーネルでは、超双曲面で各クラスを分類することができる。

$$K(\mathbf{x}_i, \mathbf{x}_k) = \mathbf{x}_i^T \mathbf{x}_k \quad (4.4)$$

$$K(\mathbf{x}_i, \mathbf{x}_k) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_k\|) \quad (4.5)$$

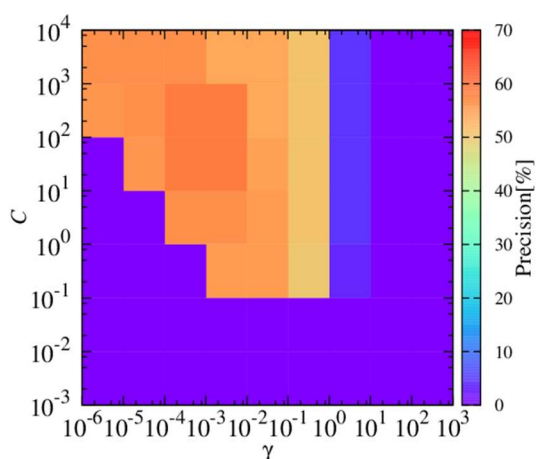
$$K(\mathbf{x}_i, \mathbf{x}_k) = (\mathbf{x}_i^T \mathbf{x}_k + 1)^\eta \quad (4.6)$$

まず、線形カーネルを用い、マージンパラメータを{0.01, 0.1, 1, 10, 100, 1000}と変化させて評価を行った。結果を表4.3に示す。被験者Aのみでは58.7% ($C = 0.1$)、被験者Bのみでは51.3% ($C = 10$)、被験者Cのみでは57.3% ($C = 1$)、全被験者では43.8% ($C = 1$)の嚙下音を正確に識別することができた。

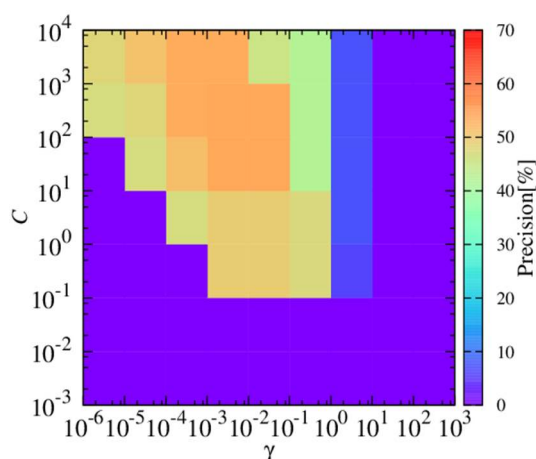
表 4.3: 線形カーネルにおける分類結果

マージンパラメータ C	0.01	0.1	1	10	100	1000
正答率 (被験者 A のみ)	51.3%	58.7%	55.3%	53.3%	54.0%	54.0%
正答率 (被験者 B のみ)	50.0%	48.7%	50.0%	51.3%	50.7%	50.7%
正答率 (被験者 C のみ)	35.3%	56.7%	57.3%	55.3%	55.3%	55.3%
正答率 (全被験者)	42.2%	42.2%	43.8%	43.3%	43.3%	43.3%

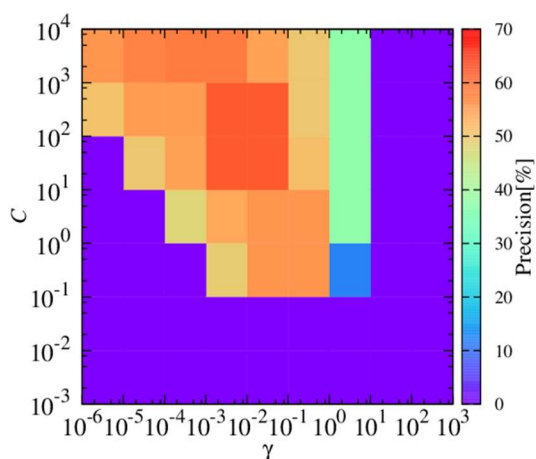
次に、RBF カーネルを用い、マージンパラメータを $\{0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000\}$ 、カーネルの定数 γ を $\{0.000001, 0.00001, 0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000\}$ と変化させて評価を行った。結果を図 4.12 に示す。被験者 A のみでは 58.7 % ($C = 100$ 、 $\gamma = 0.001$)、被験者 B のみでは 54.0 % ($C = 100$ 、 $\gamma = 0.01$)、被験者 C のみでは 62.0 % ($C = 100$ 、 $\gamma = 0.01$)、全被験者では 48.9 % ($C = 100$ 、 $\gamma = 0.01$) の嚙下音を正確に識別することができた。



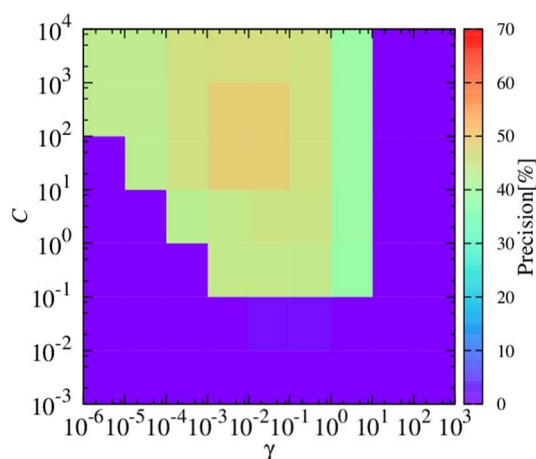
(a) 被験者 A のデータを用いた場合の分類結果



(b) 被験者 B のデータを用いた場合の分類結果



(c) 被験者 C のデータを用いた場合の分類結果



(d) 全被験者のデータを用いた場合の分類結果

図 4.12: RBF カーネルを用いた SVM による分類結果

さらに、多項式カーネルを用い、マージンパラメータを $\{0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000\}$ 、カーネルの定数 η を $\{1, 2, 3, 4, 5, 6, 7, 8\}$ と変化させて評価を行った。結果を図 4.13 に示す。被験者 A のみでは 52.7 % ($C = 10, \eta = 3$)、被験者 B のみでは 49.3 % ($C = 100, \eta = 3$)、被験者 C のみでは 46.7 % ($C = 1000, \eta = 3$)、全被験者では 44.4 % ($C = 10, \eta = 2$) の嚙下音を正確に識別することができた。

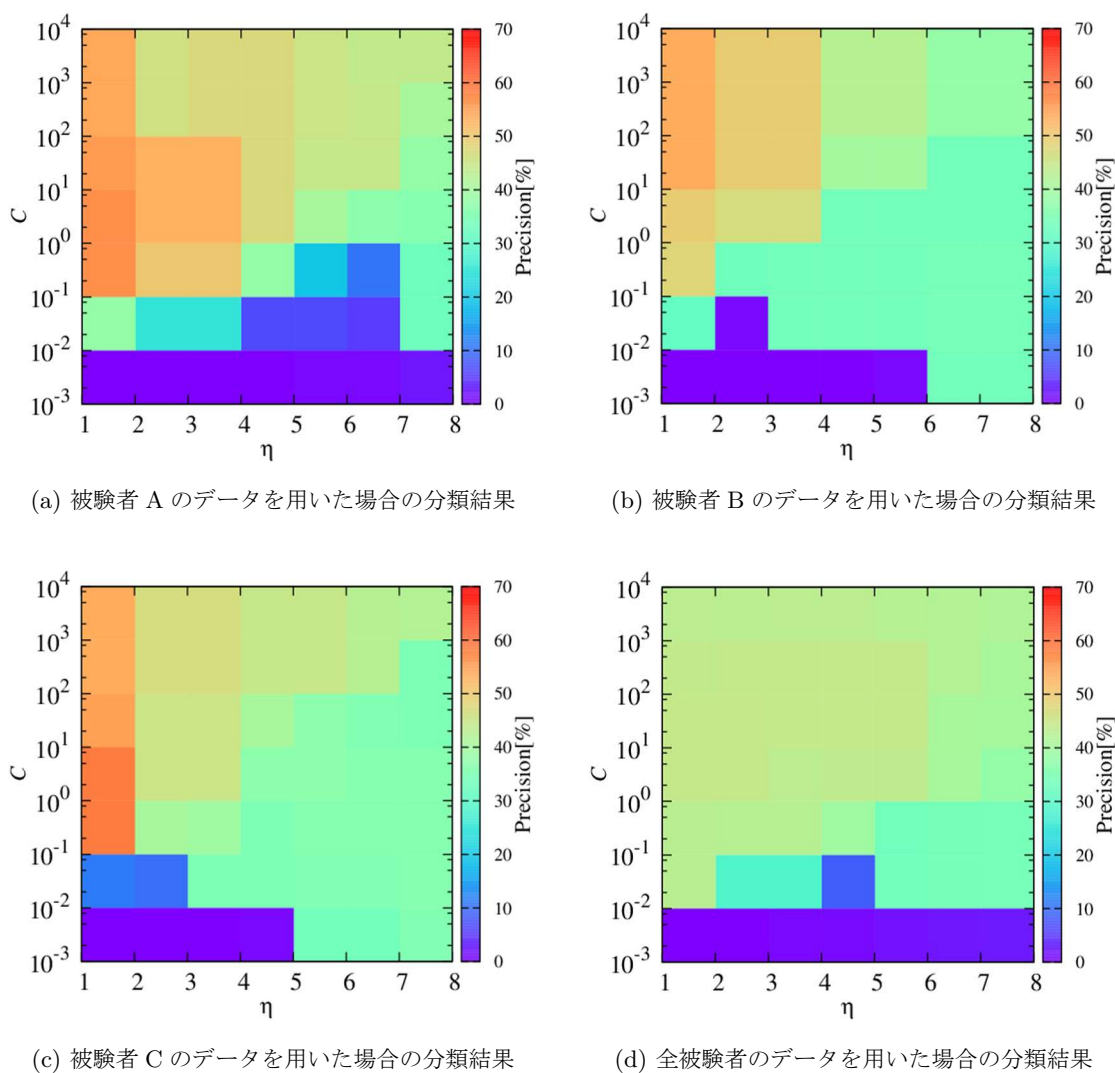


図 4.13: 多項式カーネルを用いた SVM による分類結果

得られた結果から、各被験者のみのデータを用いた場合の方が正答率が高いため、SVM による学習は被験者毎に行う。未知データに対する嚙下量の推定は、被験者 A では RBF カーネルで $C = 100, \gamma = 0.001$ 、被験者 B では RBF カーネルで $C = 100, \gamma = 0.01$ 、被験者 C では RBF カーネルで $C = 100, \gamma = 0.01$ と設定した時に得られた識別器を用いて行う。

次に、被験者3名に対して容量5 mL、10 mL、15 mLの水を嚥下した際の嚥下音を、未知データとしてさらに各10回計測した。前節で各被験者毎に決定した識別器を用いて、各嚥下音に対して摂取量の推定を行った。真値に対して、5 mL、10 mL、15 mLに分類されたデータの個数を表4.4～表4.6に示す。被験者Aでは60.0%、被験者Bでは66.7%、被験者Cでは96.7%の精度で嚥下音を正しく推定することができた。

表 4.4: 被験者 A の未知データの分類結果

推定結果 (mL)	真値 (mL)		
	5 mL	10 mL	15 mL
5 mL	7 個	2 個	1 個
10 mL	2 個	5 個	3 個
15 mL	1 個	3 個	6 個

表 4.5: 被験者 B の未知データの分類結果

推定結果 (mL)	真値 (mL)		
	5 mL	10 mL	15 mL
5 mL	7 個	0 個	1 個
10 mL	2 個	7 個	3 個
15 mL	1 個	3 個	6 個

表 4.6: 被験者 C の未知データの分類結果

推定結果 (mL)	真値 (mL)		
	5 mL	10 mL	15 mL
5 mL	10 個	1 個	0 個
10 mL	0 個	9 個	0 個
15 mL	0 個	0 個	10 個

4.4 連続的嚥下音の総摂取水分量の推定

被験者3名に対して、50 mLと100 mLの水を複数回の嚥下にわたって摂取した際の嚥下音を各10回計測した。50 mLの水は摂取に4~7回の嚥下を要し、嚥下総数は162回、100 mLの水は摂取に8~12回の嚥下を要し、嚥下総数は301回であった。計測した時の波形を図4.14に示す。矢印で示された部分が一回の嚥下が行われた時間を表す。

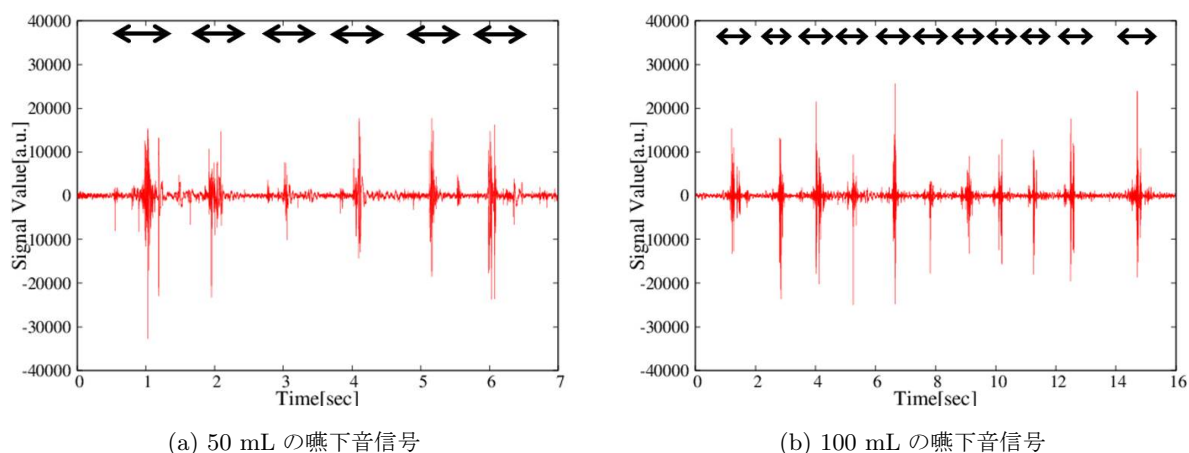
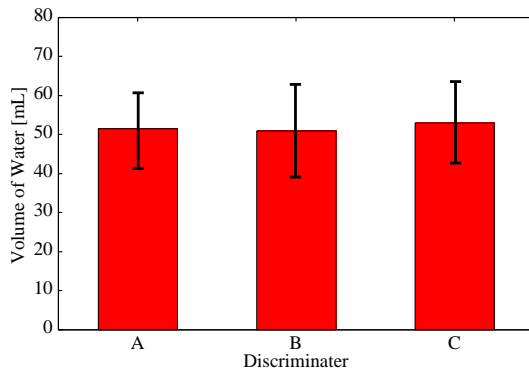


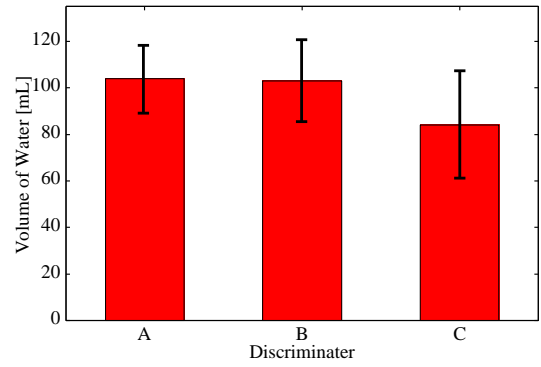
図 4.14: 50 mL と 100 mL の嚔下音信号

計測した信号を一嚔下毎に分離し、各嚔下音に対する音声信号処理結果から 160 個の特徴量を抽出した。特徴量を 15 個へと縮約し、前節で決定した識別器によって 5 mL、10 mL、15 mL のいずれであるかを分類した。分類した摂取水分量を積算することで総嚔下量の推定を行った。また、各嚔下音に対する分類は、被験者毎に決定した 3 個の識別器をそれぞれ適用し、各 3 回行った。各被験者において、3 個の識別器によって総摂取水分量を推定した結果を図 4.15 に示す。

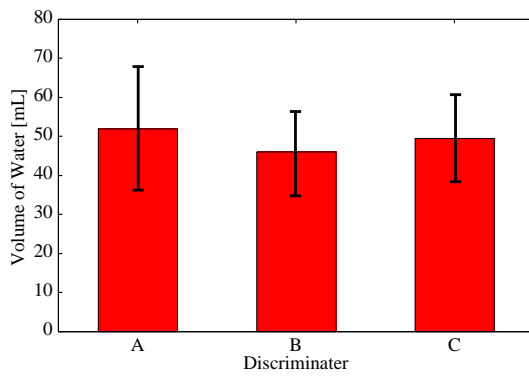
同一被験者の学習データから決定された識別器を用いた場合、50 mL の嚔下音は被験者 A では平均 51.5 mL、標準偏差 10.0 mL、被験者 B では平均 46.0 mL、標準偏差 11.0 mL、被験者 C では平均 58.5 mL、標準偏差 10.0 mL の精度で推定できることが確認できた。さらに、100 mL の嚔下音は被験者 A では平均 104.0 mL、標準偏差 14.9 mL、被験者 B では平均 94.5 mL、標準偏差 14.6 mL、被験者 C では平均 101.5 mL、標準偏差 20.8 mL の精度で推定できることが確認できた。また、異なる被験者の学習データから決定された識別器を用いた場合、平均値と真値との差が大きくなる、標準偏差が大きくなるといった傾向が見られた。



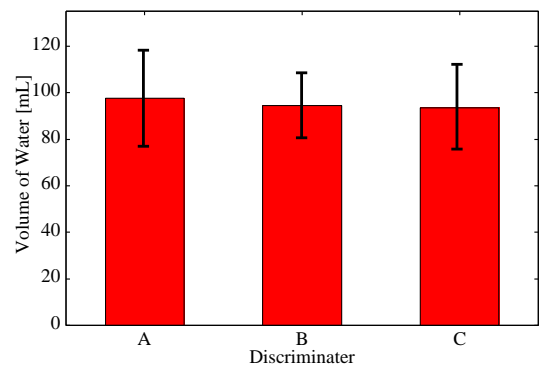
(a) 被験者 A の 50 mL の推定結果



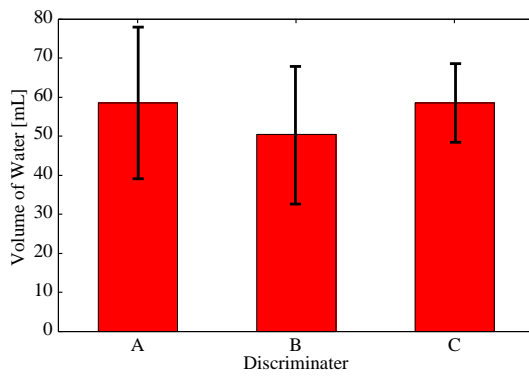
(b) 被験者 A の 100 mL の推定結果



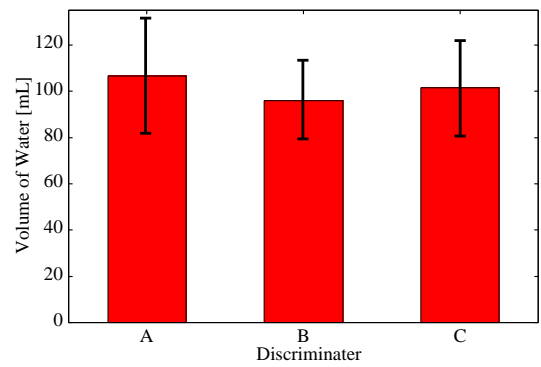
(c) 被験者 B の 50 mL の推定結果



(d) 被験者 B の 100 mL の推定結果



(e) 被験者 C の 50 mL の推定結果



(f) 被験者 C の 100 mL の推定結果

図 4.15: 被験者 3 名における摂取水分量の推定結果

次に、識別器の学習に寄与していない新たな 10 名の被験者に対して、50 mL と 100 mL の水を複数回の嚥下にわたって摂取した際の嚥下音を各 1 回計測した。前述の実験と同様に、計測した信号を一嚥下毎に分離し、被験者 A、B、C の学習データから決定した 3 個の識別器を用いて分類された嚥下量を、積算していくことで総嚥下量の推定を行った。総摂取水分量を推定した結果を図 4.16 に示す。50 mL の嚥下音は被験者 A の識別器を用いた場合、平均 52.0 mL、標準偏差 15.1 mL、被験者 B の識別器を用いた場合、平均 46.5 mL、標準偏差 12.3 mL、被験者 C の識別器を用いた場合、平均 50.5 mL、標準偏差 10.2 mL で推定することができた。さらに、100 mL の嚥下音は被験者 A の識別器を用いた場合、平均 104.5 mL、標準偏差 23.7 mL、被験者 B の識別器を用いた場合、平均 90.0 mL、標準偏差 18.4 mL、被験者 C の識別器を用いた場合、平均 95.5 mL、標準偏差 19.8 mL で推定することができた。

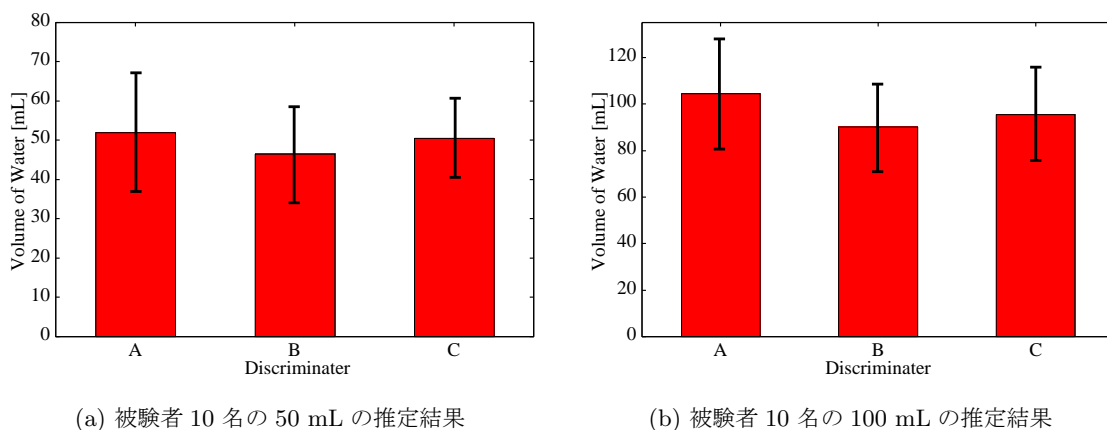


図 4.16: 異なる 10 名の被験者における摂取水分量の推定結果

4.5 考察

まず、3 名の被験者から嚥下音を計測した。次に、嚥下音に対して、平滑化、線形予測分析、離散フーリエ変換、ガボールウェーブレット変換を施し、解析を行った。解析結果から、嚥下量を反映する可能性のある 160 個の特徴量を抽出した。主成分分析と SVM によって、5 mL と 10 mL、15 mL の嚥下音を分類するための識別器を決定したところ、3 種類の嚥下音は識別可能であることが示唆された。識別の正答率が特に高かった被験者においては、学習データと未知データとの特徴が類似していたと考えられる。誤分類された 5 mL と 15 mL の嚥下音は、10 mL へ分類されることが多かったため、嚥下音は摂取量に応じた特徴を持つ

ている可能性が考えられる。主成分分析によって得られた特徴量において、線形予測係数の寄与が大きいことから、嚥下量に応じて嚥下音の発生モデルに相違があることが考えられる。また、I音、II音、III音のピーク値や間隔の寄与も大きいことから、嚥下量に応じて喉頭や咽頭が異なる運動をすることも要因として考えられる。さらに、周波数帯域が800~3200 Hzのスペクトル和に関する係数の寄与が大きいことから、嚥下量の違いは比較的高い周波数成分に影響を与えることが考えられる。

また、複数被験者のデータを用いて識別器を決定した場合、一人の被験者のデータのみを用いたときよりも分類の正答率が低下することが確認できた。したがって、嚥下する際の喉頭や咽頭の動きには被験者による個人差があることが示唆されており、分類のための識別器を作成するためには一人の被験者のデータのみを用いて行うことが望ましいと考えられる。

次に、3名の被験者に対して、50 mLと100 mLの水を複数回の嚥下によって摂取した際の嚥下音の計測を行った。計測した連続的な嚥下音を一回毎に分割し、それぞれ160個の特徴量を抽出した。主成分分析とSVMによって決定した識別器を用いて、摂取水分量を分類していくことで総水分量の推定を行った。異なる被験者の学習データから決定された識別器を用いて推定を行った場合、平均値と真値との差が大きくなる、標準偏差が大きくなるといった傾向が見られたことから、嚥下する際の喉頭や咽頭の動きには被験者による個人差があることが同様に確認できた。さらに、識別器の学習に寄与していない新たな10名の被験者に対して、50 mLと100 mLの水を摂取した際の嚥下音を計測した。前述の実験と同様の識別器を用いて、一嚥下毎に分類された摂取量を積算することで、総水分量の推定を行った。結果として、同一被験者の学習データから決定された識別器を用いた場合よりも、標準偏差が大きくなる等の傾向は見られたが、各一回の嚥下における摂取量を積算して総水分量を推定できることが示唆された。したがって、あらかじめ学習データから決定した識別器を用いることで、不特定多数の摂取水分量も推定できると考えられる。

第5章 おわりに

本研究では、嚥下音に対して信号処理を行うことで、ユーザがどの程度の水分を摂取したかを定量的に計測するシステムの構築を目的とした。本稿ではまず、被験者3名に対して5 mL、10 mL、15 mLの水を嚥下した際の嚥下音を、甲状軟骨付近において学習用データとして各50回計測した。続いて、計測した嚥下音に対して平滑化、線形予測、フーリエ変換、ウェーブレット変換によって信号処理を行い、得られた処理結果から一度の嚥下での摂取水分量を反映する160個の特徴量を抽出した。さらに、主成分分析を行い、元の情報量の90%を保持した15個の特徴量を抽出し、得られた15個の特徴量を用いて、SVMによって摂取水分量の推定するための識別器を決定した。結果として、5 mLと10 mL、15 mLの嚥下音を、各被験者においてそれぞれ70.0%、66.7%、96.7%の精度で推定することができた。次に、50 mLと100 mLの水を複数回の嚥下によって摂取した際の嚥下音の計測を各10回行った。計測した連続的な嚥下音を、一回毎の嚥下音に分割し、決定した識別器を用いて摂取した水分量を分類していくことで総摂取量の推定を行った。結果として、50 mLの嚥下音は平均52.0 mL、標準偏差11.3 mL、100 mLの嚥下音は平均100.0 mL、標準偏差16.9 mLの精度で推定することができた。したがって、本システムによって摂取した総水分量の推定が可能であると考えられる。本システムを用い、一日に摂取できる水分量が制限される患者に対して摂取水分量のモニタリングを行うことで、水分補給の補助を行う等の貢献が期待される。

謝 辞

本研究は大阪大学 大学院基礎工学研究科で行ったものである。

本研究において、研究の機会を与えて頂くと共に、人数が多くお忙しい中で、本論文の添削において丁寧にご指導頂きました大阪大学 大学院基礎工学研究科 大城理教授に篤く感謝致します。本論文の副査を引き受けて頂いた大阪大学 大学院基礎工学研究科 野村泰伸教授に深く感謝致します。また、的確なアドバイスを与えてくださり、本論文の執筆に関してもご指導頂いた大阪大学 大学院基礎工学研究科 井村誠孝准教授に感謝申し上げます。プログラムのご指導をして頂いた大阪大学 サイバーメディアセンター 黒田嘉宏准教授、食事などの際に多くの話をして頂きました京都大学医学部附属病院 浦西友樹助教、研究に取り組む姿勢に対してアドバイスをくださった大阪大学 大学院基礎工学研究科 吉元俊輔助教にも心より感謝致します。

お会いする機会は少ないものの、飲み会の際には優しく接して頂いた高畑裕美氏、三宅正夫氏、半年と短い時間ではありましたが優しく声をかけて頂いた加藤雄樹氏、研究に行き詰まった際には相談に乗ってくださり、プログラム等に関して多大なアドバイスを頂いた井手口裕太氏、篠岡亮氏に感謝致します。また、共に励ましあい、切磋琢磨してきた研究室の同期、後輩達にお礼申し上げます。

本研究における被験者実験は、大阪大学大学院基礎工学研究科人を対象とした研究に関する倫理委員会の承認 (25-4) を得ており、被験者の同意を得て行ったものである。

業績

- 国内論文

1. 中藤 寛己, 井村 誠孝, 吉元 俊輔, 大城 理, “嚙下音解析による摂取水分量の推定システム”, 生体医工学 (掲載予定).

- 国内発表

1. 中藤 寛己, 井村 誠孝, 黒田 嘉宏, 浦西 友樹, 大城 理, “摂取水分量計測のための嚙下音の解析”, 第 56 回自動制御連合講演会, 演題番号 817, 新潟, 2013 年 11 月.
2. 中藤 寛己, 井村 誠孝, 浦西 友樹, 吉元 俊輔, 大城 理, “嚙下音解析による摂取水分量の推定”, 第 53 回日本生体医工学会大会, 演題番号 O1-01-6, 仙台, 2014 年 6 月.
3. 中藤 寛己, 井村 誠孝, 吉元 俊輔, 大城 理, “嚙下音解析による摂取水分量の推定システム”, 生体医工学シンポジウム 2014, 演題番号 2P-06, 小金井, 2014 年 9 月.

- その他

1. 中藤 寛己, “嚙下音解析を用いた摂取水分量の推定”, 生体工学領域交流会, 発表番号 35, 豊中, 2014 年 7 月.

参考文献

- [1] 秋道 智彌, 小松 和彦, 中村 康夫, “人と水 水と文化”, 勉誠出版, 東京, pp.8-9, 2010.
- [2] 上平 恒, 逢坂 昭, “生体系の水”, 講談社, 東京, pp.133-134, 1989.
- [3] 小泉 直也, 田中 秀和, 上間 裕二, 稲見 昌彦, “咀嚼音提示を利用した食感拡張装置の検討”, 日本バーチャルリアリティ学会論文誌, Vol.18, pp.141-150, 2013.
- [4] 米丸 亮, 桜井 利江, “ナースのための CD による呼吸音聴診トレーニング”, 南江堂, 東京, pp.26-29, 2001.
- [5] 藤島 一郎, “よくわかる嚥下障害”, 永井書店, 大阪, pp.98-102, 2012.
- [6] 奥田 克爾, 君塚 隆太, 阿部 修, 加藤 哲男, 石原 和幸, 西條 みのり, “口腔ケアによる誤嚥性肺炎予防”, 歯科学報, Vol.105, pp.129-137, 2005.
- [7] 高橋 浩二, “嚥下障害診断法としての頸部聴診法”, 昭和歯学会雑誌, Vol.25, pp.167-171, 2005.
- [8] 村上 彩子, 新妻 可奈子, 本橋 由香, 佐藤 敏夫, 川島 徳道, 澁谷 有美, 高木 絵美子, 本橋 尊, 星野 敏久, 阿岸 鉄三, 小見 勝利, “生体雑音の時間-周波数解析による非侵襲生体機能診断”, 電子情報通信学会技術研究報告.US, 超音波, Vol.109, pp.1-4, 2009.
- [9] Joon Lee, Stefanie Blain, Mike Casas, Dave Kenny, Glenn Berall, Tom Chau, “A Radial Basis Classifier for The Automatic Detection of Aspiration in Children with Dysphagia”, BioMed Central Journal of NeuroEngineering and Rehabilitation, pp.1-17, 2006.
- [10] 渡部 雄祐, 渡邊 祐子, 浜田 晴夫, “複数のマイクロフォンによる嚥下音評価システムの検討”, 日本音響学会研究発表会, Vol.2009,ROMBUNNO.3-9-4, 2009.

- [11] 津川 英範, 山本 尚武, 中村 隆夫, “嚥下音による嚥下活動の解析に関する研究”, 電子情報通信学会技術研究報告.MBE, ME とバイオサイバネティックス, Vol.99, pp.31-36, 1999.
- [12] 小濱 隆司, 中村 尚五, “デジタルホルター心電計におけるノイズキャンセル処理”, 電子情報通信学会論文誌.A, 基礎・境界, Vol.77, pp.750-757, 1994.
- [13] 桐山 希一, “足底圧の動的変化を指標とした健常成人の歩行制御および片麻痺歩行に関する研究”, 医療保健学研究, Vol.3, pp.1-40, 2012.
- [14] 川原田 淳, 嶋津 秀昭, 山越 憲一, 戸川 達男, “無拘束非観血的連続血圧計測”, 電子情報通信学会技術研究報告.MBE, ME とバイオサイバネティックス, Vol.93, pp.41-48, 1993.
- [15] 岡崎 浩也, 辻村 肇, 土井 英明, 松村 雅史, “口腔咽喉の分析に基づく無拘束嚥下回数の計測音”, 電子情報通信学会技術研究報告.MBE, ME とバイオサイバネティックス, Vol.109, pp.15-20, 2009.
- [16] Masaki Shuzo, Komori Shintaro, Takashima Tomoko, Lopez Guillaume, Tatsuya Seiji, Yanagimoto Shintaro, Warisawa Shinichi, Delaunay Jean-Jacques, Yamada Ichiro, “Wearable Eating Habit Sensing System Using Internal Body Sound”, Journal of Advanced Mechanical Design, pp.158-166, 2010.
- [17] 上田 菜美, 野原 幹司, 小谷 泰子, 阪井 丘芳, “一回嚥下量の変化が嚥下時の舌骨運動に与える影響について”, 日本摂食・嚥下リハビリテーション学会誌, Vol.17, pp.36-44, 2008.
- [18] 田中 栄士, 小野 圭昭, 権田 悦通, “嚥下時の下顎運動と口腔内圧, 咽頭圧の解析”, 歯科医学会誌, Vol.64, pp.152-160, 2001.
- [19] 中澤 文子, “摂食から咀嚼・嚥下過程”, 日本調理科学会誌, Vol.39, pp.185-193, 2006.
- [20] 佐々木 啓一, “咀嚼・嚥下機能の検査・診断システムの確立に向けて”, 東北大学歯学雑誌, Vol.20, pp.75-90, 2001.
- [21] 藤島 一郎, 柴本 勇, “摂食・嚥下障害患者のリスクマネジメント”, 中山書店, 東京, pp.30-35, 2009.

- [22] 岩手医科大学附属歯科医療センター 口腔リハビリ外来, “<http://denture.iwate-med.ac.jp/cn21/cn17/dysphagia650.html>”(参照 2013-7-30).
- [23] 椿原 彰夫, 谷本 啓二, 馬場 尊, 藤島 一郎, 北住 映二, 岡田 澄子, “嚥下造影の標準的検査法”, 日本摂食・嚥下リハビリテーション学会誌, Vol.5, pp.166-167, 2001.
- [24] 高橋 浩二, “嚥下障害診断法としての頸部聴診法”, 昭和歯学会雑誌, Vol.25, pp.167-171, 2005.
- [25] 内田 学, 原田 亜由美, 加藤 宗規, “超音波画像診断装置と心音マイクを用いた嚥下時の咽頭運動評価の妥当性と再現性”, 理学療法科学, Vol.27, pp.539-543, 2012.
- [26] 中沢 文子, 大野 真由子, 盛田 明子, 高橋 淳子, “嚥下した飯・かゆ・水の超音波による咽頭部の流速比較”, 日本家政学会誌, Vol.51, pp.1067-1071, 2000.
- [27] 長谷川 温子, 中沢 文子, 熊谷 仁, “超音波により測定した咽頭部での液状食品の流速に及ぼす嚥下量の影響”, 日本食品科学工学会誌, Vol.55, pp.330-337, 2008.
- [28] 関 八州彦, “頸部インピーダンス波形の測定による嚥下機能評価”, 川崎医学会誌, Vol.23, pp.165-173, 1997.
- [29] 山下 光美, 松村 雅史, 小谷 泰子, 田中 信和, 野原 幹司, “廃用性委縮防止のための喉頭マイクروفオンを用いた嚥下回数計測システム”, 電子情報通信学会技術研究報告.MBE, ME とバイオサイバネティクス, Vol.107, pp.13-16, 2007.
- [30] 中村 隆夫, 楠原 俊昌, 森 恵子, 山本 尚武, “嚥き動作を伴う嚥下時のインピーダンス咽頭図の解析法”, 電子情報通信学会技術研究報告.MBE, ME とバイオサイバネティクス, Vol.103, pp.47-50, 2003.
- [31] 小谷 潔, 古川 健太, 高増 潔, “嚥下の心拍変動に与える影響の呼吸位相領域における解析”, 生体医工学, Vol.42, pp.61-65, 2004.
- [32] Lisa Lazareck, Zahra Moussavi, “Classification of normal and dysphagic swallows by acoustical means”, IEEE Transaction of Biomedical Engineering, Vol.51, pp. 2103-2112, 2004.

- [33] Ervin Sejdic, Catriona Steele, Tom Chau, “Segmentation of Dual-Axis Swallowing Accelerometry Signals in Healthy Subjects With Analysis of Anthropometric Effects on Duration of Swallowing Activities”, IEEE Transactions of Biomedical Engineering, Vol.56, pp.1090-1097, 2009.
- [34] 山下 光美, 松村 雅史, 小谷 泰子, 田中 信和, 野原 幹司, “廃用性委縮防止のための喉頭マイクロフォンを用いた嚥下回数計測システム”, 電子情報通信学会技術研究報告.MBE, ME とバイオサイバネティクス, Vol.107, pp.13-16, 2007.
- [35] Edward Sazonov, Oleksandr Makeyev, Stephanie Schuckers, Paulo Meyer, Edward Melanson, Michael Neuman, “Automatic Detection of Swallowing Events by Acoustical Means for Applications of Monitoring of Ingestive Behavior”, IEEE Transaction of Biomedical Engineering, Vol.57, pp. 626-633, 2010.
- [36] 渡部 雄祐, 渡邊 祐子, 浜田 晴夫, “複数のマイクロフォンによる嚥下音評価システムの検討”, 日本音響学会研究発表会, Vol.2009, ROMBUNNO.3-9-4, 2009.
- [37] 中山 裕司, 高橋 浩二, 宇山 理紗, 平野 薫, 深澤 美樹, 南雲 正男, “嚥下音の産生部位と音響特性の検討—健常成人を対象として—”, 昭和歯学会雑誌, Vol.26, pp.163-174, 2006.
- [38] Takahashi Koji, Grother Michael, Michi Kenichi, “Methodology for Detecting Swallowing Sounds”, Journal of Dysphagia, pp.54-62, 1994.
- [39] 新藤 恵一郎, 近藤 健男, 杉山 謙, 沖井 明, 出江 一, 小川 允洋, “嚥下音を利用し睡眠時嚥下回数測定を試み”, 日本リハビリテーション医学会誌, p.78, 2005.
- [40] 鹿野 清宏, 中村 哲, 伊勢 史郎, “音声・音情報のデジタル信号処理”, 昭晃堂, 東京, pp.10-16, 1997.
- [41] 中溝 高好, “信号解析とシステム同定”, コロナ社, 東京, pp.71-76, 1988.
- [42] 戸田 浩, 章 忠, 川畑 洋昭, “ウェーブレット実践講座 信号処理の基礎から最新理論まで”, ソフトバンククリエイティブ, 東京, pp.2-5, 2005.
- [43] 石村 貞夫, 加藤 千恵子, 劉 晨, 石村 友二郎, “多変量解析によるデータマイニング”, 共立出版, 東京, pp.170-173, 2010.

- [44] 石村 貞夫, “すぐわかる多変量解析”, 東京図書, 東京, pp.66-69, 1992.
- [45] 阿部 重夫, “パターン認識のためのサポートベクトルマシン入門”, 森北出版, 東京, pp.16-20, 2011.
- [46] 金谷 健一, “これなら分かる最適化数学”, 共立出版, 東京, pp.207-209, 2005.
- [47] 阿部 重夫, “パターン認識のためのサポートベクトルマシン入門”, 森北出版, 東京, pp.22-26, 2011.
- [48] Nello Cristianini, John Shawe-Taylor, “サポートベクターマシン入門”, 森北出版, 東京, pp.22-26, 2005.
- [49] 阿部 重夫, “パターン認識のためのサポートベクトルマシン入門”, 森北出版, 東京, pp.54-57, 2011.
- [50] 北川 源四郎, 小西 貞則, “情報量規準”, 朝倉書店, 東京, pp.34-42, 2004.
- [51] Masaki Shuzo, Komori Shintaro, Takashima Tomoko, Lopez Guillaume, Tatsuya Seiji, Yanagimoto Shintaro, Warisawa Shinichi, Delaunay Jean-Jacques, Yamada Ichiro, “Wearable Eating Habit Sensing System Using Internal Body Sound”, Journal of Advanced Mechanical Design, pp.158-166, 2010.