

筑波大学大学院博士課程

システム情報工学研究科修士論文

口の動きを利用したマスク型インタフェース
によるハンズフリーな入力手法

山崎 友翼

修士（工学）

（コンピュータサイエンス専攻）

指導教員 高橋 伸

2018年3月

概要

本研究では満員電車内のような手が使えない状況におけるモバイル端末への新たな入力手法としてマスク型インタフェースによる入力手法を提案する。既存のハンズフリーな入力手法、例えば音声入力では周囲の雑音の影響や公共の場における使用に対する抵抗などの問題があった。本手法はマスク型インタフェースによりモバイル端末の入力を行なうため、ハンズフリーで使用することができ、口の動きを周囲に知られることなく、また発声する必要がないため、公共の場における使用が可能である。また、デバイスに使用するマスク・導電糸は安価なため使い捨てのインタフェースとして手軽に使えることも利点となりうる。本研究では、マスク内に静電容量方式の格子状のセンサを取り付けることでマスク型インタフェースを実装した。実装したマスク型インタフェースを用いて口の形状を識別する実験を行なったところ、3種類の識別は高精度で行なうことができたが、4種類の識別は精度が極端に下がってしまった。しかし、この原因は口の形状の特徴量が似ていたためであることから、口の形状によってはこれ以上の識別ができることが示唆された。今後は、今回の実験での口の形状以外の口の形状についての実験を行ない正しい識別が行なえるかを調査する。そのほかにも、口の形状の識別の精度の向上および識別数の増加を行なう予定である。

目次

第1章	序論	1
1.1	背景	1
1.2	目的とアプローチ	1
1.3	貢献	2
1.4	本論文の構成	2
第2章	関連研究	3
2.1	口の動きを検出する研究	3
2.2	モバイル端末のハンズフリーな操作手法に関する研究	4
2.3	静電容量計測技術を用いた導電体検出を行なう研究	4
2.4	本論文の位置付け	5
第3章	マスク型インタフェースによるモバイル端末操作	6
3.1	概要	6
3.2	ユーザによる使い方	7
3.3	利用シナリオ	9
3.3.1	音楽プレイヤーの操作	9
3.3.2	料理ナビゲーションアプリの操作	9
3.3.3	パスワードの入力への応用	9
第4章	実装	10
4.1	マスク型インタフェース	10
4.2	システムの構成	10
4.3	センサ部における静電容量計測	11
4.4	口の形状認識	13
4.5	ブレッドボード上での実装	14
4.6	口の形状識別アプリケーション	14
第5章	評価	21
5.1	被験者	21
5.2	実験内容	21
5.3	実験の手順	21

5.4 実験の結果および考察	23
第6章 結論	28
謝辞	29
参考文献	30

目 次

3.1	提案手法の概要	6
3.2	ユーザによる利用の流れ	8
4.1	平型マスクの外観	10
4.2	システムの構成	12
4.3	信号発生回路	13
4.4	信号処理回路	13
4.5	口の形状と静電容量分布のイメージ図	14
4.6	実装したシステムの全体図	16
4.7	信号発生回路の回路図	17
4.8	信号処理回路の回路図	17
4.9	マイクロコンピュータの回路図	18
4.10	作製したセンサ部	19
4.11	静電容量分布表示および口の形状認識を行なうアプリケーション	20
5.1	実験 1 での口の形状	22
5.2	実験 2 で追加した口の形状	22
5.3	実験 1 の混同行列	26
5.4	実験 2 の混同行列	26
5.5	default の静電容量値の一例	27
5.6	open の静電容量値の一例	27
5.7	sharpen の静電容量値の一例	27
5.8	bite の静電容量値の一例	27

第1章 序論

本章において、まず満員電車内などの手を使えない状況におけるモバイル端末を操作するときの問題およびその解決策における問題を述べる。次に本研究の目的を述べ、問題解決のためのアプローチを示す。その後、本研究の貢献を示し、本論文の構成を述べる。

1.1 背景

現代におけるモバイル端末への入力手法は、タッチスクリーンを用いる手法であるソフトウェア QWERTY キーボードやフリック入力が一般的である。そのほかにも追加のデバイスを用いる手法の一例として bluetooth キーボードをモバイル端末と接続し入力に使用することもある。これらの入力手法は片手もしくは両手を使用するが、モバイル端末を使用する場合には、手を使えない状況が存在する。例えば、混雑した電車内、両手に荷物を持っているとき、手が濡れている・汚れているときなどが考えられる。

そこでこの問題を解決するための、ハンズフリーな手を使わない入力インタフェースがある。手を使わないインタフェースには、例えば視線によるポインティングや音声認識による端末の操作、舌の動きによるジェスチャ入力など、数多くの手法がある。しかし、これらの手法においては、周囲への配慮が必要であることや、公共の場における使用に対する抵抗感 [1] があること、特に音声認識においては雑音が多い場所では検出が困難であるといった問題がある。

1.2 目的とアプローチ

本研究の目的は、手を使えない状況におけるモバイル端末への新たな入力手法の提案およびその手法による入力インタフェースの開発である。このインタフェースはハンズフリーで 사용할ことができ、かつ周囲に人がいるような公共の場においても抵抗を感じることなく使用できるものであるものを目指す。

この目的を達成するための研究のアプローチとして、マスク型インタフェースにより口の動きを検出する手法を提案する。マスク型のインタフェースとすることで、ウェアラブルかつハンズフリーな使用ができ、口の動きをマスクによって隠すことができる。口をマスクによって覆うことで、周囲に人がいるような公共の場においても抵抗なく使用することができる。また、発声する必要もないため、周囲の雑音の影響もない。

1.3 貢献

本研究における貢献は，モバイル端末におけるハンズフリーな操作手法としてマスク型インタフェースの設計を行なったことである．これにより手を使えない状況におけるモバイル端末の操作を，周囲への配慮を気にすることなく，また公共の場における抵抗を感じることなく行なうことができる．マスク型インタフェースは，一般的な使い捨てマスクと導電糸によって構成されるため，安価に作製することができ，使い捨てのインタフェースとして使用することができる．また，提案手法による口の形状識別の実験を行ない，3種類の識別は高精度で行なうことが確認できた．

1.4 本論文の構成

本章以降の本論文の構成は次の通りである．第2章では，関連研究と課題，本研究との差異を示すことにより，本研究の位置付けを述べる．第3章では，本研究の提案手法であるマスク型インタフェースによる入力手法について述べる．第4章では，提案手法について実際に行なった実装の詳細を述べる．第5章では，提案手法を用いた評価について述べる．最後に第6章では，結論を述べる．

第2章 関連研究

本章では，本研究に関連する研究について述べる．本研究は，マスク型インタフェースを用いて口の動きを推定し，モバイル端末の操作に使用する手法を提案するものである．よって，まず口の動きを検出する研究を述べ，つづいて，モバイル端末のハンズフリーな操作手法に関する研究，口の動きを推定するための手法である静電容量計測技術を用いた導電体検出を行なう研究について述べる．また最後に，本研究の位置付けについて述べる．

2.1 口の動きを検出する研究

口の動きを認識する手法はカメラを用いたものが多く存在する．カメラによって撮影した顔周辺の画像から口領域を割り出し，口の位置およびその形状を取得する．その画像から口の形状を推定する．Azh ら [2] は，モバイル端末の内臓カメラにより取得した画像から，口唇の形状に操作をそれぞれ割り当て，モバイル端末の操作を行なう手法を提案している．モバイル端末の日本語の文字入力手法として，Lyons ら [3] はカメラ画像により取得した口唇の形状から母音を推定し，キーで指定した子音と組み合わせることで日本語を入力する手法を提案している．澤田ら [4] は，発話時の口の形状から母音を推定し，モバイル端末におけるアプリケーションの起動操作に用いる手法を提案している．Chan ら [5] は，既存のヘッドホンにカメラを取り付けたヘッドマウント型のカメラを口の前方に配置し口の動きを検出している．検出した口の動きと手で操作するペンによって絵を描くシステムを提案している．

口自体の動き以外にも口に関連した口周辺の部位，特に舌の動きに注目した研究がある．Miyauchi ら [6] は，ダウン症候群の子供向け舌トレーニングの研究を行っており，Kinect を用いて深度と RGB データから顔領域および口領域を特定し，口から突出させた舌の動きを認識している．Crawford ら [7] は，Web カメラによって，舌の色とテクスチャの特徴から口の領域を特定した後に突出させた舌の動きをリアルタイムで認識している．

また，カメラ以外による口および口周辺の動きを認識する研究もある．Cheng ら [8] は，頬の外側に取り付けた布製の圧力センサアレイによって口内の舌の動きを認識し，その動きを入力に用いる手法を提案している．Sasaki ら [9] は，顎の下部の筋電位 (EMG) を，多チャンネル電極が顎全体を覆うように貼り付けることによって測定し，舌の動きを推定している．同様に筋電位を用いた手法として，Zhang ら [10, 11] は 6 つの筋電位センサを顎の下部に，2 つの筋電位センサを頬に取り付けて筋電位を測定し，舌の動きを認識している．Goel ら [12] は，既存のヘッドセットに X バンドモーション検出センサを取り付け，そのセンサを口の周囲 (前方および左右の計 3 箇所) に配置することで舌の動きを検出している．Li ら [13] は，

マイクロレーダセンサを口の周囲（前方および左右の計3箇所）に配置しドップラー効果を用いて舌の動きを認識している。

2.2 モバイル端末のハンズフリーな操作手法に関する研究

手以外の体の部位を使用してモバイル端末を操作するハンズフリーな入力手法は数多く存在する。Andoら[14]は、イヤホンに埋め込んだ気圧センサを用いて、イヤホンによって密閉された外耳道の気圧値の変化から顔に関連する動作（顎、顔、頭の動き）を推定し、モバイル端末のアプリケーション操作に使用する手法を提案している。Ronkainenら[1]は、モバイル端末をポケットに入れたまま端末をタップすることで入力とする手法について研究を行っている。Scottら[15]は、ズボンのポケット内のスマートフォンの加速度センサにより足の動きを推定し、スマートフォンの操作に利用した。Strohmeierら[16]は、衣服に取り付けた2つの曲げセンサによって手首の角度を取得し、モバイル端末の入力とする手法を提案している。Nagamatsuら[17]は、2つのカメラをモバイル端末に取り付け、それらのカメラで取得した三次元画像から視線の検出を行ない、モバイル端末の操作に利用している。モバイル端末で使用できる音声認識としては、Siri¹、Google Assistant²、Cortana³、Alexa⁴などがあり、非常に高精度で音声の認識および端末の操作を行なうことが可能である。

2.3 静電容量計測技術を用いた導電体検出を行なう研究

静電容量センシングの研究は、ヒューマン・コンピュータ・インタラクション（HCI）の研究において著しい役割を果たしてきた[18]。モバイル機器、ウェアラブル機器、据え置き機器など様々な機器において使用されている。

Dietzら[19]は、センサをアレイ状に配置することで複数人のタッチ位置やジェスチャを同時に検出するDiamondTouchを提案している。Hinckelyら[20]はマウスやトラックボールといった入力デバイスに静電容量タッチセンサを取り付け、ユーザのタッチ位置検出を行なう研究を行っている。Zimmermanら[21]は静電容量計測をHCIへ応用し、人体を検出するインタフェースに適用したシステムの開発を行っている。Rekimoto[22]は、面上に設置された網状の電極と導電体（人体）との静電容量変化を検出し、その面上の複数の手の位置をセンシングするSmartSkinを提案している。Satoら[23]は、タッチ状態によってインピーダンス周波数特性が変化することから、1つの電極のみによってユーザの把持状態を認識するTouchéを提案している。Tsurutaら[24]は、単一結線という単純性を保ちながら、多電極へのそれぞれのタッチを識別する静電容量方式タッチセンサであるRootCapを開発している。Wangら[25]は、車両に搭載されたタッチスクリーンへのタッチを、座席に組み込んだセンサとタッチスクリーン間の静電容量から運転手と搭乗者とを識別する手法を提案している。

¹<https://www.apple.com/jp/ios/siri/>

²https://assistant.google.com/intl/ja_jp/

³<https://www.microsoft.com/ja-jp/windows/cortana>

⁴<https://developer.amazon.com/ja/alexa>

2.4 本論文の位置付け

これらの研究に対して本研究では、口の動きをマスク型インタフェースにより推定する。口の動きを利用するため、手を使用しないハンズフリーな入力を行なうことができる。また、マスクを装着することにより口がマスクで覆われ、口の動きを周囲の人間が視認することができなくなる。そのため、周囲に人がいるような公共の場における使用に対する抵抗感を減少することができる。入力には口の動きを使用することから実際に発声を行なう必要はなく、音声認識のような周囲の雑音の影響を受けない利点がある。

口の動きを推定する方法として、Rekimoto[22]と同様の静電容量計測技術を応用した導電体検出による手法を用いる。センサ面における導電体の近接を複数の点において認識することで、導電体（口および口周辺）の形状を認識する。

第3章 マスク型インタフェースによるモバイル 端末操作

3.1 概要

提案手法における概要図を図 3.1 に示す。本手法では、マスク内にセンサを組み込んだマスク型インタフェースによって、口の形状を認識する。マスク内のセンサによりマスク内の各点における静電容量値を測定する。測定した静電容量値を用いて、モバイル端末はユーザの口の形状を認識し、端末の操作を行なう。複数の口の形状を認識することで、口の形状によってアプリケーションの動作に対応した操作を行なうことができる。この手法を用いることによりユーザは、手を使えない状況においても口を動かすことで、モバイル端末の操作を行なうことが可能になる。

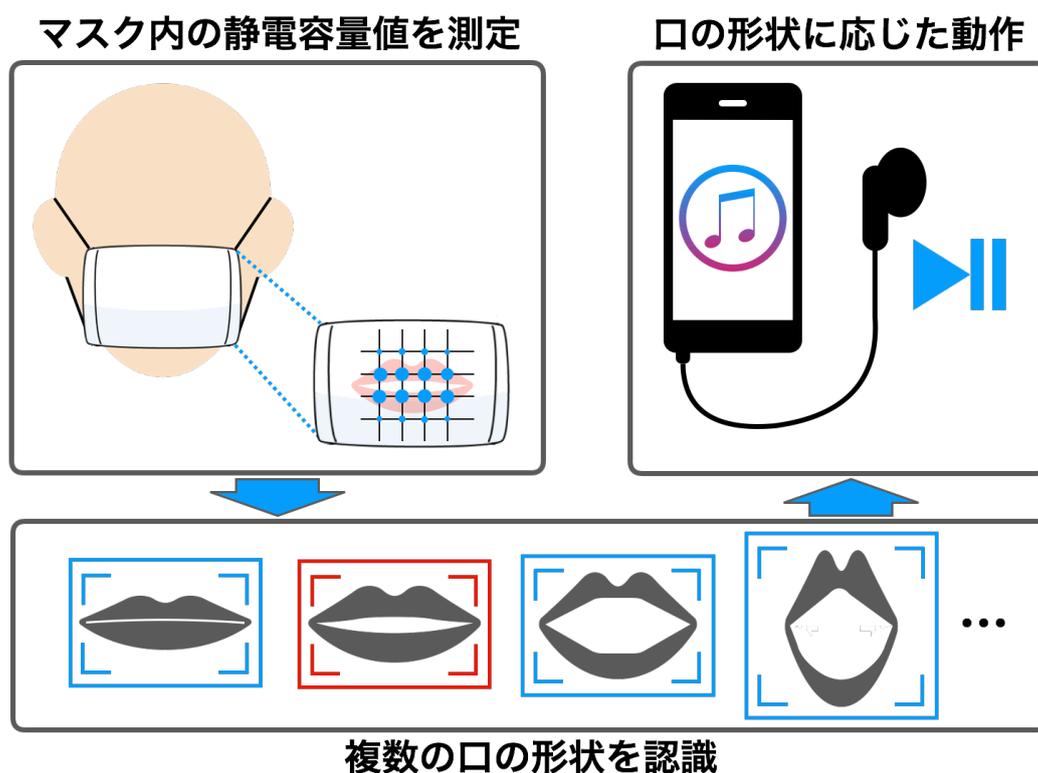


図 3.1: 提案手法の概要

3.2 ユーザによる使い方

本研究では，ユーザは日常的に公共の場においてマスクを装着するものとする．ユーザによる本手法の利用の流れは図 3.2 のようになり，以下にその説明を示す．

1. マスク型インタフェースを装着する．

マスク型インタフェースは，マスク内にセンサを挟み込むことで構成されるため，通常のマスクを装着するのと同様に装着することができる．

2. マスク内で口の形状を変化させ，複数の口の形状を登録する．

何もしていない状態，口を開いた状態，口（唇）を尖らせた状態，...とユーザに合った様々な口の形状を任意に登録することができる．モバイル端末でラベルを選択し，一定時間口の形状を固定することでその口の形状の登録を行なう．この作業を繰り返し行なうことで複数の口の形状の登録を行なう．

3. 口の形状ごとに，モバイル端末のアプリケーションで行ないたい操作を割り当てる．

口の形状とモバイル端末のアプリケーションの操作を口の形状ごとに割り当てる．例えば，口を開くと音楽アプリケーションで再生/停止を行なう．ユーザは自由に割り当てを行なうことができる．また，操作を割り当てないこともできる．

4. 口を動かすことでアプリケーションを操作する．

モバイル端末を手で操作することなく，口を動かすことでモバイル端末のアプリケーションの操作を行なう．手が使えない状況においてもモバイル端末を操作することができる．

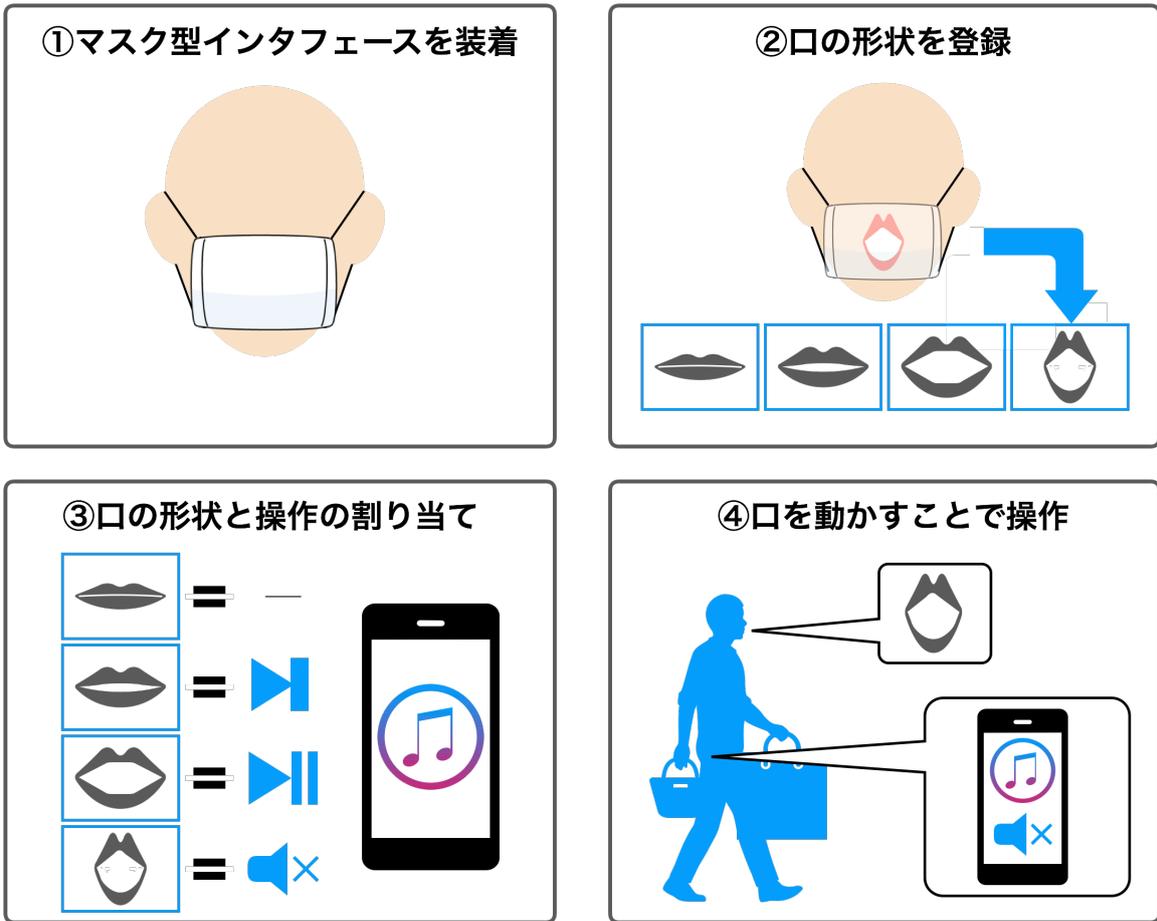


図 3.2: ユーザによる利用の流れ

3.3 利用シナリオ

3.3.1 音楽プレイヤーの操作

ユーザが音楽プレイヤーを操作する例を示す。混雑した電車内では端末を取り出すスペースがない、両手に荷物を持っている・つり革に掴まっているときは端末を手を取れない、といった状況が存在する。また、電車内は周囲に多数の人がいる公共の場であるため、声を出すことに抵抗を感じる問題もある。

音楽プレイヤーの操作に本手法を用いることで、口を大きく開くと音楽の再生/停止、口を左右にずらすことで早送り早戻し、口を尖らせると楽曲を次へ進むと割り当てる。ユーザは手を使うことなく、ハンズフリーに音楽プレイヤーの操作をすることができる。声を発することも必要ないため、操作を行なうことに抵抗を感じることもない。

3.3.2 料理ナビゲーションアプリの操作

ユーザがタブレット端末によって料理レシピを見ながら料理をする際に、料理ナビゲーションアプリを操作する例を示す。料理中は、手が濡れて端末を手で操作できないことがある。また、換気扇の動作音、食材を炒める、お湯が沸騰する、といった調理中の音が雑音となる。

料理レシピアプリの操作に本手法を用いることで、口を左右にずらすことで前後の調理工程へ移動、口を開くことでページの拡大に割り当てる。ユーザは料理をしながら料理レシピアプリの操作をすることができ、現在の状況に応じたナビゲーションを提示することができる。

3.3.3 パスワードの入力への応用

モバイル端末のロック解除や Web サイトのログイン時にパスワードが必要な場合がある。このパスワードを口の形状によって行なうことを想定する。

口はマスク型インタフェースによって覆われており、口の形状を外部から認識することは困難である。そのため、手の動きからキーボードへの入力を見てパスワード等の機密情報を盗み出すショルダーハッキングのような情報漏洩を防ぐことができるメリットがある。

第4章 実装

本章では、前章において説明した提案手法について実際に行なった実装の詳細について述べる。

4.1 マスク型インタフェース

本手法は、口の動きをマスク型インタフェースによって推定し、モバイル端末の操作を行なう手法である。インタフェースをマスク型とし、マスクによって口を覆うことで、口の動きを周囲に知られることなくモバイル端末への入力を行なうことが可能である。マスクとして、図 4.1 のような平型マスクを用いる。

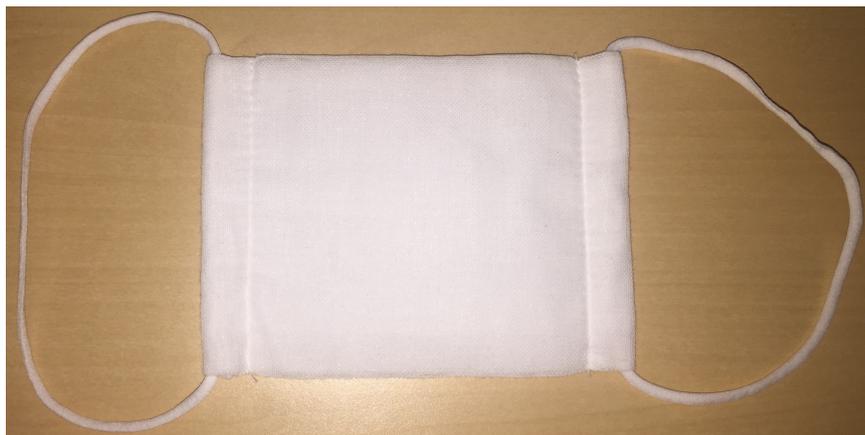


図 4.1: 平型マスクの外観

4.2 システムの構成

口の動きを認識するため、マスクの内部にタッチセンサを組み込む。本手法におけるシステムの構成を図 4.2 に示す。本システムは、高周波信号を発生する信号発生回路、格子状に配置された複数の送信電極および受信電極により作成したセンサ部、信号処理回路、マイクロコンピュータ（今回は Arduino Uno¹ を使用）により構成される。センサ部は縦方向と横方向

¹<https://store.arduino.cc/usa/arduino-uno-rev3>

それぞれに4本ずつの電線を配置し、計16個の交点をつくることで構成する。これらの交点は絶縁されており、コンデンサを構成する。センサ部の縦方向の電線が送信電極、横方向の電線が受信電極である。マスクおよびセンサ部を安価に作成することができ、使い捨て可能なインタフェースとして使用できる利点があることから、このセンサ部の電線を導電系で構成する。

このシステムの信号発生回路を図4.3に、信号処理回路を図4.4に示す。信号の発生はファンクションジェネレータICにより行なう。発生する信号の周波数 $F[Hz]$ はICに接続する抵抗の値 $R[\]$ およびコンデンサの値 $C[F]$ により決定され、以下の値となる。

$$F = \frac{1}{RC} \quad (4.1)$$

4.3 センサ部における静電容量計測

信号発生回路により約200kHzの正弦波を発生させ、その高周波信号を送信電極へ送信する。信号は送信電極と受信電極の交点を通じて受信電極へ伝わる。このときの交流信号の強度は、送信電極と受信電極による交点の静電容量に比例する。この交点に接地した導電体が接近すると、その導電体と送信電極・受信電極が静電結合し、信号の一部がその導電体にも流れる。その結果、受信電極に送信される交流信号の強度が弱まる。この変化を受信電極で計測することで、導電体の接近を推定することができる。受信電極へ伝わった交流信号は、信号処理回路において、オペアンプによって増幅されダイオードで整流した後、ローパスフィルタにかけ直流の値へ変換する。

また、生成した交流信号を送信電極へ送信する際、複数の電極へマルチプレクサを用いて時分割で送信する。同様に、受信電極についてもマルチプレクサによって複数の電極から交流信号を受け取り、信号処理回路へ接続される。マルチプレクサは共にマイクロコンピュータによって制御する。

今回の実装においては、送信電極、受信電極ともに4つでセンサ部を構成するため、交点は16点となる。これによって取得した複数の値を用いて、各交点における導電体の接近を推定し、それらの値を二次元平面にプロットすることでセンサ部全体における導電体の形状を推定することができる。

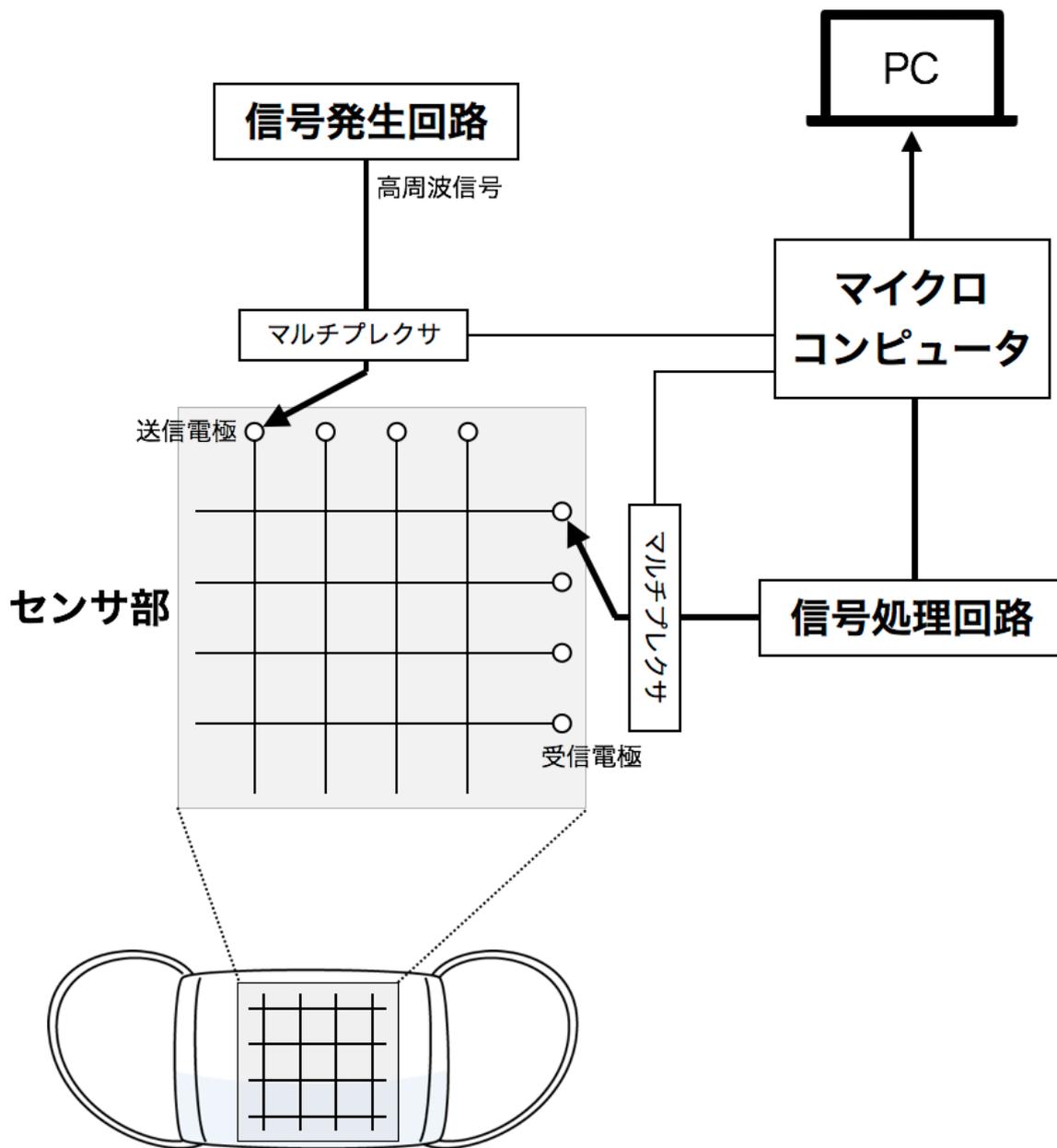


図 4.2: システムの構成．信号発生回路，センサ部，信号処理回路，マイクロコンピュータによって構成される．

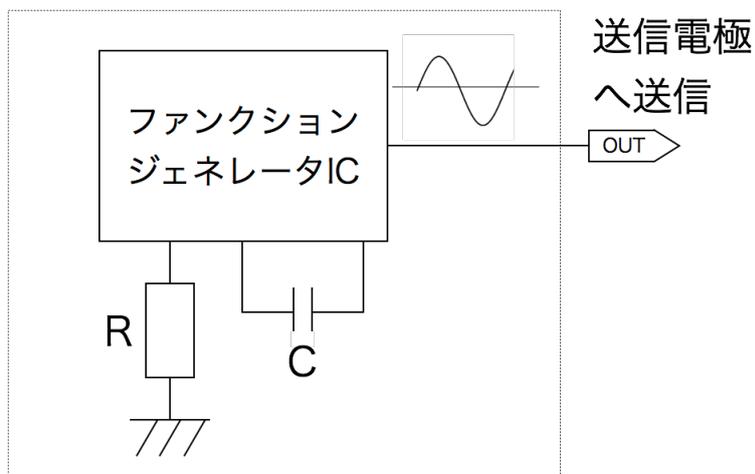


図 4.3: 信号発生回路

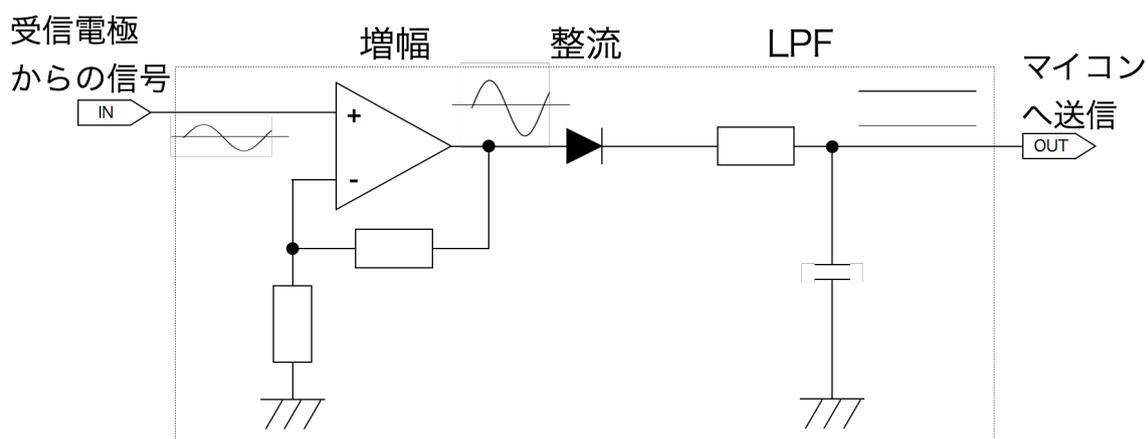


図 4.4: 信号処理回路

4.4 口の形状認識

センサ部に接地した導電体が接近することで、その導電体の形状の推定が可能である。人体は、電極に対して十分に大きい物体のため、明確に接地をしなくとも接地した導電体と同様の振る舞いをする。前述したセンサ部を平型マスクに挟み込み、その状態でマスクを装着することで、人体の一部である口の形状認識に用いる。

マスクを装着したときの口の形状によって、センサ部の送信電極と受信電極の各交点に対する人体（口および口周辺）の距離が異なるため、静電容量値が変化する。各交点の静電容量値に比例した半径の円を描画したイメージ図を図 4.5 に示す。この各交点における静電容量値の分布は口の形状ごとに異なる。

口の形状を推定するために、機械学習 (SVM: サポートベクターマシン) を用いた。データとして正規化した交点 16 点における静電容量値を使用する。まず事前に複数の口の形状を、静電容量値 16 点の値を学習データとして学習させる。その後、識別するときは現在の口の形状の静電容量値 16 点から、口の形状を推定する。

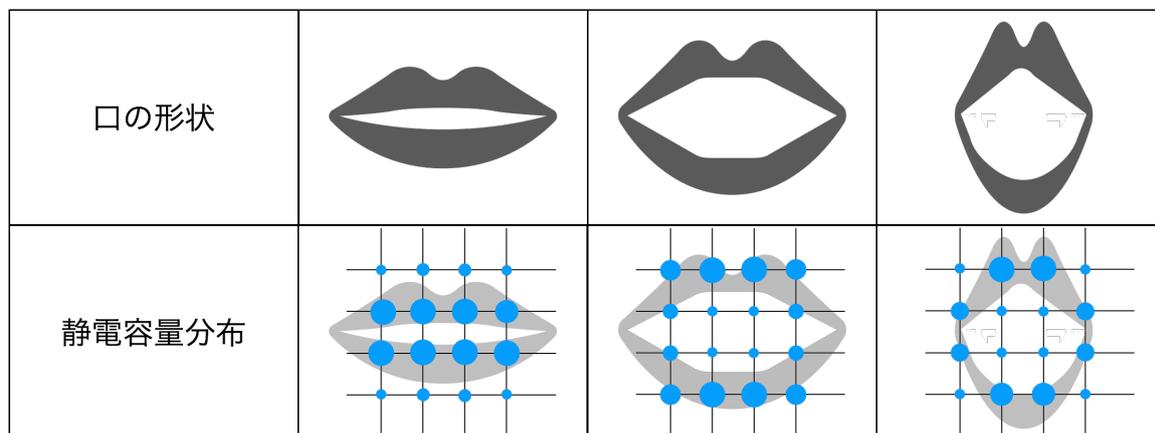


図 4.5: 口の形状と静電容量分布のイメージ図

4.5 ブレッドボード上での実装

実際にブレッドボード上に実装したシステムの全体図を図 4.6 に示す。システムは、信号発生回路、センサ部、信号処理回路、マイクロコンピュータから構成されている。実装した信号発生回路の回路図を図 4.7、信号処理回路の回路図を図 4.8、マルチプレクサの制御および信号処理回路からの信号を受信するマイクロコンピュータの回路図を図 4.9、導電糸を用いて作製したセンサ部を図 4.10 に示す。

信号発生回路で約 200kHz の正弦波を発生させ、センサ部の複数の送信電極へ送る。この際、信号は Arduino によって制御されるマルチプレクサへ一度送信され、1 つの送信電極と接続される。その後、信号は送信電極と受信電極の交点を通り受信電極へ伝わる。受信電極も Arduino によって制御されるマルチプレクサと接続されており、1 つの受信電極と信号処理回路を接続する。信号処理回路で処理された信号は、Arduino の analog ピンによって受信されその値を Arduino が読み取る。この動作を時分割で、信号発生回路と送信電極の接続、受信電極と信号処理回路の接続をそれぞれ変更していくことで、計 16 点存在する送信電極と受信電極の各交点に対する静電容量値を取得する。

4.6 口の形状識別アプリケーション

図 4.11 はセンサ部の静電容量分布の表示および口の形状認識を行なう GUI アプリケーションである。アプリケーションの開発言語には Python を使用した。口の形状の識別は SVM (サ

ポートベクターマシン)によって行なう。実装として LIBSVM[26] を使用した。このアプリケーションについて以下で説明する。

(1) ではシリアル通信によって受け取った、Arduino の analog ピンによって読み取った 16 点分の静電容量値を表示する。静電容量値は、信号処理回路によって 0V から 5V の電圧値に変換した後、その値を Arduino によって 0 から 1023 のデジタル値に変換して読み取る。その後、そのデジタル値を 1000 分の 1 の値 (0.000 から 1.023) に変換してシリアル送信する。

(2) は (1) と同様にシリアル通信で受け取ったマスク内のセンサの交点 16 点の静電容量値に比例した半径を持つ円をそれぞれ表示するエリアである。口の形状によって各点の静電容量値が変化するため、その値に比例して円のサイズも変化し、視覚的にその変化を確認することができる。

(3) では口の形状の登録および認識を行なう。最上部にあるドロップダウンリストに識別するラベルの一覧が入る。その下にあるテキストボックスに任意の文字列を入力し append ボタンを押すことで、ドロップダウンリストに任意の文字列がラベルとして追加される。その下にある Delete ボタンを押すと、ボタンを押したときにドロップダウンリストで選択されていたラベルが削除される。また、Reset ボタンを押すとドロップダウンリスト内に存在する全てのラベルが削除される。そして、Record ボタンを押すと、ボタンを押したときにドロップダウンリストで選択されていたラベル名の文字列 (図 4.11 では“あ”) を使用して「あ_train.csv」という csv ファイルが出力され、そのときの交点 16 点の静電容量値のデータ 100 フレーム分を保存する。Train ボタンを押すと出力した csv ファイルを読み込み、そのデータを元に教師データを作成し、(1) に表示される現在における交点 16 点の静電容量値を用いて識別を行なう。例えば、“default”、“open”、“sharpen” の 3 つのラベルのデータが csv ファイルに出力されていた場合、「default_train.csv」、「open_train.csv」、「sharpen_train.csv」を読み込み、それら 3 つの教師データを作成した後、現在の静電容量値を用いて識別を行なう。

(4) に識別結果を表示する。Train ボタンを押すまでは「result」と表示されており、識別が始まると文字列が識別した結果のラベルへと変化する。

最後に、(5) はコンソールである。(3) で行なった動作を表示する。例えばドロップダウンリストに“default”ラベルを append ボタンを押して追加したときは、このスペースに「default appended.」と表示される。

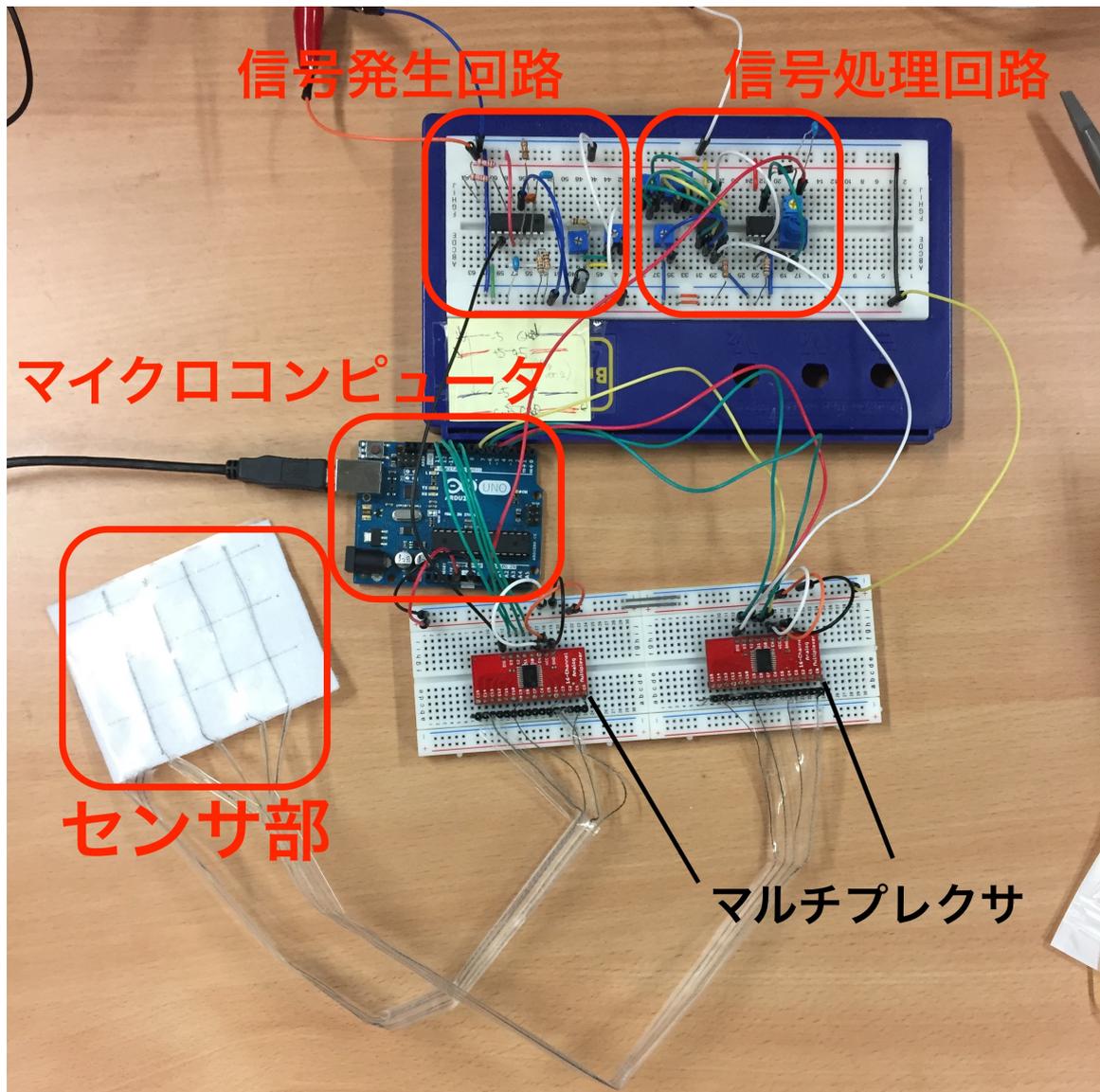


図 4.6: 実装したシステムの全体図

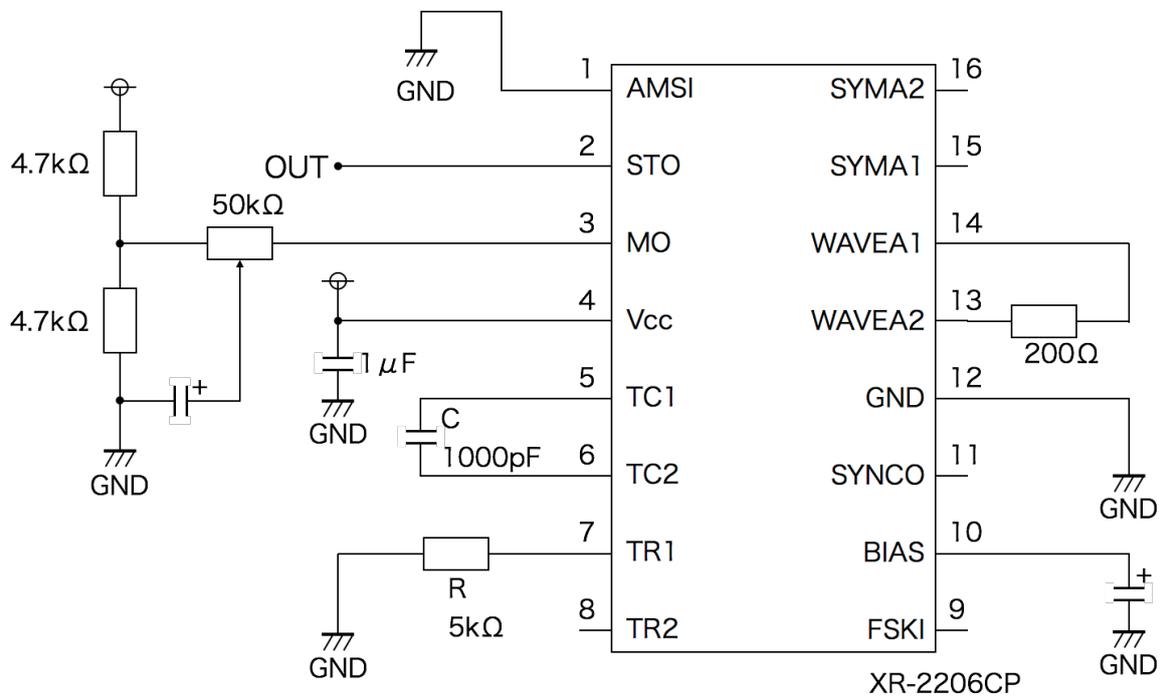


図 4.7: 信号発生回路の回路図

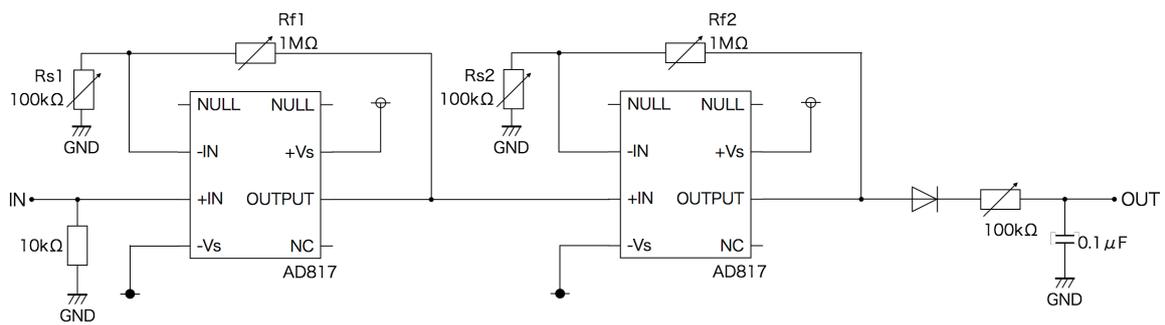


図 4.8: 信号処理回路の回路図

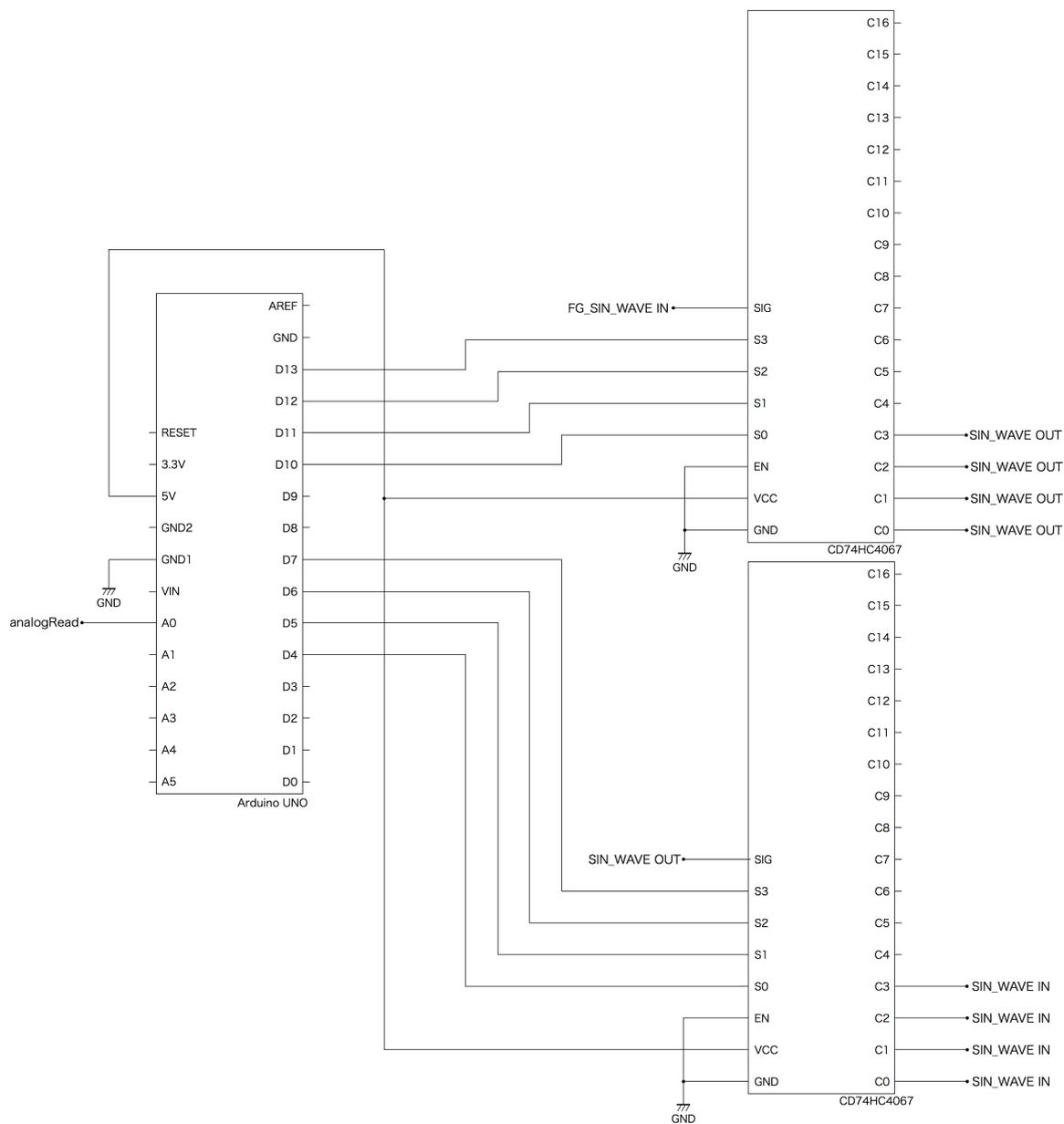


図 4.9: マイクロコンピュータの回路図 . マイクロコンピュータはマルチプレクサの制御および信号処理回路からの信号を受信する .

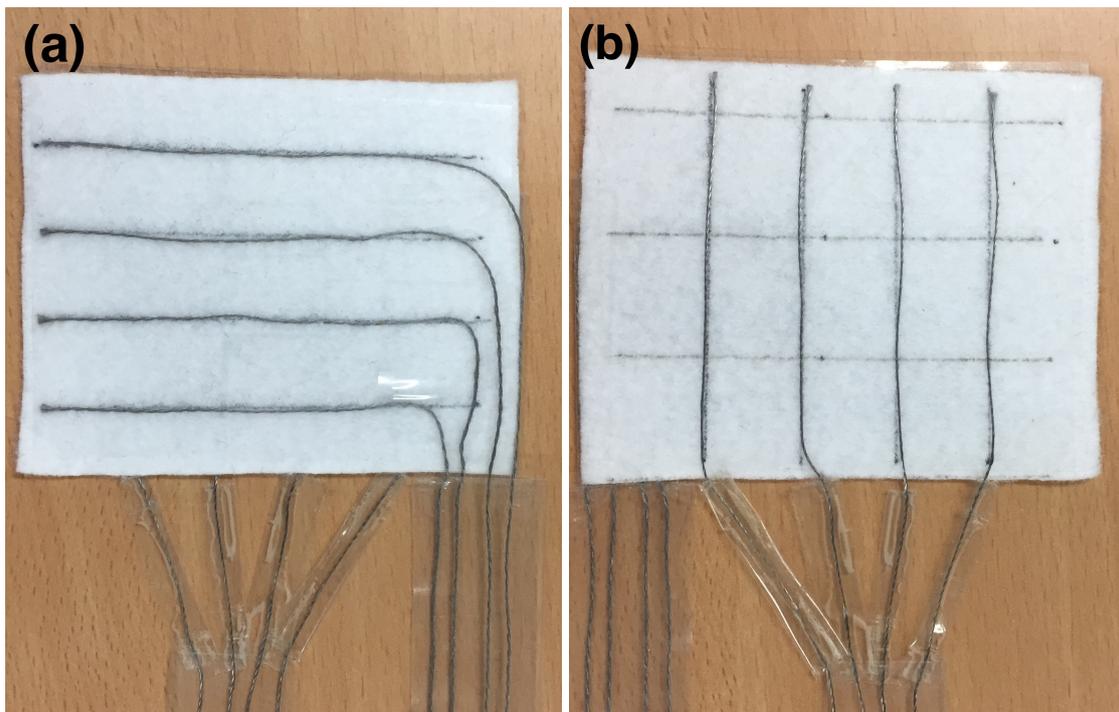


図 4.10: 作製したセンサ部．電線を導電糸によって構成し，導電糸が短絡しないように表面と裏面にそれぞれ電極を設けた．(a) 表面：送信電極，(b) 裏面：受信電極．

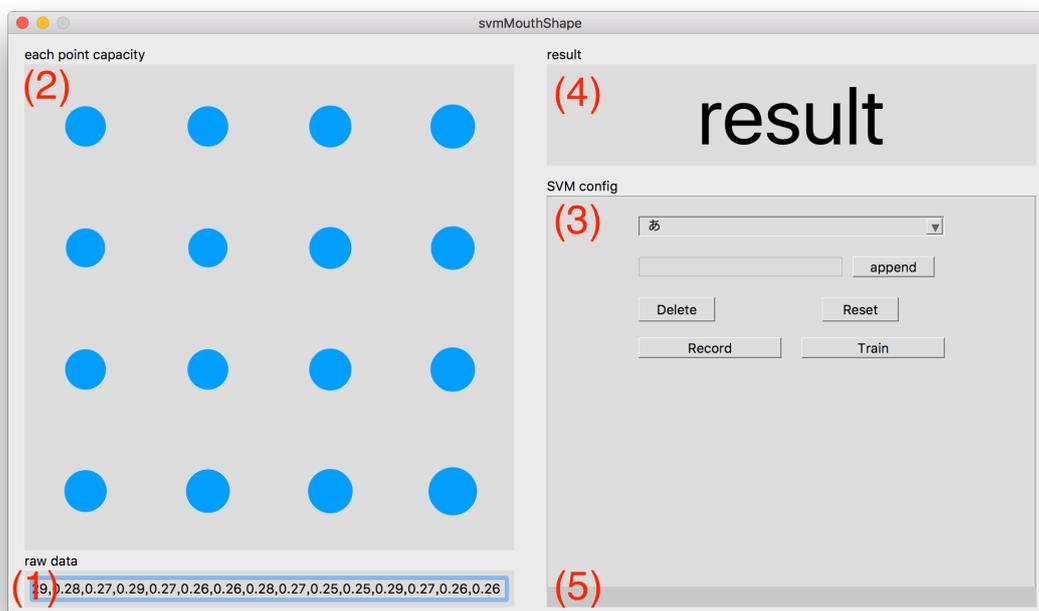


図 4.11: 静電容量分布表示および口の形状認識を行なうアプリケーション

第5章 評価

実装したマスク型インタフェースによる口の動きの認識を確認するための実験を行なった．実験には，前章で説明した GUI アプリケーションを使用した．

5.1 被験者

著者 1 名が被験者として実験を行なった．

5.2 実験内容

まず実験 1 として図 5.1 のような 3 つの口の形状を識別する実験を行なう．図 5.1 におけるラベル名は GUI アプリケーションで表示されるラベル名と同じである．default は何もしない状態，open は口を開いた状態，sharpen は口（唇）を尖らせた状態である．

次に実験 2 として，識別するラベルを 1 つ増やし，計 4 つの口の形状を識別する実験を行なう．追加する口の形状を図 5.2 に示す．bite は，上下の両方の唇を噛んだ状態である．実験方法はラベルを 1 つ増やし，実験 1 と同様に行なう．

5.3 実験の手順

GUI アプリケーションによる実験の手順は以下の通りである．実験 1 においてラベルの順番は default , open , sharpen の順とし，実験 2 においての順番は default , open , sharpen , bite の順とする．また，口の形状を学習する際の学習データ数は口の形状ごとに 100 フレーム分とする．

1. マスクを装着する．
2. ラベル名を決定しラベルを登録する．
3. 口の形状をラベルごとに学習する（実験 1 は 3 つ，実験 2 は 4 つ）．
4. 口の形状を 1 つめのラベルに変化させる．
5. 口の形状をその状態で固定し，口の形状の識別を 20 回行なう．
6. 口の形状を次のラベルに変化させる．

7. 5. と 6. を繰り返し、最後のラベルまで行なう。

8. 4. から 7. を 5 回繰り返す。

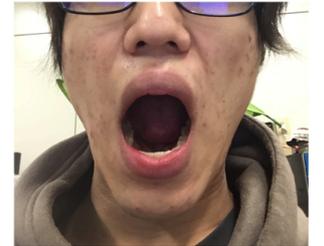
ラベル名	口の形状	マスク装着時の外観
default		
open		
sharpen		

図 5.1: 実験 1 での口の形状

ラベル名	口の形状	マスク装着時の外観
bite		

図 5.2: 実験 2 で追加した口の形状

5.4 実験の結果および考察

実験1の結果を表5.1にまとめる．同様に実験2の結果を表5.2にまとめる．表5.1,表5.2の見方は, 行のラベルの口の形状を行なったときに, 列の識別となった割合および正答率を表している．例えば, 図5.1では, 1回目の default と口の形状を変化させたとき, default と20回中19回識別され, open と20回中0回識別され (open とは識別されず), sharpen と20回中1回識別され, 結果として正答率が20回中19回の95%となったことを示す．

図5.1より, 実験1ではラベル3種類での識別を行ない, 1, 3, 4回目の default の識別のみ sharpen と誤認識されている．この結果から, 高い認識率を得ることができたといえる．

次に図5.2より, 実験2ではラベル4種類での識別を行なった．open および sharpen については, 正答率が100%であり, 誤認識されることはなかった．しかし, default および bite については正答率が著しく低く, default については bite と誤認識され, bite については default と誤認識されていることが多いことがわかる．このことから, default の口の形状と bite の口の形状が似ており, 識別が困難であったと考えられる．

また, 図5.1と5.2の結果から, 混同行列を作成した．その結果をそれぞれ図5.3と図5.4に示す．行のラベルの口の形状を行なったときに, 列の識別となった割合を表している．各ラベルの値は5回分を1つにまとめたものとなっている．

ここで, 実験で使用した GUI アプリケーションで表示した各点における default, open, sharpen, bite の静電容量値の一例をそれぞれ図5.5, 図5.6, 図5.7, 図5.8に示す．これらの図から, default と bite の静電容量値はどちらも16点が同じような値となっていたため, 識別が困難であったと考えられる．また, open の静電容量値は中央の4つの点が他の点より小さくなっており, 識別が精度よく行なわれたと考える．sharpen については図から目視で特徴を見出すことが困難であったが識別は精度よく行なわれていた．

実験2において default と bite の正答率が低く正しい識別ができなかった．この理由は, この2つの口の形状の特徴量 (交点16点の静電容量値) が似ていたためである．よって, 特徴量の違いの大きい口の形状であれば識別することが可能であると考えられる．これはこの2つ以外の口の形状 (open と sharpen) および実験1の結果を見れば明らかである．口の形状によっては4つ以上の識別ができる可能性もある．今後は今回行なった口の形状以外の口の形状についても実験を行ない, 識別可能な口の形状数の調査や口の形状ごとの特徴量の違いの調査を行なう必要がある．

表 5.1: 実験 1 の結果

回数	ラベル	default	open	sharpen	正答率 [%]
1	default	19/20	0/20	1/20	95
1	open	0/20	20/20	0/20	100
1	sharpen	0/20	0/20	20/20	100
2	default	20/20	0/20	0/20	100
2	open	0/20	20/20	0/20	100
2	sharpen	0/20	0/20	20/20	100
3	default	19/20	0/20	1/20	95
3	open	0/20	20/20	0/20	100
3	sharpen	0/20	0/20	20/20	100
4	default	19/20	0/20	1/20	95
4	open	0/20	20/20	0/20	100
4	sharpen	0/20	0/20	20/20	100
5	default	20/20	0/20	0/20	100
5	open	0/20	20/20	0/20	100
5	sharpen	0/20	0/20	20/20	100

表 5.2: 実験 2 の結果

回数	ラベル	default	open	sharpen	bite	正答率 [%]
1	default	8/20	0/20	0/20	12/20	40
1	open	0/20	20/20	0/20	0/20	100
1	sharpen	0/20	0/20	20/20	0/20	100
1	bite	13/20	0/20	1/20	6/20	30
2	default	9/20	0/20	0/20	11/20	45
2	open	0/20	20/20	0/20	0/20	100
2	sharpen	0/20	0/20	20/20	0/20	100
2	bite	7/20	0/20	1/20	12/20	60
3	default	10/20	0/20	0/20	10/20	50
3	open	0/20	20/20	0/20	0/20	100
3	sharpen	0/20	0/20	20/20	0/20	100
3	bite	9/20	0/20	0/20	11/20	55
4	default	9/20	0/20	1/20	10/20	45
4	open	0/20	20/20	0/20	0/20	100
4	sharpen	0/20	0/20	20/20	0/20	100
4	bite	10/20	0/20	0/20	10/20	50
5	default	9/20	0/20	0/20	11/20	45
5	open	0/20	20/20	0/20	0/20	100
5	sharpen	0/20	0/20	20/20	0/20	100
5	bite	8/20	0/20	0/20	12/20	60

	default	open	sharpen
default	0.97	0.00	0.03
open	0.00	1.00	0.00
sharpen	0.00	0.00	1.00

図 5.3: 実験 1 の混同行列

	default	open	sharpen	bite
default	0.45	0.00	0.01	0.54
open	0.00	1.00	0.00	0.00
sharpen	0.00	0.00	1.00	0.00
bite	0.47	0.00	0.02	0.51

図 5.4: 実験 2 の混同行列

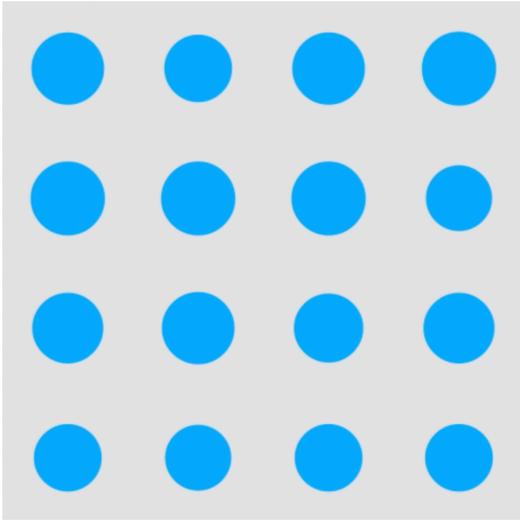


図 5.5: default の静電容量値の一例

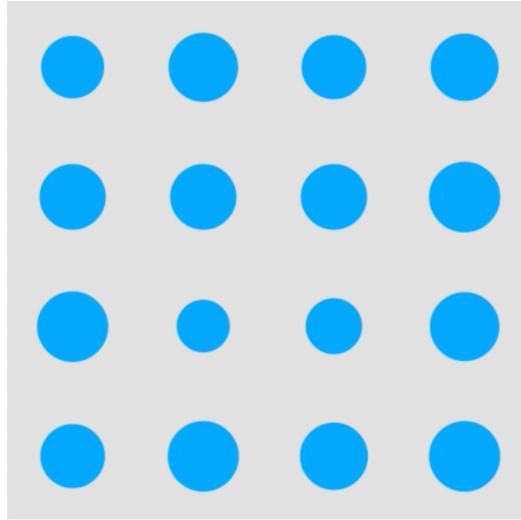


図 5.6: open の静電容量値の一例

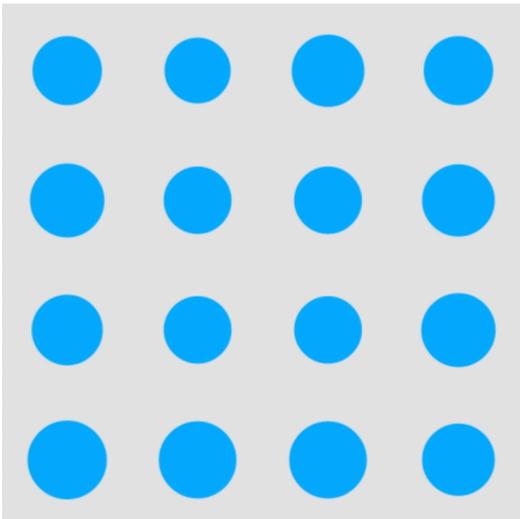


図 5.7: sharpen の静電容量値の一例

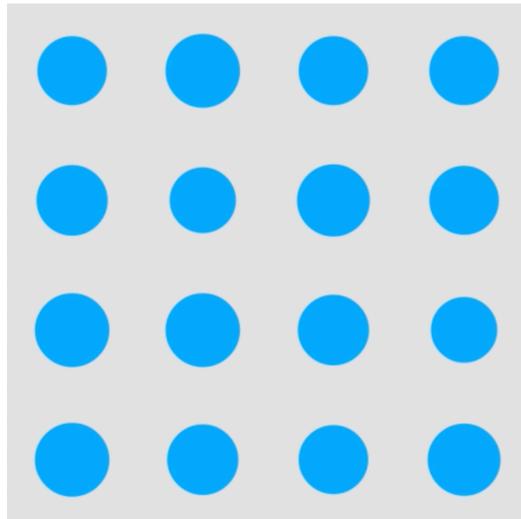


図 5.8: bite の静電容量値の一例

第6章 結論

本研究では満員電車内のような手が使えない状況におけるモバイル端末への新たな入力手法として、マスク型インタフェースにより口の動きを検出する手法を提案した。既存のハンズフリーな入力手法では、特に音声入力においては周囲の雑音の影響や公共の場における使用に対する抵抗感がある問題があった。本手法では、マスク型のインタフェースとすることで、ウェアラブルかつハンズフリーな使用ができ、口の動きをマスクによって隠すことができる。口をマスクで覆うことで、周囲に人がいるような公共の場においても抵抗なく使用することができ、また発声する必要もないことから周囲の雑音の影響を受けることもない。また、デバイスに使用するマスク・導電糸は安価なため使い捨てのインタフェースとして手軽に使えることも利点である。本研究では、マスク内に静電容量方式の格子状のセンサを取り付けることでマスク型インタフェースを実装した。

実装したマスク型インタフェースを用いて口の形状を識別する実験を行なったところ、3種類の識別は高精度で行なうことができたが、4種類の識別は精度が低い結果となった。しかし、識別精度が低い理由としては特徴量が似た口の形状の識別を試みたためと考えられるため、4種類以上の識別が行えることが示唆された。

今後の課題として、今回の実験での口の形状だけでなく、その他の口の形状についても実験を行ない、正しい識別を行なうことができるかを調査することが挙げられる。そのほかにも識別の精度を向上すること、識別数を増加することが挙げられる。また、今後は本手法を利用したモバイル端末の操作を行なうアプリケーションの開発を行なう予定である。

謝辞

本研究を進めるにあたり，指導教員である高橋伸准教授には多くのご指導およびご助言を頂きました．また，志築文太郎准教授，嵯峨智准教授，Simona Vasilache 助教には研究室内の全体ゼミなどを通して大変貴重な意見を頂きました．ここに深く御礼申し上げます．

インタラクティブ研究室の皆様には日頃の研究活動から様々なアドバイスを頂きました．特に，Saga チームの富田氏からはハードウェア面における多くの知識や示唆を頂きました．また特に，同 Ubiquitous チームの高木氏からはソフトウェア面におけるアドバイスやプログラミング知識の提供を頂きました．ここに感謝の意を表します．

最後に，自分の生活を支えてくださった家族，日常生活や研究生活において支えてくださった皆様に深く感謝いたします．

参考文献

- [1] Sami Ronkainen, Jonna Häkkinen, Saana Kaleva, Ashley Colley, and Jukka Linjama. Tap Input as an Embedded Interaction Method for Mobile Devices. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, TEI '07, pp. 263–270, New York, NY, USA, 2007. ACM.
- [2] Maryam Azh and Shengdong Zhao. LUI: Lip in Multimodal Mobile GUI Interaction. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, ICMI '12, pp. 551–554, New York, NY, USA, 2012. ACM.
- [3] Michael J. Lyons, Chi-Ho Chan, and Nobuji Tetsutani. MouthType: Text Entry by Hand and Mouth. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '04, pp. 1383–1386, New York, NY, USA, 2004. ACM.
- [4] 澤田佳樹, 高橋伸, 田中二郎. 母音形状を用いたモバイル端末向けコマンド入力手法. 情報処理学会インタラクシオン 2016, pp. 1038–1041, 2016.
- [5] Chi-ho Chan, Michael J. Lyons, and Nobuji Tetsutani. Mouthbrush: Drawing and Painting by Hand and Mouth. In *Proceedings of the 5th International Conference on Multimodal Interfaces*, ICMI '03, pp. 277–280, New York, NY, USA, 2003. ACM.
- [6] Masato Miyauchi, Takashi Kimura, and Takuya Nojima. A Tongue Training System for Children with Down Syndrome. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, pp. 373–376, New York, NY, USA, 2013. ACM.
- [7] Chris S. Crawford, Stephen W. Bailey, Carmen Badea, and Juan E. Gilbert. Using Cr-Y Components to Detect Tongue Protrusion Gestures. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, pp. 1331–1336, New York, NY, USA, 2015. ACM.
- [8] Jingyuan Cheng, Ayano Okoso, Kai Kunze, Niels Henze, Albrecht Schmidt, Paul Lukowicz, and Koichi Kise. On the Tip of My Tongue: A Non-Invasive Pressure-Based Tongue Interface. In *Proceedings of the 5th Augmented Human International Conference*, AH '14, pp. 12:1–12:4, New York, NY, USA, 2014. ACM.

- [9] Makoto Sasaki, Kohei Onishi, Stefanov Dimitar, Katsuhiko Kamata, Atsushi Nakayama, Masahiro Yoshikawa, and Goro Obinata. Tongue interface based on surface EMG signals of suprahyoid muscles. In *ROBOMECH Journal 2016*, pp. 3–9, 2016.
- [10] Qiao Zhang, Shyamnath Gollakota, Ben Taskar, and Raj P.N. Rao. Non-Intrusive Tongue Machine Interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 2555–2558, New York, NY, USA, 2014. ACM.
- [11] Qiao Zhang, Shyamnath Gollakota, Ben Taskar, and Raj P.N. Rao. Non-Intrusive Tongue Machine Interface. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pp. 2555–2558, New York, NY, USA, 2014. ACM.
- [12] Mayank Goel, Chen Zhao, Ruth Vinisha, and Shwetak N. Patel. Tongue-in-Cheek: Using Wireless Signals to Enable Non-Intrusive and Flexible Facial Gestures Detection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pp. 255–258, New York, NY, USA, 2015. ACM.
- [13] Zheng Li, Ryan Robucci, Nilanjan Banerjee, and Chintan Patel. Tongue-n-Cheek: Non-contact Tongue Gesture Recognition. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*, IPSN '15, pp. 95–105, New York, NY, USA, 2015. ACM.
- [14] Toshiyuki Ando, Yuki Kubo, Buntarou Shizuki, and Shin Takahashi. CanalSense: Face-Related Movement Recognition System Based on Sensing Air Pressure in Ear Canals. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pp. 679–689, New York, NY, USA, 2017. ACM.
- [15] Jeremy Scott, David Dearman, Koji Yatani, and Khai N. Truong. Sensing Foot Gestures from the Pocket. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pp. 199–208, New York, NY, USA, 2010. ACM.
- [16] Paul Strohmeier, Roel Vertegaal, and Audrey Girouard. With a Flick of the Wrist: Stretch Sensors as Lightweight Input for Mobile Devices. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, TEI '12, pp. 307–308, New York, NY, USA, 2012. ACM.
- [17] Takashi Nagamatsu, Michiya Yamamoto, and Hiroshi Sato. MobiGaze: Development of a Gaze Interface for Handheld Mobile Devices. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '10, pp. 3349–3354, New York, NY, USA, 2010. ACM.
- [18] Tobias Grosse-Puppenthal, Christian Holz, Gabe Cohn, Raphael Wimmer, Oskar Bechtold, Steve Hodges, Matthew S. Reynolds, and Joshua R. Smith. Finding Common Ground: A

- Survey of Capacitive Sensing in Human-Computer Interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 3293–3315, New York, NY, USA, 2017. ACM.
- [19] Paul Dietz and Darren Leigh. DiamondTouch: A Multi-User Touch Technology. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, UIST '01, pp. 219–226, New York, NY, USA, 2001. ACM.
- [20] Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. Passive Real-World Interface Props for Neurosurgical Visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '94, pp. 452–458, New York, NY, USA, 1994. ACM.
- [21] Thomas G. Zimmerman, Joshua R. Smith, Joseph A. Paradiso, David Allport, and Neil Gershenfeld. Applying Electric Field Sensing to Human-Computer Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, pp. 280–287, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.
- [22] Jun Rekimoto. SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, pp. 113–120, New York, NY, USA, 2002. ACM.
- [23] Munehiko Sato, Ivan Poupyrev, and Chris Harrison. Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pp. 483–492, New York, NY, USA, 2012. ACM.
- [24] Masaya Tsuruta, Shuta Nakamae, and Buntarou Shizuki. Rootcap: Touch detection on multi-electrodes using single-line connected capacitive sensing. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*, ISS '16, pp. 23–32, New York, NY, USA, 2016. ACM.
- [25] Edward Jay Wang, Jake Garrison, Eric Whitmire, Mayank Goel, and Shwetak Patel. Carpacio: Repurposing Capacitive Sensors to Distinguish Driver and Passenger Touches on In-Vehicle Screens. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pp. 49–55, New York, NY, USA, 2017. ACM.
- [26] Chih-Chung Chang and Chih-Jen Lin. LIBSVM – A Library for Support Vector Machines. <https://www.csie.ntu.edu.tw/~cjlin/libsvm/> (accessed 2018-1-6).