

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

# スポイトメタファに基づくタッチパネル操作

堀 竜慈

(コンピュータサイエンス専攻)

指導教員 田中 二郎

2012年03月

## 概要

我々は、我々の先行研究において、スポイトメタファに基づくタッチパネル操作（以下スポイト操作）を提案した。スポイト操作とは、タッチパネル面に触れた指に同じ手の親指を付け、その親指によりタッチパネル面に触れた指を摩擦しながら押し下げる動作または引き上げる動作を用いた操作である。従来のタッチパネル操作に、我々が提案するスポイト操作を追加することにより、タッチパネル面において、押し出すまたは吸い上げる操作感を得る操作、ある一点に対する奥行き方向の操作感を得る操作、ドラッグアンドドロップ操作中の命令への割り当てに適した操作を実現等の応用が可能になる。我々の先行研究において、タッチパネル面に伝わる、スポイト操作による固体音を、隠れマルコフモデルを用いて認識することにより、スポイト操作を認識可能であることがわかった。

本研究では、タッチパネル面に伝わる、スポイト操作による固体音を、別手法を用いて認識することにより、スポイト操作をより高精度、かつ低レイテンシな認識が可能となる。これについて述べる。また、スポイト操作を利用したアプリケーションを作成し、実際に使用し、考察したため、その考察について述べる。

# 目次

|              |                                                            |           |
|--------------|------------------------------------------------------------|-----------|
| <b>第 1 章</b> | <b>はじめに</b>                                                | <b>1</b>  |
| <b>第 2 章</b> | <b>スポイト操作</b>                                              | <b>4</b>  |
| 2.1          | スポイト操作とは . . . . .                                         | 4         |
| 2.2          | スポイト操作の特徴 . . . . .                                        | 5         |
| 2.2.1        | 押し出すまたは吸い上げる操作感 . . . . .                                  | 5         |
| 2.2.2        | 奥行き方向の操作感 . . . . .                                        | 5         |
| 2.2.3        | タッチパネル平面上の操作から独立 . . . . .                                 | 5         |
| 2.2.4        | 指の摩擦移動距離をパラメータに利用可能 . . . . .                              | 5         |
| 2.3          | スポイト操作の応用例 . . . . .                                       | 5         |
| 2.3.1        | データを取り込む、取り出す命令への割り当て . . . . .                            | 7         |
| 2.3.2        | ある一点に対する奥行き方向の命令への割り当て . . . . .                           | 7         |
| 2.3.3        | ドラッグアンドドロップ中の操作への割り当て . . . . .                            | 9         |
| <b>第 3 章</b> | <b>スポイト操作の認識</b>                                           | <b>10</b> |
| 3.1          | スポイト操作認識のための基本アイデア . . . . .                               | 10        |
| 3.2          | スポイト操作認識実験にあたってのシステム構成 . . . . .                           | 11        |
| 3.2.1        | ハードウェア . . . . .                                           | 11        |
| 3.2.2        | ソフトウェア . . . . .                                           | 13        |
| 3.3          | 実験 1: 認識アルゴリズムに SVM を用いた認識実験 . . . . .                     | 13        |
| 3.3.1        | 手法 . . . . .                                               | 15        |
|              | 固体音の FFT . . . . .                                         | 15        |
|              | SVM 認識モデルの生成 . . . . .                                     | 15        |
|              | SVM 認識モデルを用いた認識 . . . . .                                  | 15        |
| 3.3.2        | 条件 . . . . .                                               | 15        |
| 3.3.3        | 結果 . . . . .                                               | 16        |
| 3.3.4        | 考察 . . . . .                                               | 17        |
| 3.4          | 実験 2: 認識アルゴリズムに spectral centroid の最小二乗法を用いた認識実験 . . . . . | 18        |
| 3.4.1        | 手法 . . . . .                                               | 18        |
| 3.4.2        | 条件 . . . . .                                               | 21        |
| 3.4.3        | 結果 . . . . .                                               | 21        |
| 3.4.4        | 考察 . . . . .                                               | 22        |

|            |                                 |           |
|------------|---------------------------------|-----------|
| <b>第4章</b> | <b>スポイト操作を用いたアプリケーション</b>       | <b>23</b> |
| 4.1        | システム構成 . . . . .                | 23        |
| 4.2        | ドローアプリケーションの機能説明 . . . . .      | 27        |
| 4.3        | ドローアプリケーションを使用した結果 . . . . .    | 27        |
| <b>第5章</b> | <b>関連研究</b>                     | <b>28</b> |
| 5.1        | 固体音を用いた操作 . . . . .             | 28        |
| 5.2        | スポイトメタファを用いた操作 . . . . .        | 29        |
| 5.3        | タッチパネルにおける奥行き操作感を得る操作 . . . . . | 29        |
| 5.4        | ドラッグアンドドロップ中の操作 . . . . .       | 29        |
| 5.5        | タッチパネル操作への操作の追加 . . . . .       | 30        |
| <b>第6章</b> | <b>まとめと今後の課題</b>                | <b>31</b> |
|            | <b>謝辞</b>                       | <b>32</b> |
|            | <b>参考文献</b>                     | <b>33</b> |

# 目次

|      |                                                                                         |    |
|------|-----------------------------------------------------------------------------------------|----|
| 1.1  | 従来のタッチパネル操作-左：タップ操作、右：ドラッグ操作 . . . . .                                                  | 1  |
| 1.2  | マルチタッチ操作-上：回転操作、左下：ピンチイン操作、右下：ピンチアウト<br>操作 . . . . .                                    | 2  |
| 2.1  | スポイト操作説明-a:push out の動作, b:pump up の動作 . . . . .                                        | 4  |
| 2.2  | スポイトを用いて液体を押し出している様子 -a の状態においてスポイトの尾<br>部を押すことにより、スポイトの中の液体が押し出されb の状態になる . . .        | 6  |
| 2.3  | データを取り込む、取り出す命令へ割り当てられたスポイト操作のイメージ .                                                    | 7  |
| 2.4  | スポイト操作を用いた 3D 空間オブジェクト移動操作のイメージ . . . . .                                               | 8  |
| 2.5  | モデリングツールに用いられるマルチビューの一例 . . . . .                                                       | 8  |
| 3.1  | 固体音の伝搬 . . . . .                                                                        | 10 |
| 3.2  | 認識実験装置側面 . . . . .                                                                      | 11 |
| 3.3  | 認識実験装置上面 . . . . .                                                                      | 12 |
| 3.4  | コンクリートマイク . . . . .                                                                     | 12 |
| 3.5  | オーディオインタフェース . . . . .                                                                  | 13 |
| 3.6  | FFT 窓をずらしながら FFT を行い、各 FFT 窓ごとにスペクトルデータを生成<br>する . . . . .                              | 14 |
| 3.7  | pump up により生じる固体音のスペクトログラム（矢印区間に pump up のラベ<br>リングを行った） . . . . .                      | 16 |
| 3.8  | push out により生じる固体音のスペクトログラム（矢印区間に push out のラベ<br>リングを行った） . . . . .                    | 17 |
| 3.9  | 正解ラベルと認識結果の違い- 始端部分と終端部分の違い . . . . .                                                   | 17 |
| 3.10 | 正解ラベルと認識結果の違い-よくある誤認識パターン . . . . .                                                     | 18 |
| 3.11 | push out スペクトログラム . . . . .                                                             | 19 |
| 3.12 | pump up スペクトログラム . . . . .                                                              | 19 |
| 3.13 | pump up スペクトログラムから一定振幅以上の音声の連続を切り出している様子                                                | 19 |
| 3.14 | 切り出した pump up スペクトログラムの spectral centroid を各フレームごとに<br>求めている様子 . . . . .                | 20 |
| 3.15 | 求めた pump up スペクトログラムの spectral centroid に 1 次方程式の最小二乗<br>法を適用し、近似式の傾きを求めている様子 . . . . . | 20 |

|     |                                             |    |
|-----|---------------------------------------------|----|
| 4.1 | アプリケーションシステム構成 . . . . .                    | 23 |
| 4.2 | タブレット PC にコンクリートマイクを取り付けた様子 . . . . .       | 24 |
| 4.3 | タブレット PC にコンクリートマイクを取り付けた様子 (側面図) . . . . . | 25 |
| 4.4 | Sumsung 社のタブレット端末 GALAXY Tab . . . . .      | 25 |
| 4.5 | ドローアプリケーションでの pump up . . . . .             | 26 |
| 4.6 | ドローアプリケーションでの push out . . . . .            | 26 |
| 4.7 | スポイト操作を用いたドローアプリケーションを実際に使用している様子 . .       | 27 |

## 第1章 はじめに



図 1.1: 従来のタッチパネル操作-左：タップ操作、右：ドラッグ操作

近年、手や指を用いるタッチパネル操作（以後はタッチパネル操作）に新たな操作を追加する研究、開発が行われている。これにより、図 1.1 に示す、タップ操作やドラッグ操作等の従来のタッチパネル操作に加えて、新しい操作が実現されてきている。実用化されている例としては、図 1.2 に示す、ピンチイン操作、ピンチアウト操作、回転操作等のマルチタッチがあげられる。また、タッチパネルへの指の近接を用いた操作の実現 [18]、指を傾ける動作を用いた操作の実現 [9] といった研究も為されている。

これらに対し我々は以前に、タッチパネルへの追加操作としてスポイト操作に基づくタッチパネル操作（以下、スポイト操作）を示した [16]。スポイト操作とは、タッチパネル面に触れた指に同じ手の親指を当て、その親指によりタッチパネル面に触れた指を摩擦しながら押し下げる動作または引き上げる動作を用いた操作である。従来のタッチパネル操作に我々が提案するスポイト操作を追加することにより、タッチパネル面において、押し出すまたは吸い上げる操作感を得る操作、ある一点に対する奥行き方向の操作感を得る操作、ドラッグアンドドロップ操作中の命令の割り当てに適した操作を実現する。また、スポイト操作の認識方法として、タッチパネルを伝わる操作音（以下固体音）を隠れマルコフモデルを用いて解析する方法を試みた。

今回、我々は、タッチパネルへ固体音として伝わるスポイト操作の摩擦音を高速フーリエ

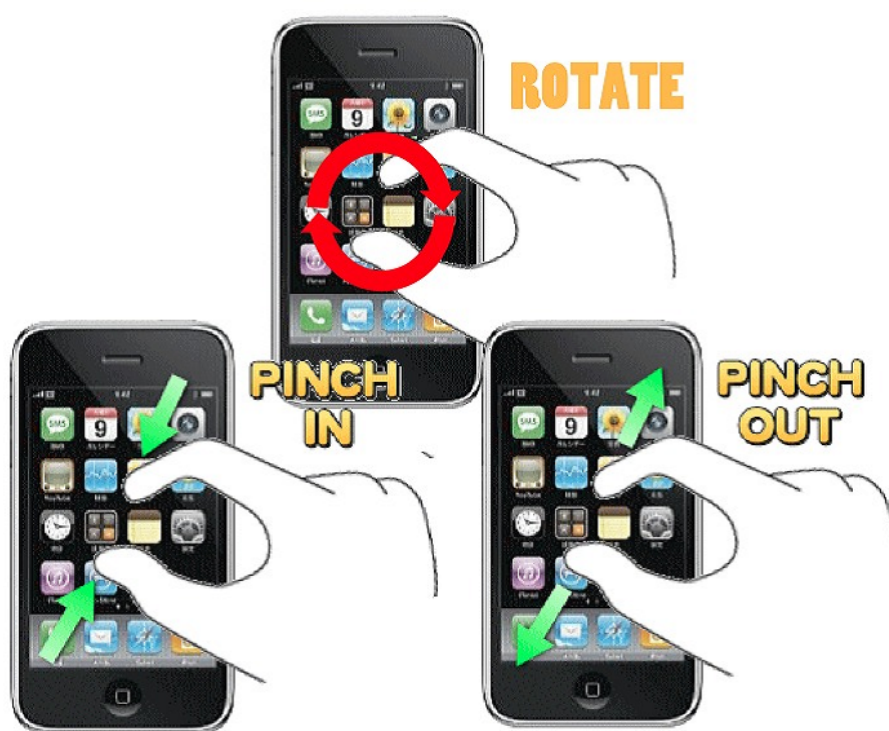


図 1.2: マルチタッチ操作-上：回転操作、左下：ピンチイン操作、右下：ピンチアウト操作

変換（以下 FFT）し、その結果を support vector machine（SVM）を用いて解析することにより、スポイト操作を実時間認識する手法を検討した。また、タッチパネルへ伝わる固体音の spectral centroid に最小二乗法を用いることにより認識する手法も検討した。

本論文ではスポイト操作の概略およびその応用例、2つの認識手法の詳細、および認識結果について述べる。また、作成したスポイト操作を用いた簡易ドローアプリケーションについても述べる。

本論文の構成を以下に示す。第2章で本研究において提案するスポイト操作を述べる。第3章では、スポイト操作の認識を述べる。第4章では、スポイト操作を用いたアプリケーションを述べる。第5章では、本研究に関連する研究と、それに基づく本研究の位置づけを述べる。そして第6章では、本論文の結論と今後の課題を述べる。

## 第2章 スポイト操作

本性では、スポイト操作、その特徴及び応用を述べる。

### 2.1 スポイト操作とは

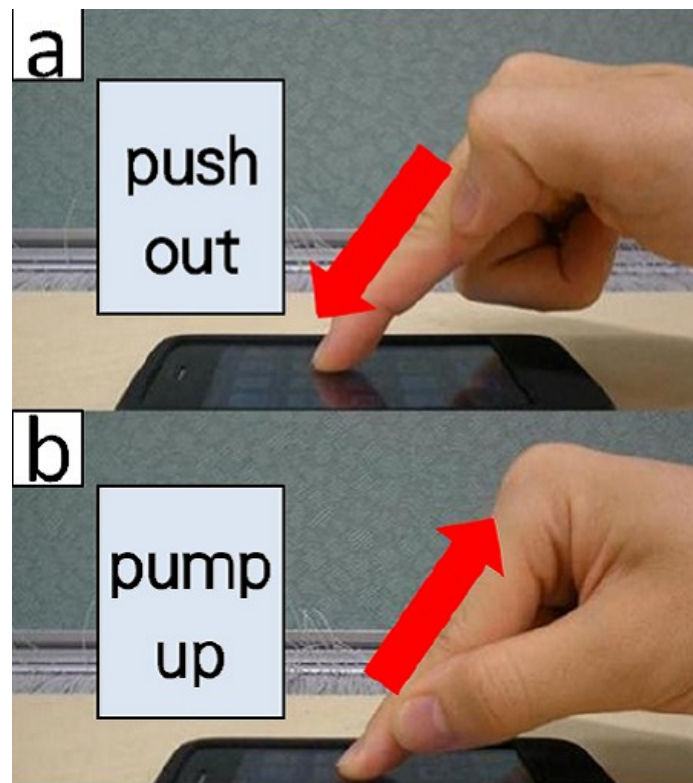


図 2.1: スポイト操作説明-a:push out の動作, b:pump up の動作

スポイト操作は、押し出し操作（以下 push out）及び吸い上げ操作（以下 pump up）の総称である。図 2.1a のように、タッチパネル面に触れた指に、同じ手の親指を当てた状態で、親指を指先方向に摩擦する動作を push out と呼ぶ。また、図 2.1b のように、タッチパネル面に触れた指に、同じ手の親指を当てた状態で、親指を指の根本方向に摩擦する動作を pump up と呼ぶ。

## 2.2 スポイト操作の特徴

本節では、スポイト操作の特徴について述べる。

### 2.2.1 押し出すまたは吸い上げる操作感

実際のスポイトを用いて液体を押し出すまたは吸い上げる際は、図 2.1 に示すようにスポイトの尾部をつまむ操作を用いる。一方、本研究が提案するスポイト操作は、スポイトにより押し出されているまたは吸い上げられている液面の動きを模倣したものである。図 2.1 に示した push out は、図 2.2 に示すようにスポイトの液体が押し出されていく際の液体の動きと同じであり、pump up は、スポイトの液体が吸い上げられていく際の液体の動きと同じである。このため、使用者はスポイトを使用時に得られる、押し出すまたは吸い上げる操作感を得る。

### 2.2.2 奥行き方向の操作感

push out 時、操作に用いている親指は、タッチパネル面に近づく。また、pump up 時、その親指はタッチパネル面から遠ざかる。すなわち、使用者はタッチパネル面に対し奥行き方向に親指を動かす。このため、使用者は操作時に、タッチパネル面のある一点に対する奥行き方向の操作感を得る。

### 2.2.3 タッチパネル平面上の操作から独立

スポイト操作に用いる動作は、指を対象とする動作である。そのため、タッチパネル平面を動作の対象とする操作とは独立しており、その操作に割り当てられた命令とは別の命令を同時に実行、あるいは単独で実行することが可能である。

### 2.2.4 指の摩擦移動距離をパラメータに利用可能

スポイト操作時の親指の摩擦移動距離という連続値をパラメータとして命令に用いることができる。

## 2.3 スポイト操作の応用例

スポイト操作の特徴を用いることにより、以下のようなスポイト操作の応用例が考えられる。

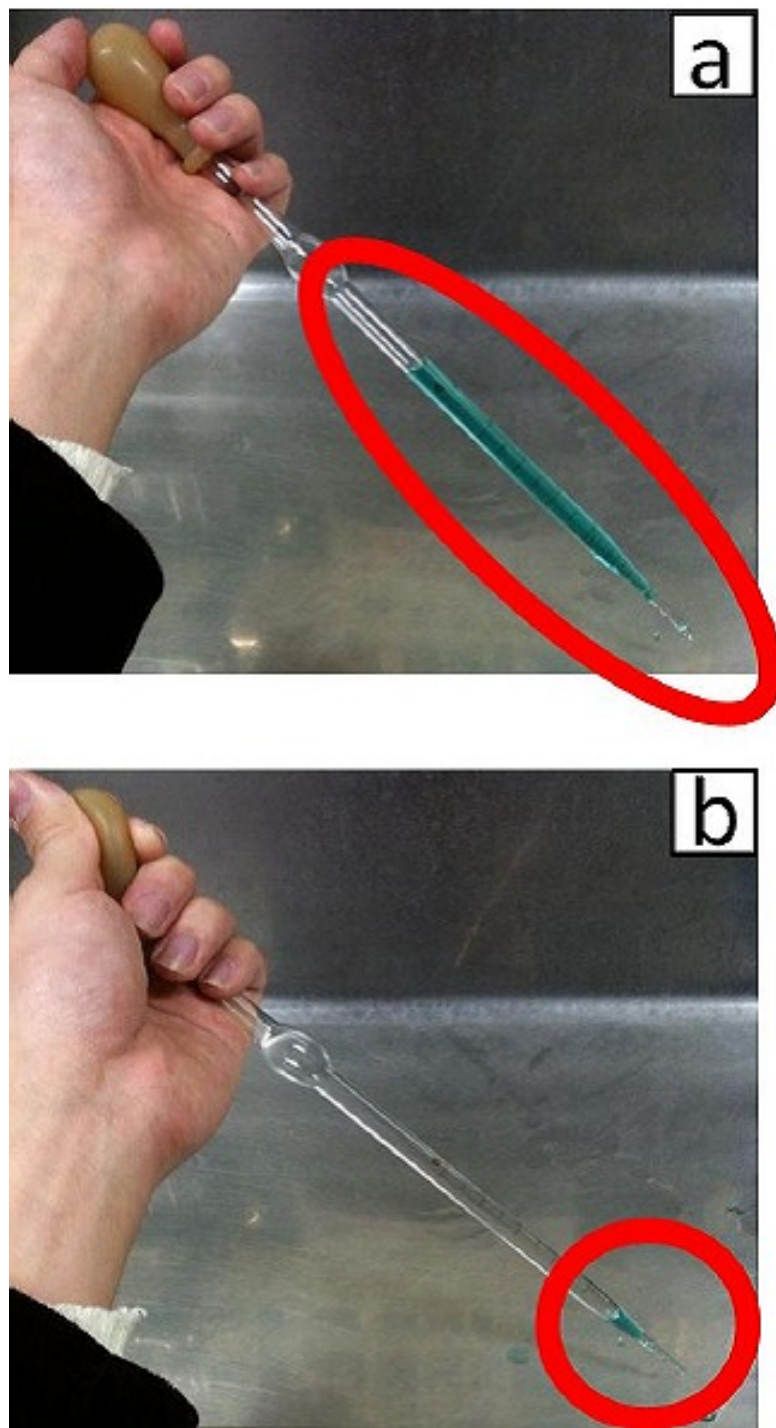


図 2.2: スポイトを用いて液体を押し出している様子 -a の状態においてスポイトの尾部を押すことにより、スポイトの中の液体が押し出され b の状態になる

### 2.3.1 データを取り込む、取り出す命令への割り当て

押し出すまたは吸い上げる操作感を活かす応用として、pump up をデータを取り込む命令、push out をデータをデータを取り出す命令に割り当てることが考えられる。具体例として、図 2.3 に示すように、カット命令に pump up を、ペースト命令に push out を割り当てるという応用が考えられる。また、画像編集ソフトのスポイトツールを用いた色の取得命令に pump up を割り当てると、という応用も考えられる。これにより、長押し操作やコンテキストメニューを用いて行なう必要があった命令を、より簡易かつ直感的なスポイト操作を用いて行うことにより、スムーズに一連の操作を行うことが可能となる。特に、データを取り込む、取り出す命令を頻繁に行う際に効果的だと考えられる。

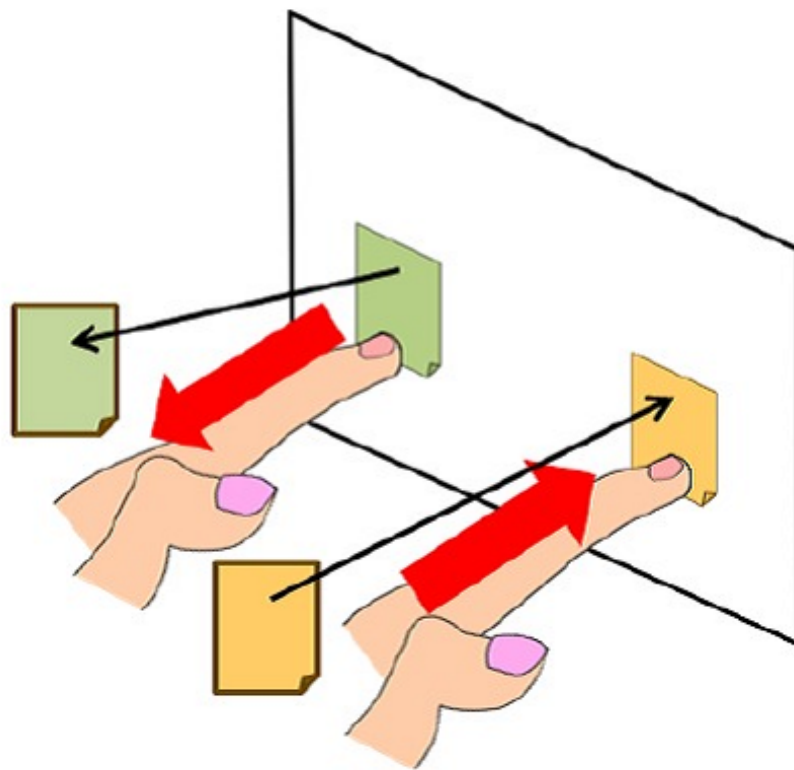


図 2.3: データを取り込む、取り出す命令へ割り当てられたスポイト操作のイメージ

### 2.3.2 ある一点に対する奥行き方向の命令への割り当て

奥行き方向の操作感を活かす応用として、3D 空間におけるオブジェクト操作に活用することができると考えられる。図 2.4 にスポイト操作を 3D 空間のオブジェクト移動操作に適用

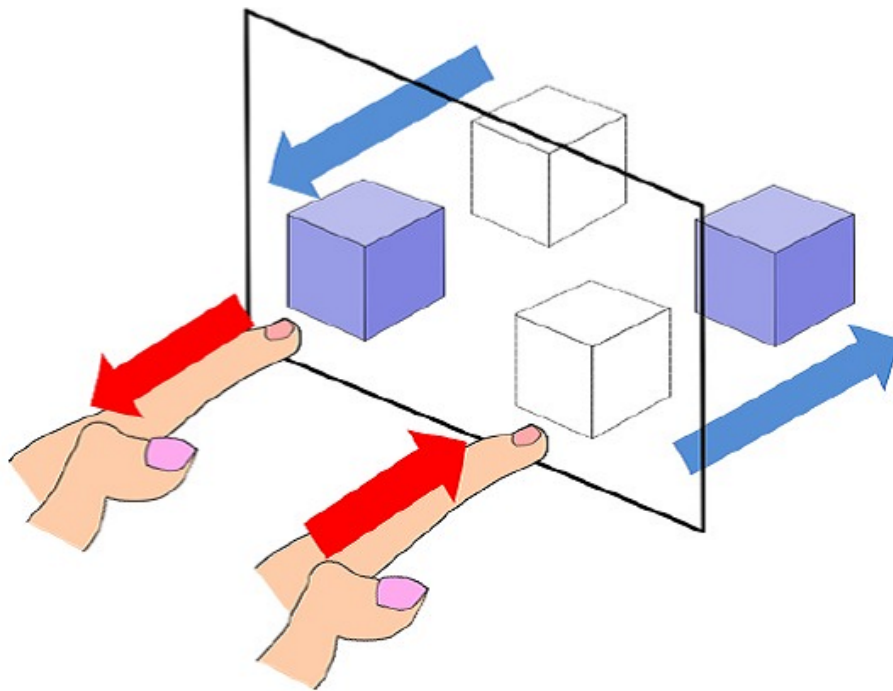


図 2.4: スポイト操作を用いた 3D 空間オブジェクト移動操作のイメージ

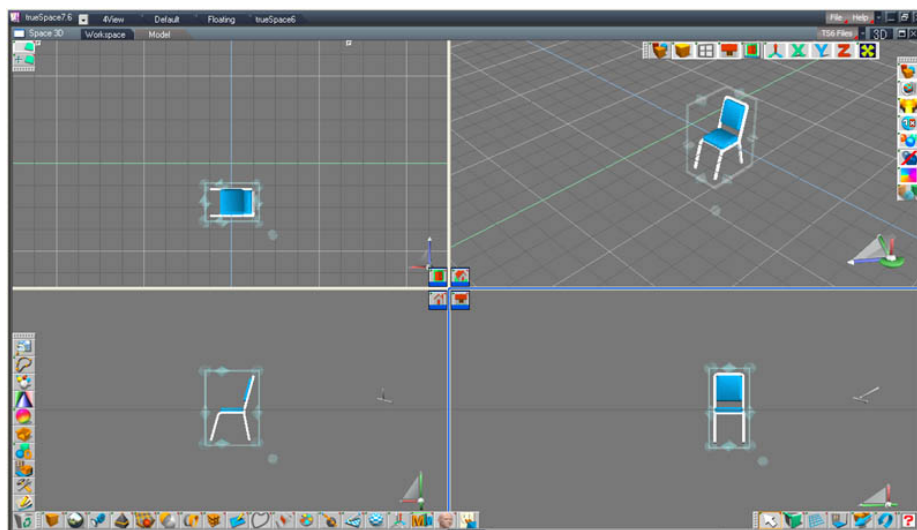


図 2.5: モデリングツールに用いられるマルチビューの一例

したイメージを示す。具体的には、push out を、指定オブジェクトを奥側に移動させる命令、pump up を、指定オブジェクトを手前側に移動させる命令に割り当てることができる。これにより、3D 空間の 3 軸を平面から操作する際、図 2.5 に示すようなマルチビューを用いずに、一つのビューのみで操作することが可能となる。図 2.5 に用いられている 3D モデリングソフトウェアは Microsoft 社<sup>1</sup>の Caligari trueSpace である。

### 2.3.3 ドラッグアンドドロップ中の操作への割り当て

タッチパネル平面上の操作から独立している特徴を活かす応用として、ドラッグアンドドロップ中に行えると便利な機能の実行にスポイト操作を割り当てることが考えられる。具体例として、ウィンドウのスクロールが考えられる。使用者が pump up または push out を行うと、操作の摩擦移動距離に応じてウィンドウを上方向または下方向にスクロールさせる。また、デスクトップ環境におけるフォルダの階層移動に割り当てることも考えられる。具体的には、pump up または push out を完了すると、現在のフォルダ階層から上位層または選択しているフォルダの層へ移動する。これにより、ドラッグアンドドロップを途中中断することなく、目的の命令を遂行することができるため、スムーズな操作が可能となる。

---

<sup>1</sup><http://www.microsoft.com>

## 第3章 スポイト操作の認識

本性では、スポイト操作を認識する各種法と、それぞれの手法の特徴、認識結果について述べる。

### 3.1 スポイト操作認識のための基本アイデア

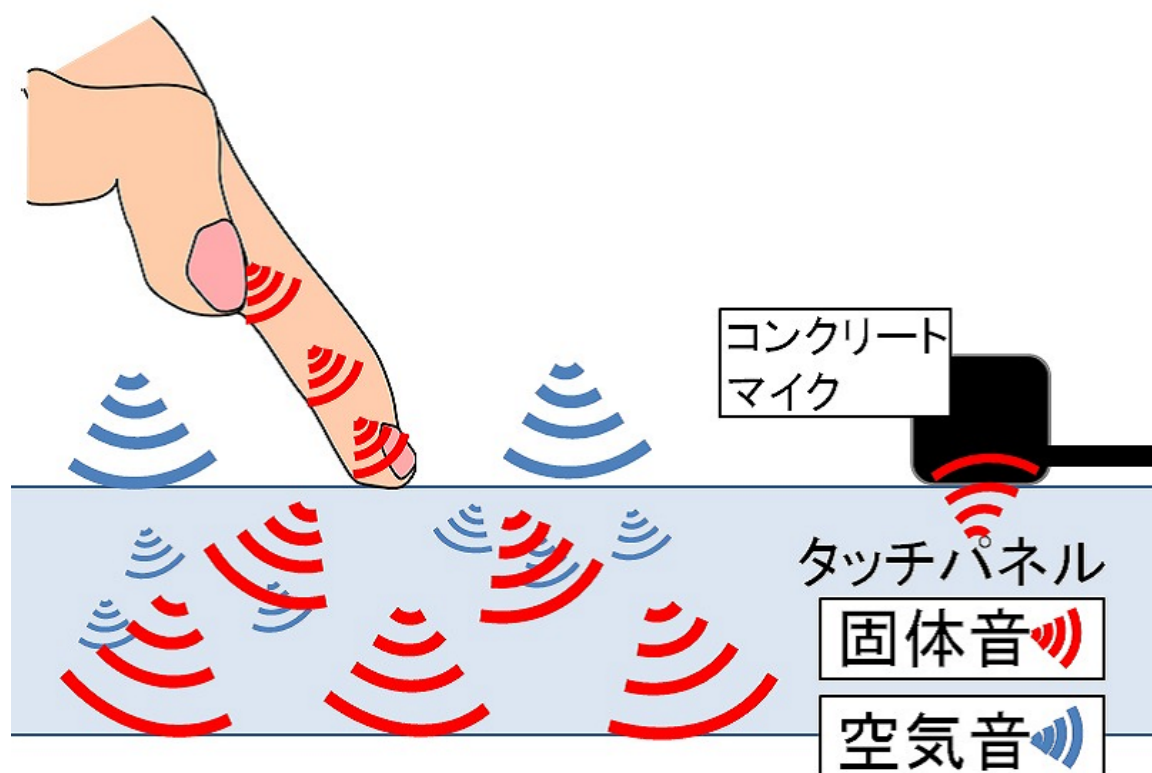


図 3.1: 固体音の伝搬

図 3.1 に示すように、スポイト操作を行うと、スポイト操作による固体音がタッチパネルに伝わる。我々は、この固体音を取得し、スポイト操作の認識に用いる。固体音を操作の認識に用いた理由を以下に挙げる。

- 固体は気体にくらべて音の伝導率が良いため、指のわずかな摩擦音でも容易に取得することができる。
- push out による固体音と、pump up による固体音の違いを用いて、スポイト操作の認識が可能である。
- スポイト操作の認識をタッチパネル側において行えるため、使用者はデバイス等を装着する必要がない。
- 携帯情報端末は内蔵のマイクを所持しているため、特別なセンサ類を携帯端末に取り付けることなく携帯端末においてスポイト操作を実現することができる。

## 3.2 スポイト操作認識実験にあたってのシステム構成

### 3.2.1 ハードウェア

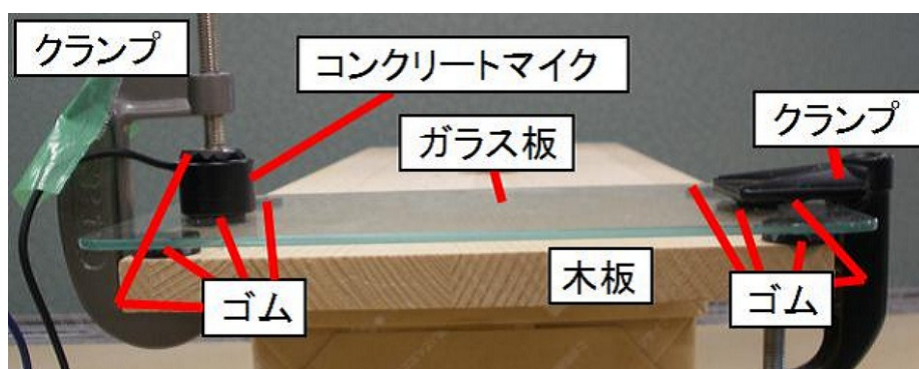


図 3.2: 認識実験装置側面

スポイト操作を認識することが可能かを確かめるために、図 3.2、図 3.3 に示す装置を製作した。タッチパネルの代わりとして、タッチパネルに使われているものと同様の透明なガラス板（横:230mm、縦:132mm、厚み:3mm）を木板の上に設置した。固体音の取得には、物体の振動の取得に特化したマイクであるコンクリートマイク（SMT 社製 SM-222）を用いた（図 3.4）。また、コンクリートマイクをクランプによりガラス板と木板に固定した。固定の際、ガラス板以外からの固体音がマイクに伝達しにくいように、クランプ、木板とガラス板との間にゴム（和気産業社製 EGW001）を挟んだ。また、スポイト操作による固体音を取得するためのオーディオインタフェースとして、Roland 社の CAKEWALK UA-1G を用いた（図 3.5）。今回 CAKEWALK UA-1G を用いたのは、PC 付属のオーディオインタフェースより良い音質を保つためである。また、アンプとして、図に示す簡易なアンプを用いた。

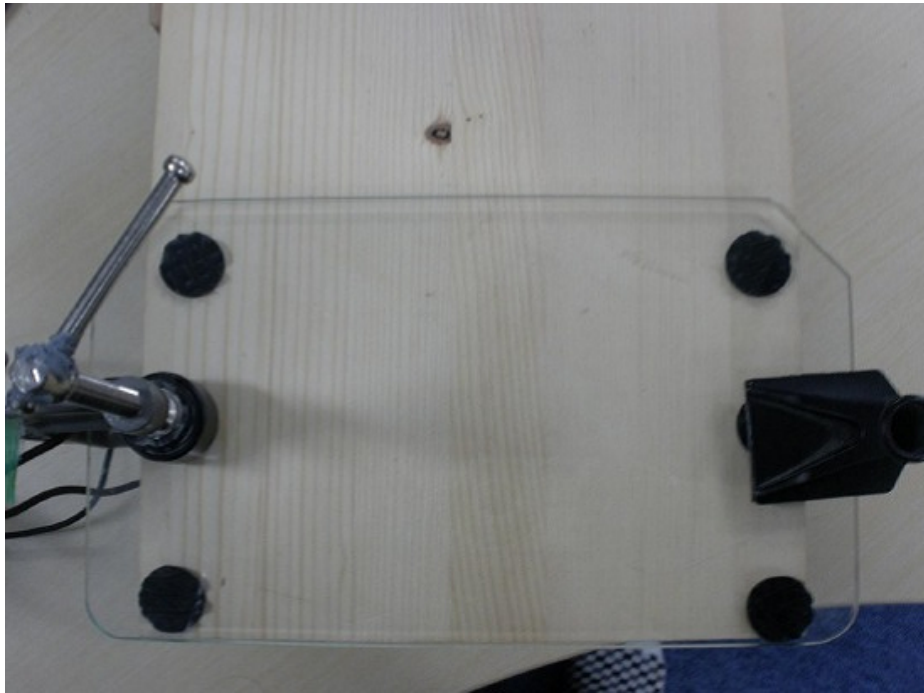


図 3.3: 認識実験装置上面



図 3.4: コンクリートマイク



図 3.5: オーディオインタフェース

### 3.2.2 ソフトウェア

オーディオドライバとして、低レイテンシ性に優れている（レイテンシが 10ms 以下）ASIO ドライバを用いた。またオーディオ I/O ライブラリとして portaudio<sup>1</sup> を用いた。FFT ライブラリとして FFTW<sup>2</sup> を、SVM ライブラリとして LIBSVM<sup>3</sup> を用いた。

## 3.3 実験 1: 認識アルゴリズムに SVM を用いた認識実験

我々は、スポイト操作の認識を、スポイト操作の際に生じる摩擦音が固体音としてマイクに伝わるもの取得し、その音声波形を FFT した結果を解析することによって行っている。本節ではこの解析を実時間で行うために、FFT の結果を SVM を使って解析する方法を試みた。3.3.1 節に、SVM を使って認識を行うための準備として、スポイト操作のための SVM 認識モデルを生成する方法を述べる。次に、3.3.1 節に、生成した SVM 認識モデルを用いてスポイト操作を認識する方法を述べる。

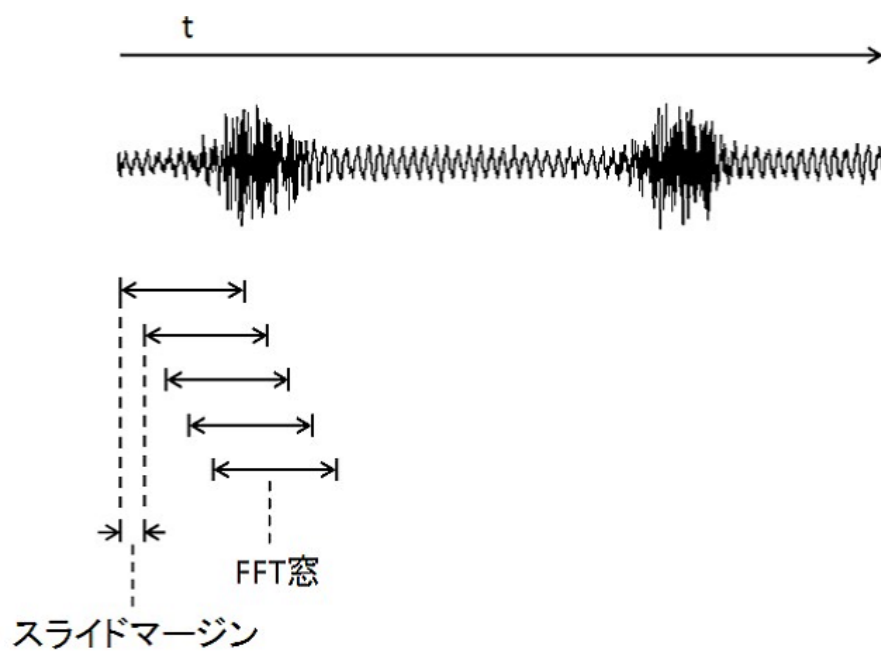


図 3.6: FFT 窓をずらしながら FFT を行い、各 FFT 窓ごとにスペクトルデータを生成する

### 3.3.1 手法

#### 固体音の FFT

取得した固体音を、FFT 窓長ごとに FFT し、スペクトルデータを生成する。スペクトルデータを生成する際、FFT 窓長を長くするとスペクトルデータの周波数分解能は向上するが、FFT の時間分解能が低くなってしまふ。一方、FFT の時間分解能を向上させるために FFT 窓長を短くすると、スペクトルデータの周波数分解能が低くなってしまふ。そこで我々は、このトレードオフを解消するため、図 3.6 に示すように、FFT 窓を少しずつスライドマージン分だけスライドさせては FFT することによって、スペクトルデータを生成した。FFT 窓をスライドさせながら FFT することにより、スペクトルデータの周波数分解能を保ちつつも、FFT の時間分解能を向上させることができる。

#### SVM 認識モデルの生成

3.3.1 節にて得られた各スペクトルデータが、それぞれどの操作に対応しているのかラベリングを行う。今回は、各スペクトルデータが、push out が行われている状態、pump up が手作業にて行われている状態、スポイト操作が行われていない状態（以降 none）のいずれかになるように手作業にてラベリングした。ラベリングを行ったスペクトルデータセットを、SVM 学習機にかけると、SVM 認識モデルが生成される。

#### SVM 認識モデルを用いた認識

3.3.1 節にて生成された SVM 認識モデルを用いて、スポイト操作の認識を行う。具体的には、3.3.1 節に示す手法を用いて、認識したいスポイト操作のスペクトルデータを取得し、各スペクトルデータを SVM 認識モデルを用いて認識する。

### 3.3.2 条件

実験は大学の研究室内の静音環境にて行った。被験者は著者の一人である。実験システムは PC (DELL Intel Core2) において動作させた。固体音の取得はサンプリングレート 44.1kHz にて行い、FFT を行う際は、FFT 窓長を 2048 サンプル数、スライドマージン長を 128 サンプル数とした。また、SVM に与えるスペクトルデータは、0Hz~11050Hz の範囲のデータを 512 次元の特徴量として与えた。カーネルトリックを SVM に適用し学習する際、Gauss カーネルを用いた。

---

<sup>1</sup><http://www.portaudio.com/>

<sup>2</sup><http://www.fftw.org/>

<sup>3</sup><http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

### 3.3.3 結果

まず、被験者が push out、pump up をそれぞれ 15 回ずつ行い、操作音を取得した。次に、これらのスペクトルデータに対し、手作業で 1027 フレームの pump up、1223 フレームの push out、4719 フレーム分の none をラベリングした。実際のスペクトルデータに対して行ったラベリングの例を、図 3.7、図 3.8 に示す。図 3.7、図 3.8 はそれぞれ pump up、push out により生じる固体音のスペクトログラムを示しており、矢印区間に pump up または push out のラベリングを行った。また、それ以外の区間はすべて none とした。次に、これらのデータを、pump up、push out、none の割合が同じである 3 つのデータセットに分け、最後に、これら 3 つのデータセットを用いて、3 分割交差検定を行った。これによって得られた認識結果を、表 3.1 に示す。

固体音の取得を終えた時点から、認識結果を出力するまでにかかる時間を測定したところ、最大 9.75ms であった。また、固体音の取得にオーディオインタフェースとして ASIO ドライバを用いているため、固体音の取得にかかる時間が 10ms 以下であることを考慮すると、スポイト操作の認識にかかる時間は最長でも 19.75ms 程度となる。したがって、使用者がスポイト操作を行う上で、操作の遅延を感じることはないと言える。

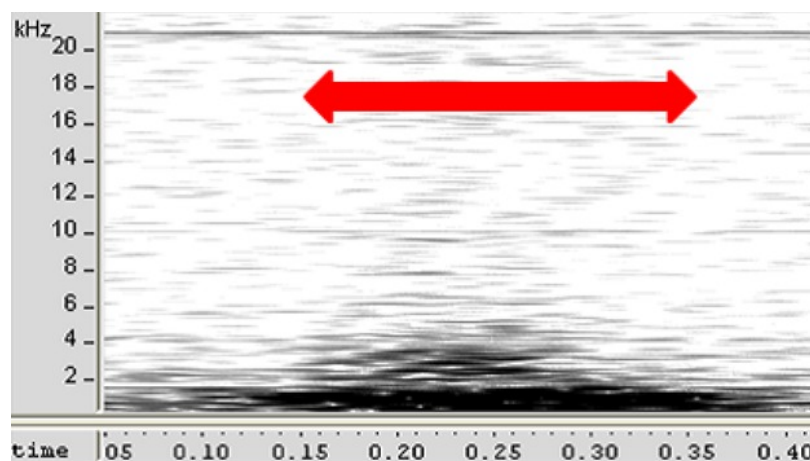


図 3.7: pump up により生じる固体音のスペクトログラム（矢印区間に pump up のラベリングを行った）

表 3.1: 実験 1 の認識結果

| Input \ Output | Output  |          |        |
|----------------|---------|----------|--------|
|                | pump up | push out | none   |
| pump up        | 57.8 %  | 10.7 %   | 31.4 % |
| push out       | 22.2 %  | 82.2 %   | 15.5 % |

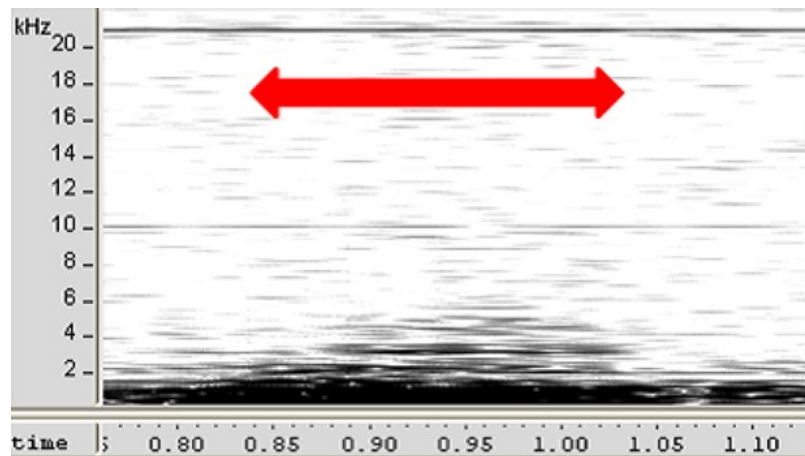


図 3.8: push out により生じる固体音のスペクトログラム（矢印区間に push out のラベリングを行った）

### 3.3.4 考察

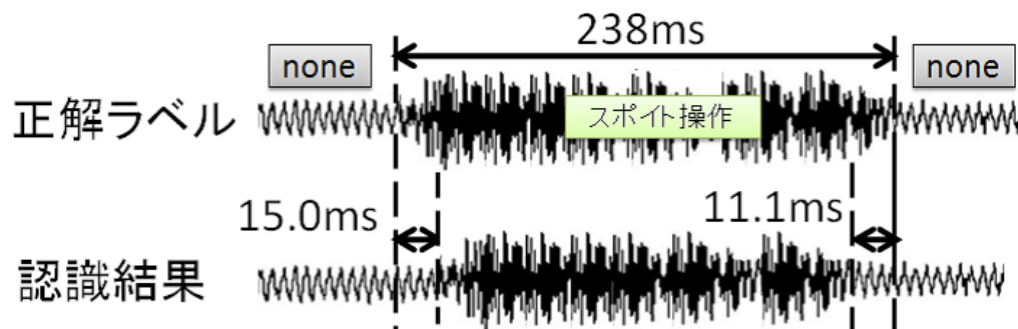


図 3.9: 正解ラベルと認識結果の違い- 始端部分と終端部分の違い

表 1 を見ると、push out は 82.2 % と高精度に認識されている。一方 pump up の精度は 57.8 % と高くない。pump up の行を見てみると、pump up は 31.4 % の割合で none と判定されている。

そこで、どのタイミングにおいて pump up、push out が none と誤認識されているか調べるために、実際の認識結果と正解データを時系列に並べて比較して調べた。調査の結果、スポイト操作による固体音の、初期部分または終端部分のみ、またはその両方の部分が、none と判定されていることがわかった。図 3.9 に、そのイメージ図を示す。また、スポイト操作の際に生じる固体音の平均時間は 238ms であり、そのうち none と判定されてしまうのは、始端部分の 15.0ms と終端部分の 11.1ms であった。none と認識されるということは、操作が検出さ

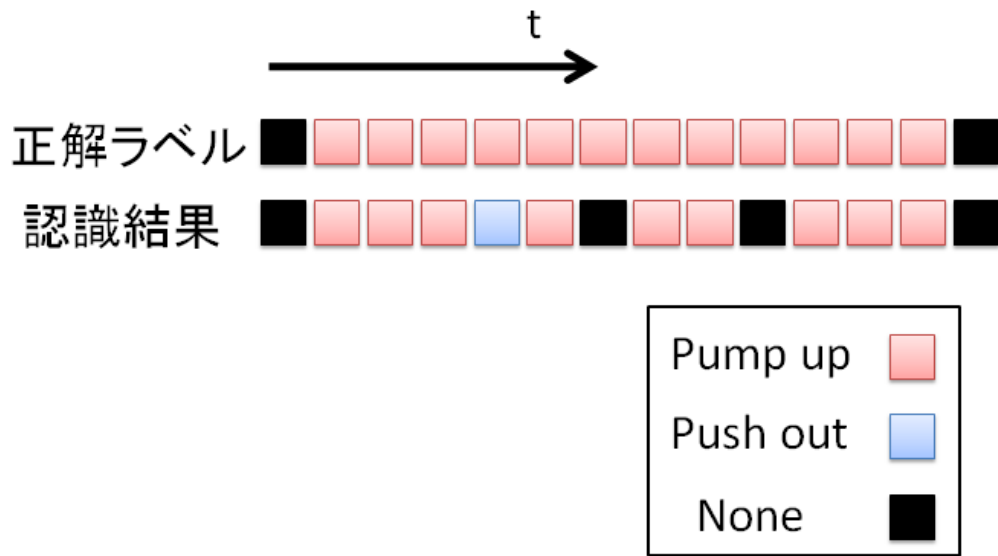


図 3.10: 正解ラベルと認識結果の違い-よくある誤認識パターン

れないことを意味する。つまり、この誤認識による影響は、スポイト操作の認識に 15.0ms 遅延が発生し、かつ認識が 11.1ms 早く終わることである。両時間は使用者が知覚出来ない程であるため、問題ないと考えられる。

また図 3.10 に示すように、スポイト操作が誤認識される際は、以下のように連続した認識結果の一部が誤認識される場合が多かった。これは、前後の認識結果を用いることにより、ご認識率を軽減できると考えられる。

### 3.4 実験 2: 認識アルゴリズムに spectral centroid の最小二乗法を用いた認識実験

本節では、スポイト操作の精度のより高い認識を模索するため、spectral centroid の最小二乗法を用いた認識実験を行った。

#### 3.4.1 手法

push out と pump up による摩擦音の差異を認識するアルゴリズムについて述べる。

固体音としてタッチパネルに伝わるスポイト操作による摩擦音を切り出し、その切り出した波形の spectral centroid を各フレームごとに求め、その spectral centroid に一次方程式の最小二乗法を適用し近似式の傾きを求め、その傾きを認識に用いる手法を考えた。spectral centroid は、スペクトラムの重心を求める手法である。Spectral spectral centroid は、スペクトラムを周

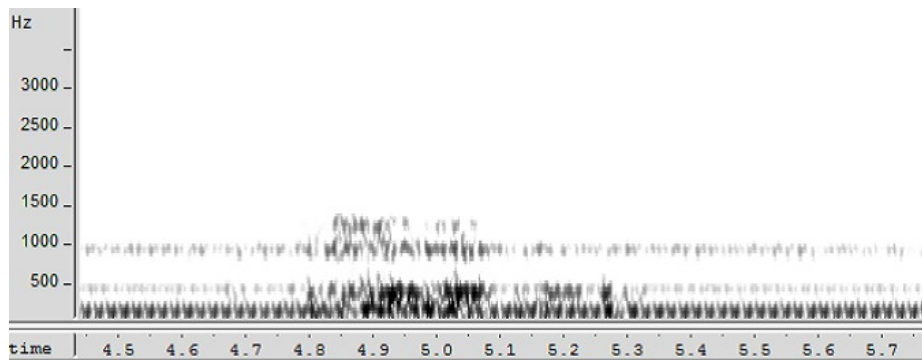


図 3.11: push out スペクトログラム

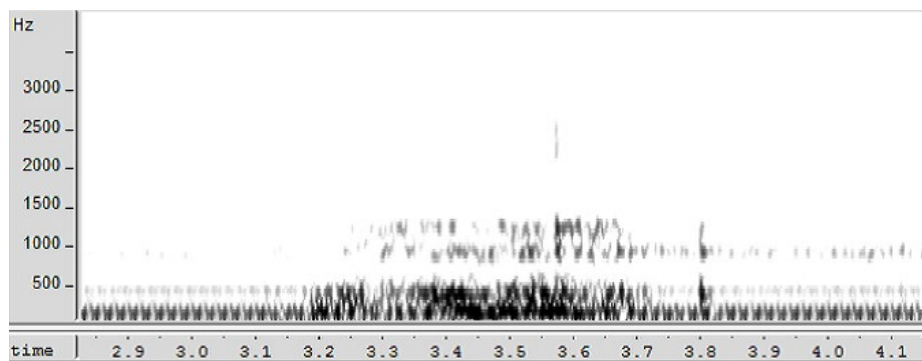


図 3.12: pump up スペクトログラム

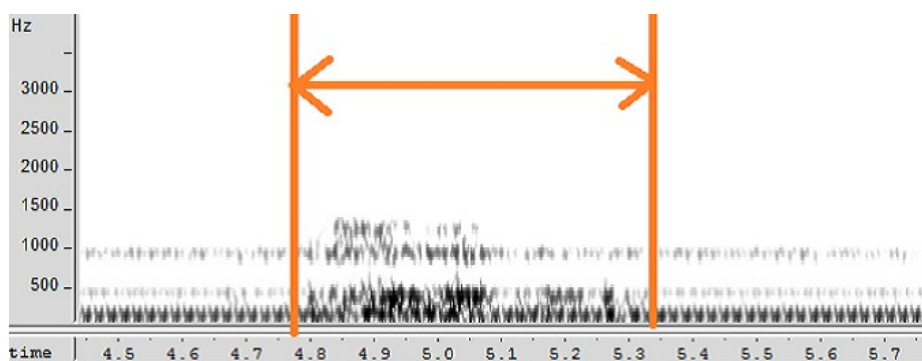


図 3.13: pump up スペクトログラムから一定振幅以上の音声の連続を切り出している様子

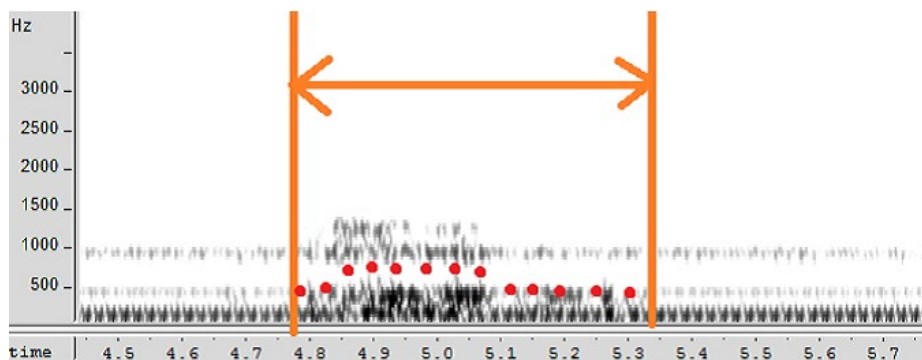


図 3.14: 切り出した pump up スペクトログラムの spectral centroid を各フレームごとに求めている様子

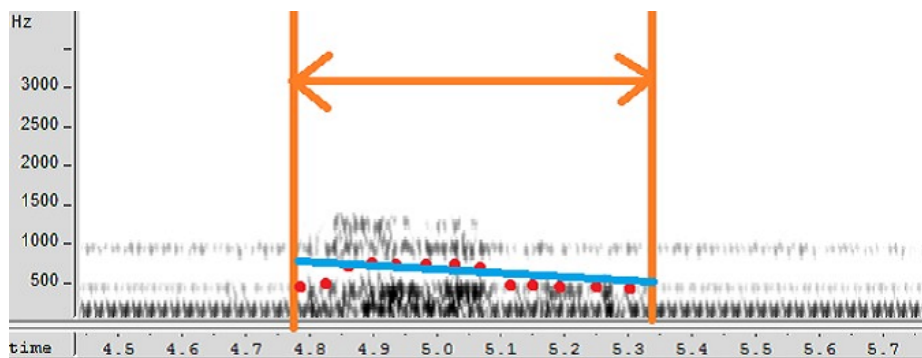


図 3.15: 求めた pump up スペクトログラムの spectral centroid に 1 次方程式の最小二乗法を適用し、近似式の傾きを求めている様子

波数軸において  $N$  個に標本化した上で、3.1 式にて求められる。 $x(n)$  は  $n$  番目に時点における標本化値、 $f(n)$  は  $n$  番目の時点における周波数値を示す。

次に最小二乗法について述べる。今、測定結果として以下の式が得られたとする。

$$(x, y) = (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

求めたい一次方程式の式を  $y = ax + b$  とおくと、 $a$  と  $b$  は式 (2) で求められる。

pump up と push out による摩擦音のスペクトログラムの比較により、pump up による固体音は高周波成分が徐々に消失し、push out による固体音は高周波成分が徐々に現れる傾向があることがわかった。pump up による固体音のスペクトログラムを図 3.11、push out による固体音のスペクトログラムを図 3.12 に示す。

この知見から、一定振幅以上の音声の連続を切り出し（図 3.13）、切り出した音声の各フレームの spectral centroid を求め（図 3.14）、その各フレームの spectral centroid に最小二乗法を適用し一次方程式を導き（3.2 式）、その一次方程式の  $a$  の値（3.3 式）、つまり傾きから push out と pump up を識別することとした。

$$Centroid = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)} \quad (3.1)$$

$$a = \frac{n \sum_{k=1}^n x_k y_k - \sum_{k=1}^n x_k \sum_{k=1}^n y_k}{n \sum_{k=1}^n x_k^2 - (\sum_{k=1}^n x_k)^2} \quad (3.2)$$

$$b = \frac{\sum_{k=1}^n x_k^2 \sum_{k=1}^n y_k - \sum_{k=1}^n x_k y_k \sum_{k=1}^n y_k}{n \sum_{k=1}^n x_k^2 - (\sum_{k=1}^n x_k)^2} \quad (3.3)$$

### 3.4.2 条件

本プロジェクトで開発した認識手法を用いて、push out と pump up の認識精度の実験を行った。実験データとして、開発者の push out、pump up の音を、それぞれ 40 データずつ用いた。ハードウェアとしては、過去に用いてきたタッチパネルのプロトタイプを用いた。場所は、静音状態における研究室において行った。その他の実験条件を以下に示す。

- サンプリングレート: 44100 Hz
- FFT フレーム長: 1024 (23.2msec)
- FFT スライドマージン長: 512 (11.6msec)
- 窓関数: ハニング窓

### 3.4.3 結果

各操作の認識率を 3.2 に示す。

表 3.2: 実験 2 の認識結果

| Input \ Output | Output  |          |
|----------------|---------|----------|
|                | pump up | push out |
| pump up        | 90 %    | 10 %     |
| push out       | 10 %    | 90 %     |

#### 3.4.4 考察

認識精度は、各操作 90 %となっており、高精度であると思われる。

今後は実用化に向けて、より高い精度の確保を行う必要がある。また、現在の設計では、認識を失敗すると、別の命令が実行されてしまうため、可能であればこの点も改善する必要がある。

## 第4章 スポイト操作を用いたアプリケーション



図 4.1: アプリケーションシステム構成

スポイト操作を用いてオブジェクトのレイヤを切り替えることのできるドローアプリケーションを、携帯端末上に実装した。

### 4.1 システム構成

システム構成図を図 4.1 に示す。またタブレット端末にコンクリートマイクをどのように取り付けたかを、図 4.2、図 4.3 に示す。コンクリートマイクの集音部を携帯端末に密着させるために、太い輪ゴムで携帯端末にコンクリートを縛り付けた。今の実現において携帯端末として図 4.4 に示す Samsung 社<sup>1</sup>の galaxyTab<sup>2</sup>を用いた。

今回は、タッチパネルに伝わる固体音を PC に取り込み、その認識結果を Bluetooth を用いて携帯端末へ送り、携帯端末のドローアプリケーションに認識させた。

<sup>1</sup><http://www.samsung.com/jp/>

<sup>2</sup><http://www.samsung.com/global/microsite/galaxytab/>



図 4.2: タブレット PC にコンクリートマイクを取り付けた様子



図 4.3: タブレット PC にコンクリートマイクを取り付けた様子 (側面図)



図 4.4: Sumsung 社のタブレット端末 GALAXY Tab

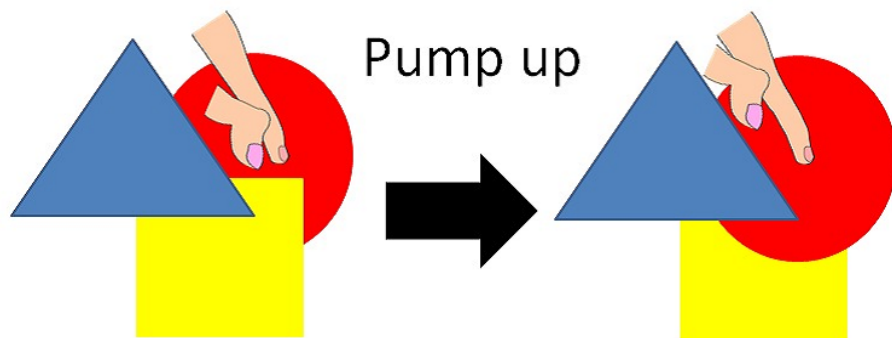


図 4.5: ドローアプリケーションでの pump up

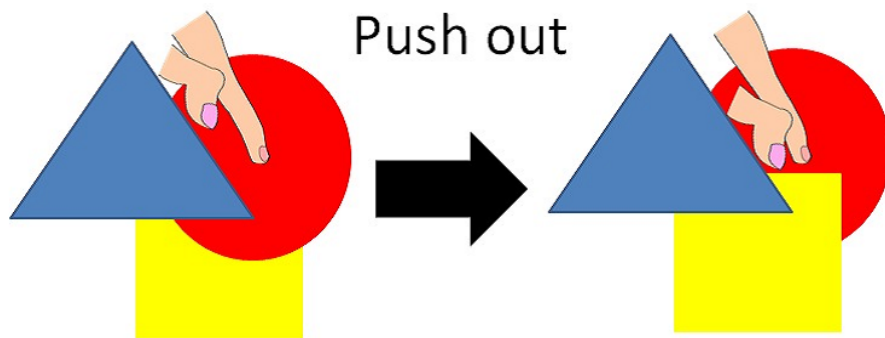


図 4.6: ドローアプリケーションでの push out

## 4.2 ドローアプリケーションの機能説明

本アプリケーションは簡易ドローアプリケーションである。使用者は、メニューから図形を選択し、画面をタップすることによりオブジェクトを生成、ドラッグによりオブジェクトを移動、マルチタッチ操作のひとつであるピンチ操作によって図形の大きさを変更しマルチタッチ操作のひとつである回転操作によって、図形の色を変更する。本アプリケーション上のオブジェクトは、それぞれ階層構造をもっており、pump up により一段上のレイヤへ移動し（図 4.5）、push out により一段下のレイヤへ移動する（図 4.6）。実際に操作している様子を図 4.7 に示す。

## 4.3 ドローアプリケーションを使用した結果

筆者が実際にアプリを操作してみたところ、pump up を用いてオブジェクトを一段上のレイヤに移動させる操作、push out を用いてオブジェクトを一段下のレイヤに移動させる操作は、とても自然な操作であると感じた。また、ドラッグによるオブジェクトの移動操作と、スポイト操作によるレイヤの切り替え操作は、非常にスムーズに切り替えことができ、感覚的には連続的であった。この利点を活かし、かつてはレイヤの切り替え操作のために作業を中断しなければならなかったが、レイヤ切り替えを交えても連続的に作業を行うことが可能となる。

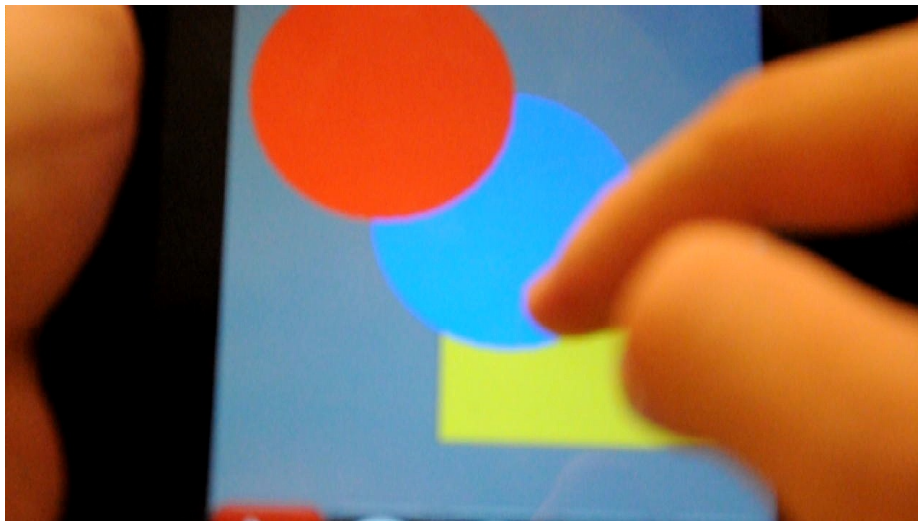


図 4.7: スポイト操作を用いたドローアプリケーションを実際に使用している様子

## 第5章 関連研究

### 5.1 固体音を用いた操作

固体音を操作に用いることにより、操作面として利用されることの無かった面が、操作面として利用することができるようになる。

Amento らは、指先を操作面としたジェスチャを認識するシステムについて述べた [1]。このシステムにより、Amento らは携帯デバイスの操作に、タップする、フリックする、親指を擦るといった、ジェスチャ操作を、携帯デバイスの操作に割り当てることを実現した。ジェスチャの認識には、手首に装着したピエゾマイクを用いて取得し、骨を媒介として指先から手首に伝わる固体音を用いている。

綾塚らは、壁や机などに対する固体音から、ノックの位置を検出するシステムについて述べた [17]。このシステムにより、壁面上に仮想ボタンを配置し、操作可能にする、棚上に並んだ物に対して音声メモを残すなどのインタラクションを可能にした。固体音の位置検出には、複数の振動センサを壁や机に設置し、各振動センサへ伝わる固体音の時間差を用いている。

Harrison らは、あらゆる表面上においても、爪を用いて引っ掻くジェスチャを認識する Scratch Input について述べた [5]。このシステムにより、携帯デバイスへの命令に、あらゆる表面からの、爪を用いて引っ掻くジェスチャによる操作の割り当てを実現した。しかし、あらゆる表面を引っ掻くジェスチャによる固体音からジェスチャを認識したため、引っ掻いた方向を認識することはできなかった。

また、Harrison らは、バイオ・アコースティック感知アレイを腕に装着することで、皮膚上のタッチ位置を検出するシステムを開発している [6]。

根岸ソ citenegisi らは、信号解析による実世界の音イベント認識を、部品のように手軽かつ低コストに利用可能にするスマートセンサと、対象音イベントに適した認識処理を本センサへ自動的に設定する Instant Learning 手法を提案した。根岸らの手法では、DP マッチングに基づく認識処理において、最も性能が良い特徴量の種類や窓長などの各パラメータの組み合わせを、認識率と誤認識率の評価の試行により自動的に選択する。これらにより、信号処理プログラミングを行うことなく、高度な実世界イベント認識が手軽に利用可能にした。

本研究は、タッチパネル面に伝わるスポイト操作による固体音を用いるため、デバイスを装着せずに操作の追加を実現することができる。また、タッチパネル平面状の操作から独立した操作が行えるため、タッチパネルの利用に適した操作が可能である。

## 5.2 スポイトメタファを用いた操作

スポイトの操作感を得るデバイスを用いることにより、操作にスポイトのメタファを割り当てることができる。

神武らは、スティック状のデバイスを用いることにより、現実世界の対象物から情報を取り込む、取り出すシステムを開発した [15]。これにより、コンピュータ等の電子世界におけるドラッグアンドドロップ操作を現実世界へ拡張する操作を実現した。

また、Zigelbaum らは、スポイト型のデバイスを用いた、現実世界の物質とコンピュータディスプレイ間の、データの移動を行うシステムを開発した。

本研究は、指の動きを用いてスポイトの動きを模倣することにより、操作に装置を用いることなく、タッチパネル操作にスポイトメタファの操作を追加することを実現した。

## 5.3 タッチパネルにおける奥行き操作感を得る操作

近年、3次元データを扱うアプリケーションが多くなっている。これに供い、タッチパネル操作において3次元方向の操作を行う研究が挙げられる。

國田 [19] らは、高田らが開発した3次元表示装置である DFD[13] 表示装置を用いて、2次元画像のみにより表現可能である以上の視覚的フィードバックを与えることにより、ある一点に対する奥行き方向の操作感を与える操作を実現した。しかし、触覚フィードバックによる奥行き操作感は全く考慮されていなかった。

また、内藤 [8] らは、円筒型マルチタッチインターフェースにおいて、ボールをつかむ、はなすジェスチャを用いて、ある一点に対する奥行き方向の操作感を与える操作を実現した。しかし、ボールをつかむ、はなすジェスチャは、五本の指を用いる必要があるため、正確にある一点を狙うのは難しく、特に、小型のタッチパネルデバイスに用いるのは困難である。

竹岡 [12] らは、複数の層のレーザーにより構成されるレーザーアレイを用いて、ディスプレイ表面付近の指先の3次元座標に加え、指の姿勢も認識することができる Z-touch プラットフォームを開発した。Z-touch により、従来のマルチタッチではできなかった「ものをつかむ」「持ち上げる」「近づける／離す」「指の姿勢の認識」など、指先がディスプレイから離れている状態の3次元的动作を利用した多種多様なインタラクションが可能になる。そのため、より実世界に近いような操作方法をユーザが行うことができるようになる。

本研究のスポイト操作は、従来研究と異なり、通常の平面タッチパネルのある一点に対し、スポイト操作の動作による、触覚フィードバックと視覚的フィードバックを用いた奥行き操作感を得る操作を実現することができる。

## 5.4 ドラッグアンドドロップ中の操作

小林らの Boomerang[7] は、ブーメランのメタファを用いることによって、ドラッグアンドドロップ操作におけるドラッグ操作を一時中断させる操作手法である。従来ドラッグアンド

ドロップ中に出来なかった操作を可能にする。しかし、通常のドラッグアンドドロップ操作の際の、今、どのアイコンをドラッグしているか、という視覚的フィードバックによる情報がない。

fold-and-drop[4] は、紙を折り曲げるメタファを用いてドラッグアンドドロップ中におけるウィンドウ切り替え操作を可能にした。しかし、目的のウィンドウを表示するために、ウィンドウを折り曲げる操作には、一旦つかんでいるアイコンをウィンドウの端まで移動させる必要があった。

本研究のスポイト操作は、通常のドラッグアンドドロップ操作中に行えるため、ドラッグアンドドロップ操作を中断することなく、ウィンドウの切り替え、ウィンドウのスクロールやフォルダの階層移動を行える。

## 5.5 タッチパネル操作への操作の追加

鷲野らは、タッチパネル操作に、指の近接によるマウスオーバー機能を追加するシステムについて述べた [18]。このシステムにより、タッチ操作を前提としているためのボタンの大きさの制約の解除、タッチしようとしているボタンのフォーカス機能、タッチしようとしているボタンのプレビュー機能を実現した。指の近接の取得には、特殊なタッチパネル装置を用いた。

本研究のスポイト操作は、特に特殊なタッチパネルを用いず、タッチパネルに固体音を集音するマイクを取り付けることにより、タッチパネル操作への操作の追加を行う。

## 第6章 まとめと今後の課題

今回、我々は、タッチパネルへ伝わる固体音を FFT し、その結果を SVM を用いて解析することにより、スポイト操作を実時間認識する手法によりスポイト操作の実時間認識が可能であることがわかった。

また、タッチパネルへ伝わる固体音の spectral centroid に最小二乗法を用いることにより認識する手法により、実時間ではないものの SVM を用いた手法よりもより精度の高い認識が可能であることがわかった。また、作成したスポイト操作を用いたドローアプリケーションが有用であることもわかった。

今後は、認識の精度の向上、およびスポイト操作を用いたアプリケーションの有効性の評価を行う予定である。

## 謝辞

本論文の執筆にあたり、指導教員である田中二郎教授、志築文太郎先生をはじめ、三末和男准教授、高橋伸講師ならびに同大学のマルチメディア研究室の山田武志准教授には多くのご助言やご指導をいただきました。心より感謝申し上げます。

同大学の嵯川友宏先生には、私の論文の副査を担当して頂きました。

また、本研究の一部は、2010 年度未踏 IT 人材発掘・育成事業の助成を受けて行われました。IPA の皆様をはじめ、プロジェクトマネージャの後藤先生には、プロジェクトの進行にあたって多くのご助言とご助力を頂きました。

また、IPLAB の皆様にも大変お世話になりました。特に WAVE チームの皆様にはチームゼミだけでなく日常的にご意見を頂きました。最後に日頃より私を支えてくれました家族や友人たちに心より感謝いたします。

## 参考文献

- [1] Brian Amento, Will Hill, and Loren Terveen. The sound of one hand: a wrist-mounted bio-acoustic fingertip gesture interface. In *CHI '02 extended abstracts on Human factors in computing systems*, CHI EA '02, pp. 724–725, New York, NY, USA, 2002. ACM.
- [2] Daniel Ashbrook, Patrick Baudisch, and Sean White. Nanya: subtle and eyes-free mobile input with a magnetically-tracked finger ring. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pp. 2043–2046, New York, NY, USA, 2011. ACM.
- [3] Gabe Cohn, Daniel Morris, Shwetak N. Patel, and Desney S. Tan. Your noise is my command: sensing gestures using the body as an antenna. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pp. 791–800, New York, NY, USA, 2011. ACM.
- [4] Pierre Dragicevic. Combining crossing-based and paper-based interaction paradigms for dragging and dropping between overlapping windows. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, UIST '04, pp. 193–196, New York, NY, USA, 2004. ACM.
- [5] Chris Harrison and Scott E. Hudson. Scratch input: creating large, inexpensive, unpowered and mobile finger input surfaces. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, UIST '08, pp. 205–208, New York, NY, USA, 2008. ACM.
- [6] Chris Harrison, Desney Tan, and Dan Morris. Skininput: appropriating the body as an input surface. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, pp. 453–462, New York, NY, USA, 2010. ACM.
- [7] Masatomo Kobayashi and Takeo Igarashi. Boomerang: suspendable drag-and-drop interactions based on a throw-and-catch metaphor. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, UIST '07, pp. 187–190, New York, NY, USA, 2007. ACM.
- [8] Masaki Naito, Buntarou Shizuki, Jiro Tanaka, and Hiroshi Hosobe. Interaction techniques using a spherical cursor for 3d targets acquisition and indicating in volumetric displays. In

- Proceedings of the 2009 13th International Conference Information Visualisation*, pp. 607–612, Washington, DC, USA, 2009. IEEE Computer Society.
- [9] Anne Roudaut, Eric Lecolinet, and Yves Guiard. Microrolls: expanding touch-screen input vocabulary by distinguishing rolls vs. slides of the thumb. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pp. 927–936, New York, NY, USA, 2009. ACM.
  - [10] Yu Suzuki, Kazuo Misue, and Jiro Tanaka. Interaction technique for a pen-based interface using finger motions. In *Proceedings of the 13th International Conference on Human-Computer Interaction. Part II: Novel Interaction Methods and Techniques*, pp. 503–512, Berlin, Heidelberg, 2009. Springer-Verlag.
  - [11] Yu Suzuki, Kazuo Misue, and Jiro Tanaka. Interaction technique combining gripping and pen pressures. In *Proceedings of the 14th international conference on Knowledge-based and intelligent information and engineering systems: Part IV*, KES'10, pp. 440–448, Berlin, Heidelberg, 2010. Springer-Verlag.
  - [12] Yoshiki Takeoka, Takashi Miyaki, and Jun Rekimoto. Z-touch: an infrastructure for 3d gesture interaction in the proximity of tabletop surfaces. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, pp. 91–94, New York, NY, USA, 2010. ACM.
  - [13] 高田英明, 陶山史朗, 伊達宗和, 昼間香織, 中澤憲二. 前後2面のLCDを積層した小型DFDディスプレイ. 映像情報メディア学会誌, Vol. 58, No. 6, pp. 807–810, 2004.
  - [14] 根岸佑也, 河口信夫. 高度な実世界イベント認識を手軽に利用可能にする instant learning sound sensor の提案. 情報処理学会論文誌, pp. 1272–1286, 2009.
  - [15] 神武直彦, 暦本純一, 安西祐一郎. 現実世界での情報移動を支援する InfoSpuit. 第5回日本ソフトウェア科学会インタラクティブとソフトウェアに関するワークショップ (WISS 1997) 予稿集, pp. 151–166, 1997.
  - [16] 堀竜慈, 志築文太郎, 田中二郎. タッチパネル面におけるスポイト操作の認識の実現. 情報処理学会第72回大会, Vol. 4, pp. 89–90, 2010.
  - [17] 綾塚裕二, 暦本純一. 簡易なノック位置検出装置を用いた実世界アプリケーション. 第9回日本ソフトウェア科学会インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2001) 予稿集, pp. 191–196, 2001.
  - [18] 鷲野浩之, 岡野祐一, 川又武典. 3次元タッチパネルインターフェイス. インタラクシオン 2009 論文集, pp. 37–38, 2009.

- [19] 國田豊, 中平篤, 鈴木尚文, 木村一夫, 中沢憲二. DFD ディスプレイとタッチパネルを用いた3次元ユーザインタフェース. 第12回日本ソフトウェア科学会インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2004) 予稿集, pp. 111–116, 2004.