

タッチパネル面に伝わる固体音の SVM を用いた解析によるスポイト操作の実時間認識

堀 竜慈 志築 文太郎 田中 二郎

我々はこれまで、タッチパネル上における新たな操作手法であるスポイト操作を提案してきた。スポイト操作とは、タッチパネル面に触れた指の側面に同じ手の親指を付け、指の根元方向へ擦り上げる、または指先へ擦り下げる動作による操作である。スポイト操作を用いることにより、タッチパネル操作において、スポイトのように「吸い上げる」または「押し出す」ような操作を実現できる。従来のタッチパネル操作に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、吸い上げるまたは押し出す操作感を得る操作、ある一点に対する奥行き方向の操作感を得る操作、タッチパネル面上の操作から独立した操作のような応用が可能になる。今回、我々は、タッチパネルへ伝わる固体音を 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を検討した。本稿ではその認識手法の詳細、および認識結果について述べる。

1 はじめに

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と記述する) 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、従来のタッチパネル操作において新しい操作が実現されてきている。例としては、タッチパネルへの指の近接を用いた操作の実現 [6]、指を傾げる動作を用いた操作の実現 [2] が挙げられる。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を示した [7]。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、タッチパネルへ伝わる固体音を高速フーリエ変換 (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。本稿ではスポイト操作の概略およびその応用例、認識

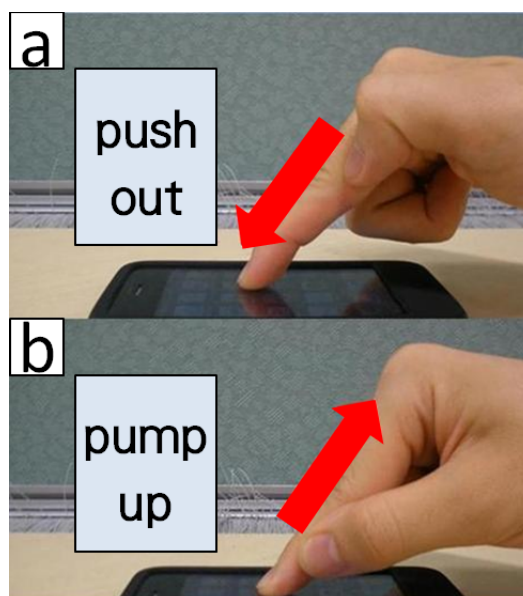


図 1 スポイト操作の動作 -
a : push out の動作, b : pump up の動作

手法の詳細、および認識結果について述べる。

Real-time recognition of " Dropper operation " analyzing structure-borne sound on touch panel by SVM.
Ryuji Hori, Buntarou Shizuki, Jiro Tanaka, 筑波大学大学院システム情報工学研究科, Graduate School of Systems and Information Engineering, University of Tsukuba.

2 スポイト操作

2.1 スポイト操作とは

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 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 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

スポイト操作は、以下の特徴を持つ。

押し出すまたは吸い上げる操作感 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。このため、使用者は操作時に、押し出すまたは吸い上げる操作感を得る。**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。

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。

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。

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し、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。具体例として、図 2 に示すように、カット命令に pump up を、ペースト命令に push out を割り当てるといった応用が考えられる。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、という応用も考えられる。

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し、3D 空間におけるオブジェクト操作に活用できる。図 3 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。

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し、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。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られ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。

3 スポイト操作の実時間認識手法

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固体音を取得 (3.1 節) し、FFT した (3.2 節) 結果を解析することによって行っている。今回、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 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、3.4 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。

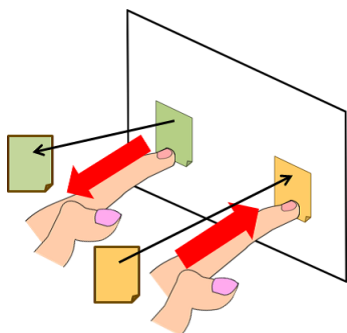


図 2 データを取り込む、取り出す命令へ割り当てられた
スポイト操作

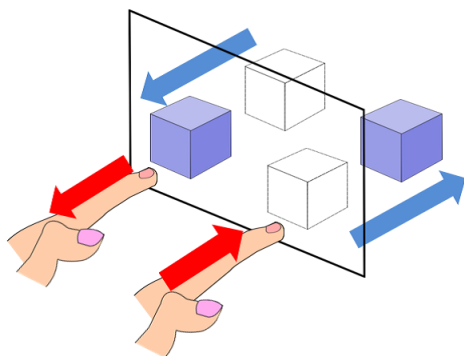


図 3 スポイト操作を用いた 3D 空間オブジェクト移動
操作

3.1 スポイト操作により生じる固体音の取得

図 5 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可能である。
- 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。

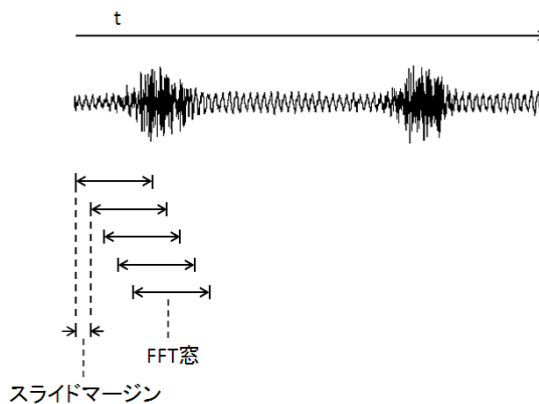


図 4 FFT 窓をずらしながら FFT を行う

3.2 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 4 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。

3.3 SVM 認識モデルの生成

3.1、3.2 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。

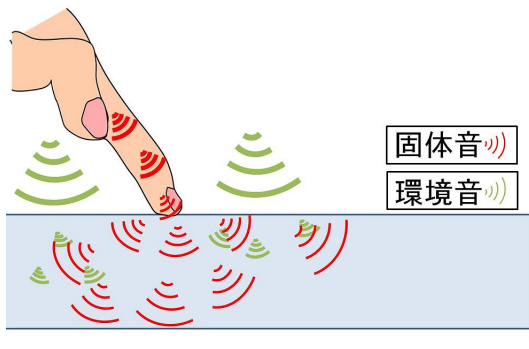


図 5 固体音の伝搬

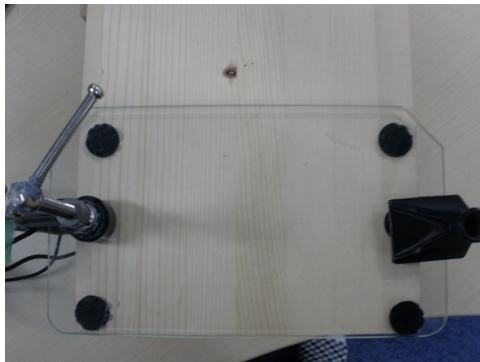


図 7 認識装置上面図

3.4 SVM 認識モデルを用いた認識

3.3 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、3.2 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。

4 認識システムの実装

4.1 ハードウェア

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 6、図 7 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。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。

4.2 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 を用いた。FFT ライブラリとして FFTW を、SVM ライブラリとして LIBSVM を用いた。

5 認識実験

提案手法によりスポイト操作を認識可能であるか検証するため、実験を行った。

5.1 実験条件

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。

5.2 スポイト操作の認識

5.2.1 認識精度の測定

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、そ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 8、図 9 に示す。図 8、図 9 はそれぞれ pump up、push out

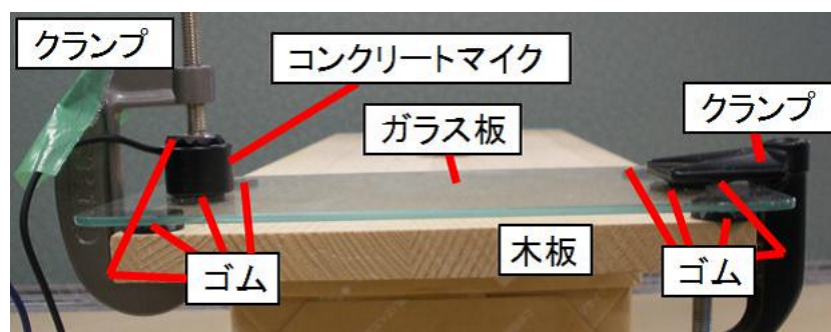


図 6 認識装置側面図

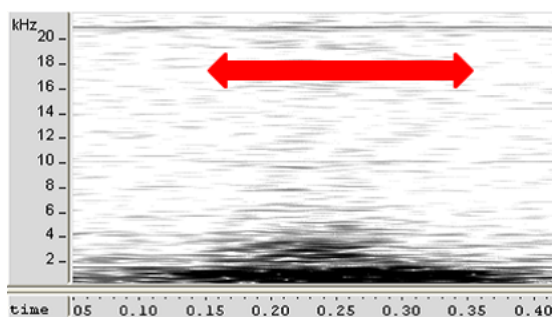


図 8 pump up により生じる固体音のスペクトログラム
(矢印区間に pump up のラベリングを行った)

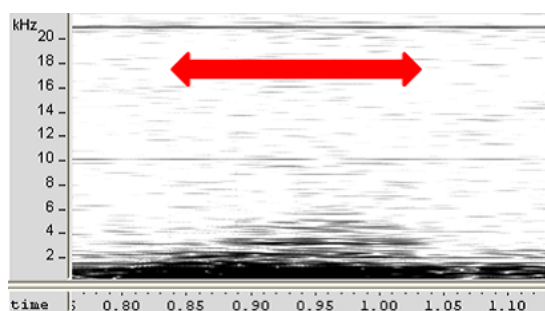


図 9 push out により生じる固体音のスペクトログラム
(矢印区間に push out のラベリングを行った)

により生じる固体音のスペクトル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

表 1 スポイト操作の認識率

Input \ Output	Output		
	pump up	push out	none
pump up	57.8 %	10.7 %	31.4 %
push out	2.2 %	82.2 %	15.5 %

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 1 に示す。

5.2.2 認識時間の測定

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。

5.3 考察

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。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、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。

6 関連研究

6.1 固体音を用いた操作

固体音を操作に用いる研究が報告されている。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 [3]。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 [8]。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、または棚上に並んだ物に対して音声メモを残すといった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、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 [4]。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 [1]。

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。

6.2 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 [9]。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 [5]。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。

7 まとめと今後の課題

本論文では、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を検討した。本提案手法による認識を用いると、改良の余地があるものの、スポイト操作を実時間認識できることがわかった。

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。

参考文献

- [1] Chris Harrison, Desney Tan, and Dan Morris. Skinput: appropriating the body as an input surface. In *CHI, '10: Proceedings of the 28th international conference on Human factors in computing systems*, pp. 453-462, New York, NY, USA, 2010. ACM.
- [2] Roudaut Anne, Lecolinet Eric, and Guiard Yves. Microrolls: expanding touch-screen input vocabulary by distinguishing rolls vs. slides of the thumb. In *CHI, '09: Proceedings of the 27th international conference on Human factors in computing systems*, pp. 927-936, New York, NY, USA, 2009. ACM.
- [3] Briant Amento, Will Hill, and Loren Terveen. The sound of one hand: a wrist-mouted bio-acoustic fingertip gesture interface. In *CHI, '02: CHI '02 extended abstracts on Human factors in computing systems*, pp. 724-725, New York, NY, USA, 2002. ACM.
- [4] Chris Harrison and Scott E. Hudson. Scrath input: creating large, inexpensive, unpowered and mobile finger input surfaces. In *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology*, pp. 205-208, New York, NY, USA, 2008. ACM.
- [5] Jamie Zigelbaum, Adam Kumpf, Alejandro Vazquez, and Hiroshi Ishii. Slurp: tangibility spatiality and an eyedropper. In *CHI, '08: CHI '08 extended abstracts on Human factors in computing systems*, pp. 2565-2574, New York, NY, USA, 2008. ACM.
- [6] 鷲野浩之, 岡野祐一, 川又武典. 3次元タッチパネルインタフェース. *インタラクション 2009*, 2009. pp. 37-38. 情報処理学会.
- [7] 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*, 2010. Vol.4, pp. 89-90. 情報処理学会.
- [8] 綾塚祐二, 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. 日本ソフトウェア科学会.
- [9] 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システム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. 日本ソフトウェア科学会.