

# アクティブ音響センシングを用いた既存物体への インタラクティブ性の付与

大野 誠<sup>a)</sup>

**概要:** 本稿は、アクティブ音響センシングを用いたタッチ認識技術について述べる。これは既存の物体に1組のコンタクトスピーカとコンタクトマイクをセンサとして取り付け、機械学習を行うことによりその物体上のタッチや把持を認識する。これにより、タッチ入力を扱うインタラクティブオブジェクトをプロトタイピングする際のハードウェア構築を容易にする。本手法の認識率を調べる評価実験を行ったところ、プラスチック製の玩具における5種類のタッチジェスチャの認識率が per-user test において 99.6%、general-user test において 97.8% となった。また、携帯情報端末における6種類の把持姿勢の認識率が、それぞれ 86.3%、71.2% となった。これらの結果から、本手法はタッチ入力を扱うインタラクティブオブジェクトのプロトタイピングへの活用が期待できる。

## Adding Interactivity to Existing Objects using Active Acoustic Sensing

**Abstract:** In this paper, we present touch sensing technique using active acoustic sensing. It recognize a rich context of touches including grasp on existing objects by attaching only a contact speaker and a contact microphone paired as a sensor. It provides easy hardware configuration for prototyping interactive objects that have touch input capability. We conducted a controlled experiment to measure the accuracy for our technique. From its results, per-user recognition accuracies with five touch gestures for a plastic toy and six hand postures for a mobile phone were 99.6% and 86.3%, respectively. Walk up user recognition accuracies for the two applications were 97.8% and 71.2%, respectively. Since the results of our experiment showed promising accuracy for the recognition of touch gestures and hand postures, this technique should be feasible for prototyping interactive objects that have touch input capability.

### 1. はじめに

携帯情報端末やタブレット等の製品においてタッチ入力がよく使われるようになったが、タッチ入力を扱うインタラクティブオブジェクトをプロトタイピングすることは、回路設計やハードウェア構築のために多くの知識や時間を要する。

また、タッチより多くのコンテキストを含む把持を認識して HCI への応用を試みる研究が数多く行われている。例えば Kim らは 64 個の静電容量センサを携帯情報端末に組み込むことにより、端末の把持姿勢を認識する [20]。また、MTPen[26] は 200 個の静電容量センサアレイを用いてペ

ンの把持姿勢を認識する。これらのデバイスは大量のセンサや特殊なハードウェアを用いており複雑な構成となる。したがって、把持認識を扱うインタラクティブオブジェクトをプロトタイピングすることも同様に困難である。

本稿では、タッチ入力を扱うインタラクティブオブジェクトのプロトタイピングを高速かつ容易にするタッチセンシング技術について述べる。これは図 1 に示すような既存の物体に対してコンタクトスピーカとコンタクトマイクをセンサとして貼り付けるだけで、それらの物体上において把持を含むタッチを認識する。これにより、開発者は回路設計やハードウェア構築よりもタッチ入力のデザインに時間を費やすことができる。本手法の原理は物体の音響特性に基づく。物体の音響特性はその物体がどのようにタッチされるかによって変化する。多くの物体はこの特性を持つため、本手法は既存の日用品等に対しても適用可能である。また、本手法はそのセンシティブリティにより大まかなタッ

<sup>1</sup> 筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻

Department of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba

<sup>a)</sup> ono@iplab.cs.tsukuba.ac.jp

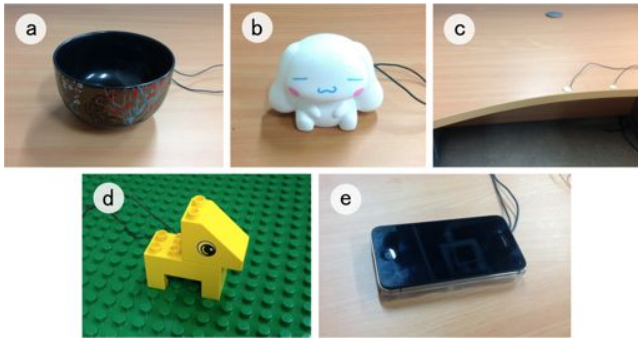


図 1 本手法を適用可能な物体の例：a) セラミック製の碗，b) プラスチック製の玩具，c) 木製の机，d) Duplo ブロック，e) 携帯情報端末（ハードケース）。

Fig. 1 Examples of applicable objects: a) ceramic bowl, b) plastic toy, c) wood desk, d) Duplo block, e) mobile device (hard case).

チ位置や、タッチの仕方（摘んだり、押し込んだり）を推定することができる。これらの利点により、開発者は物体に対するタッチインタラクションを柔軟に設計することができる。

## 2. 関連研究

### 2.1 タッチ入力を扱うプロトタイプの実成支援

タッチ入力を扱うプロトタイプの実成を支援するツールはこれまでにいくつか提案されている。BOXES[18]は画鋲やアルミホイルを物体に取り付けることにより、既存の物体からタッチ入力を扱うプロトタイプを構築可能なツールキットである。Tactile Tape[17]は既存の物体に1次元のタッチ入力領域を追加可能なタッチセンサである。WimerらはTime Domain Reflectometryを用いた変形・モジュール化可能なタッチセンサを開発し、既存の物体をタッチ入力可能に拡張した[32]。Midas[24]はGUIエディタによってタッチセンサのレイアウトの実成を支援する。これらの先行研究は導電性の素材を既存の物体に取り付けることによりタッチ入力を扱うプロトタイプを構築する。一方で、本手法は1組のコンタクトスピーカとコンタクトマイクを除いてそのような導電性の素材を必用としない。これは先行研究の手法よりもハードウェアの構築が簡単になるだけでなく、物体の元の外観が損なわれる問題を軽減する。

### 2.2 タッチ認識技術

近年のタッチ認識技術はタッチ認識をプロトタイプに用いるための敷居を下げ、そのセンシング能力も広がっている。FTIR[11]はFrustrated total internal reflectionを用いてマルチタッチサーフェスを低コストに作成する手法である。OmniTouch[12]はあらゆる面においてマルチタッチアプリケーションを利用可能にするウェアラブルなデプスセンシング・プロジェクションシステムである。これらのビジョンベースのタッチ認識技術はタッチ位

置を高精度に認識できるものの、ハードウェアの規模が大きく、画像処理のための高速なプロセッサを必要とする。Touché[23]やCapacitive Fingerprinting[13]はSwept Frequency Capacitive Sensingにより、既存の導電性の物体に触れる際の把持姿勢の認識や人の識別が可能であることを明らかにした。しかしながらTouchéは物体に導電性を要求するため、非導電性の物体上でそれらの認識を行うためには導電インクやテープを用いたコーティングを必要とする。

### 2.3 アクティブ音響センシング

アクティブ音響センシングはスピーカを用いて音を出力し、マイクでその応答を取得し解析するセンシング手法である。潜水艦等に用いられるActive Sonar[27]は古くからある典型的な例であり、音響パルスを送ってからそれが反射して返ってくるまでの時間を測定することにより周辺の物体を検出する。SAW(Surface Acoustic Wave)タッチスクリーン[2]はもうひとつの代表的な例であり、これは1枚のガラスシート、発信子、受信子、リフレクタによって構成される。ガラスの縁に取り付けられた発信子から表面弾性波を放出し、その波はリフレクタによって受信子の方向へ反射した後に受信子へと入力される。ガラス面をタッチした際に生じる振動の減衰を検出することによりタッチ位置を推定する。

また、近年においてもアクティブ音響センシングを利用した研究がいくつか提案されている。SoundWave[10]はドップラー効果を計測することによって、既存のラップトップの上空においてハンドジェスチャの認識を行う。Takemuraらは、骨伝導音をアクティブにセンシングすることにより、肘の角度を検出するウェアラブルシステムを提案した。Collinsはガラス平面を発信子で振動させ、その周波数応答を解析することによりその平面上のタッチ位置を推定した[6]。本手法は[6]に類似するが、あらゆる既存の物体上において、タッチ位置だけでなく把持のような複雑なコンテキストを含むイベントの認識を行い、インタラクティブオブジェクトのプロトタイピングを主眼とする点で異なる。

## 3. センシング原理

すべての物体は固有の音響特性を持ち、共振モード、共振周波数、モード減衰比によって表される。この特性は物体の形状、材質、境界条件に依存して決定される。現在、この特性は構造力学等の分野において、耐久性の検証や振動の抑制等に应用されている[8]、[25]。

本研究では、音響特性を決定する要素の1つである境界条件に着目する。物体に触れることは、その物体の境界条件に変化を与えることであり、これによって音響特性が変化する。この変化は共振スペクトルの変化として観測でき

る。図2にその例を示す。これはセラミック性の碗を異なる触れ方(図2a-d)で触れた際の共振スペクトル(図2A-D)を示す。したがって、物体の形状や材質に変化が無ければタッチや把持に伴う境界条件の変化は共振スペクトルの変化として観測可能である。

本手法はこの現象を利用して、音響特性を解析することにより物体がどのようにタッチされたかを推定する。音響特性の解析には実験的周波数応答解析を用いる。これは一般に、インパクトハンマーや振動スピーカのようなアクチュエータを用いて物体を様々な周波数で振動させ、その周波数応答をピエゾマイクや加速度計の様なセンサで計測する手法である。構造力学の分野ではより詳細な境界条件を求めるために、有限要素法のような理論モデル解析が用いられるが、これには物体の形状や材質が既知でなければならない。計算コストも高い。そこで、本手法では機械学習を用いて、共振スペクトルによって表される物体の境界条件とタッチジェスチャの対応付けを行う。これにより、構造が未知の物体に対してもタッチジェスチャの推定が可能であり、計算コストも抑えられる。

#### 4. 実装

以上の原理に基づいてプロトタイプシステムを実装した。図3に実装したプロトタイプシステムの全体図を示す。

##### 4.1 ハードウェア

プロトタイプシステムはコンタクトスピーカ、コンタクトマイクロフォン、信号処理と機械学習を行うコンピュータによって構成される。

物体を振動させるためのコンタクトスピーカとしてバイモルフ圧電素子を用いた(図4左; Thrive OMR20F10-BP310, 厚さ0.3mm, 直径21mm)。また、振動応答を取得するためのコンタクトマイクとして、ユニモルフ圧電素子(図4右; Murata 7BB-20-6L0, 厚さ0.2mm, 直径20mm)を用いた。圧電素子によって物体を十分な力で振動させる場合、バイモルフ構造はユニモルフ構造よりも好ましい。そのため、スピーカ側はバイモルフ構造のものを用いた。一方で振動を検知するマイクロフォン側はより安価である。

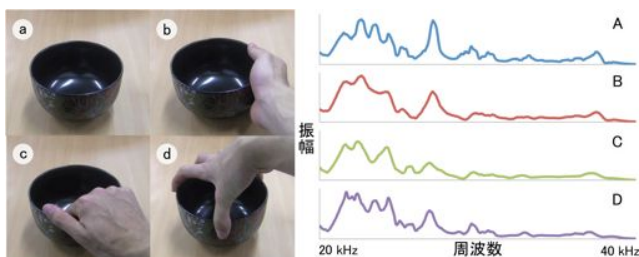


Fig. 1 セラミック製の碗に対して異なる触れ方をした際の共振スペクトル。

図2 Resonant spectra of a ceramic bowl with different ways to touch.

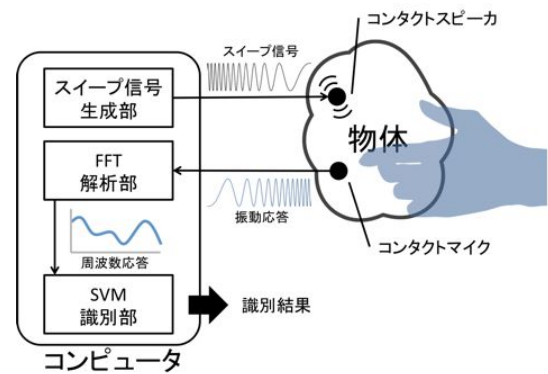


図3 プロトタイプシステムの全体図。

Fig. 3 Overview of our prototype system.

ユニモルフ構造のものを用いた。これらの圧電素子を両面テープを用いて、適用対象となる物体に取り付けた。

圧電素子を流れる信号はUSBオーディオインタフェース(Native Instruments Audio Komplete 6)を介して増幅され、コンピュータ(Apple MacBook Air, CPU: Intel Core 2 Duo 1.4GHz, RAM: 2GB)へと取り込まれる。取り込まれた信号は以下のソフトウェアによって処理が行われる。

##### 4.2 ソフトウェア

信号処理と機械学習を行うソフトウェアをC++を用いて実装した。このソフトウェアはスイープ信号生成部、FFT解析部、SVM識別部の3つのモジュールから構成される。

スイープ信号生成部は96kHzのサンプリングレートで20ms間に20kHzから40kHzまでスイープする正弦曲線のスイープ信号を生成する。生成された信号はUSBオーディオインタフェースを介してコンタクトスピーカから繰り返し再生される。繰り返しの瞬間の周波数ギャップによりインパルスノイズが発生してしまうため、それを抑制するために繰り返しの瞬間にクロスフェード処理を行う。コンタクトスピーカから出力した振動は空气中に伝達する際に音となるが、20kHzから40kHzの音は人間の非可聴域であるため聞こえない[5]。また、この周波数帯は物体に触れた際に生じる振動の周波数とも異なるため、音響特性を取り出す際の周波数帯として適している。

FFT解析部はコンタクトマイクから取得した振動応答を

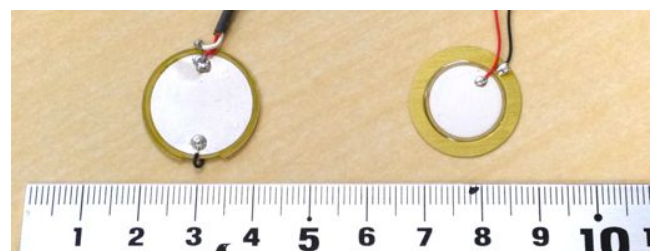


図4 コンタクトスピーカとコンタクトマイクロフォン。

Fig. 4 Contact speaker and contact microphone.



図 5 プラスチック製の玩具における 5 種類のタッチジェスチャ。

Fig. 5 Five touch gestures applied to the plastic toy.



図 6 木製の机における 5 種類のタッチジェスチャ。

Fig. 6 Five touch gestures that are applied to the wood desk.

周波数応答へと変換する。このモジュールはスイープ信号生成部と同様に 96kHz でサンプリングを行い、8192 サンプルごとに FFT を行う。FFT を行う際の窓関数としてハンギング窓を使用する。これにより 4096 点の 0 から 48kHz の周波数スペクトルが抽出される。この中から 20kHz から 40kHz の部分のみを取り出し、ダウンサンプリングすることによって 400 点の特徴点を抽出する。

SVM 識別部は LIBSVM[3] によって実装されたサポートベクタマシンを用いてタッチジェスチャを識別する。SVM のカーネルにはデフォルトパラメータの RBF カーネルを用いた。このパラメータは認識率の改善のためにグリッドサーチを用いて最適化することができる。FFT 解析部によって抽出された 400 点の特徴点は 20ms ごとに識別器に入力される。識別を行う前に、識別器は学習を行う必要がある。本システムでは、ユーザがタッチジェスチャを行う間にコンピュータのキーボードを押すことにより学習を行う。キーを押している間に入力される特徴量はそのキーコードをラベルとして対応付けられ、訓練データとしてコンピュータに蓄積される。すべてのジェスチャの学習を終えた後、スペースキーを押すことによって、識別モデルが構築され、各タッチジェスチャを識別可能になる。

## 5. アプリケーション

本節では本手法の幅広いセンシング能力を示すものとして、いくつかのアプリケーションを紹介する。日用品、Duplo ブロック、携帯情報端末からなる 3 種類の物体に対して適用した例を示す。

### 5.1 日用品への適用

本手法は様々な既存の日用品に対しても適用可能である。これを示すものとして、それぞれ材質や形状の異なる 3 種類の物体に対して本手法を適用したアプリケーションを作成した。

#### 5.1.1 セラミック製の椀

本手法をセラミック製の椀（高さ 8cm、直径 14cm）に適用することにより、インタラクティブメディアアートを作成した。椀はコンテキストやユーザの癖によっていくつかの持ち方がある（例えば、もし椀の中に熱いものが入っていれば図 2d のように持つ）。そのような慣習やコンテキストはしばしば、インタラクティブメディアアートの中で扱われることがある。本手法はアーティストによる実世界の物体を用いたインタラクティブアートの作成も支援する。このアプリケーションでは、図 2 中の 4 種類の触れ方に対して、それぞれ連想される効果音の対応付けを行った。

#### 5.1.2 プラスチック製の玩具

玩具は子供の日常生活に不可欠な要素である。子供のアクションに反応して、音を出したり、動いたりするようなインタラクティブな玩具は普及してきたものの、静的な玩具に比べてまだ高価である。本手法を応用することによって、子供は自分のお気に入りの静的な玩具に対してインタラクティブ性を加える事ができる。その例として、本来は貯金箱として使われるプラスチック製の玩具に対して本手法を適用し、5 種類のタッチジェスチャに対して異なる声を発するように学習を行った。

#### 5.1.3 木製の机

机のような大きな物体に適用する場合、センシング可能な領域はスピーカとマイクの間小さな領域に制限される。これは物体に伝わる振動がその距離に伴って減衰するためである。研究室の木製の机（厚さ 1.5cm、幅 120cm、奥行き 75cm）の端に 15cm の間隔を開けてスピーカとマイクを配置し、それらの間の面における 5 種類のタッチジェスチャの学習を行った。これにより、机上のスピーカとマイク間の領域においてそれらのジェスチャを認識可能であることを確認した。

### 5.2 Duplo ブロックへの適用

Lego もしくは Duplo ブロックは様々な形状モデルを簡単に組むことができ、プロトタイピングのためによく使用される。本節では Duplo ブロックを使ったインタラクティブオブジェクトをプロトタイピング例を示す。そのために、まず図 7a のように 2 つの圧電素子を 1 つの Duplo ブロック（2 × 4、アーチ型）の内側に取り付けられた。そして、

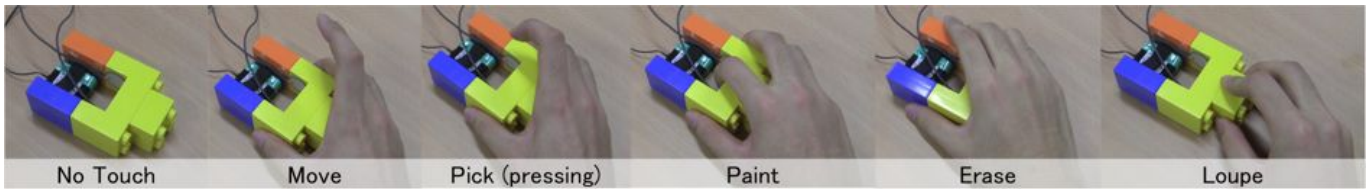


図 8 多機能入力デバイスにおける入力ツールと対応する把持状態。

Fig. 8 Tools of the multi-functional input device and corresponding ways to grasp.

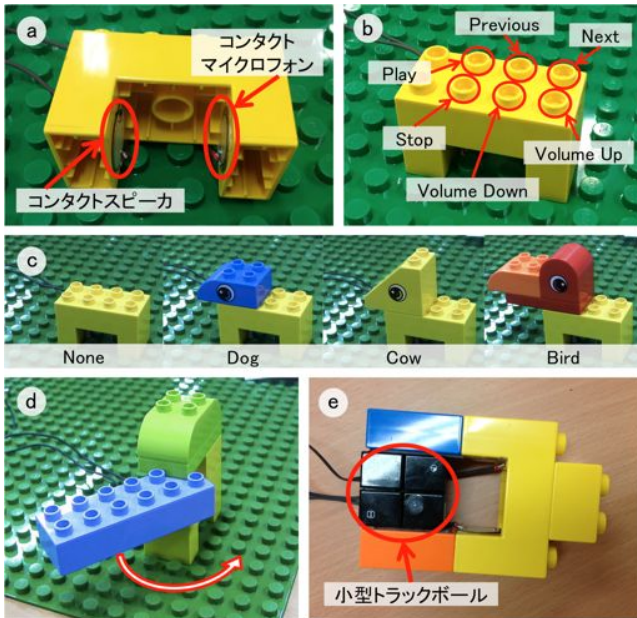


図 7 Duplo ブロックを用いたアプリケーションの例: a) 圧電素子を取り付けられたベースブロック, b) 音楽プレイヤー, c) Interactive Animal Body, d) 6段階レバー, e) 多機能入力デバイス。

Fig. 7 Examples of applications using Duplo blocks: a) a base block attached piezo-electric elements, b) music player, c) interactive animal body, d) six stage lever, and e) multi-functional input device.

このブロックをベースとして4種類のインタラクティブオブジェクトを作成した。

### 5.2.1 音楽プレイヤー

最初に、ベースブロック上の6つの突起をボタンとする簡単な音楽プレイヤーを作成した(図7b)。これらの突起にはそれぞれ音楽プレイヤーの機能(Play, Stop, Previous, Next, Volume Up, Volume Down)が割り当てられている。システムは各突起上のタッチイベントを検出し、Apple Scriptを介してiTunesをコントロールするコマンドを送信する。

### 5.2.2 Interactive Animal Body

組まれたブロックの構造や配置を推定することができる。これはブロックの形状や配置によってことなる音響特性が生じるためである。我々はこれを利用してInteractive Animal Bodyと呼ばれるアプリケーションを作成した。このアプリケーションはベースブロックを動物の体に見立て

て、そのブロックに異なる動物の顔を象徴するブロックを追加したときに、対応する動物の鳴き声を再生する。

### 5.2.3 6段階レバー

タッチや把持以外による音響特性の変化を用いるもう一つのアプリケーション例として、6段階レバーを作成した(図7d)。これはレバーとして振る舞うブロックがベースブロックに取り付けられており、そのレバーを動かすことによってブロックの音響特性に変化が生じることを利用したアプリケーションである。6つの異なる角度の状態を6つの異なるラベルとして学習を行った。これにより、システムはレバーの角度を認識可能となった。このレバーは先の音楽プレイヤーの音量を操作に適用した。

### 5.2.4 多機能入力デバイス

デバイスの把持状態によって機能を選択する入力デバイスがいくつか提案されている[22], [30]。ブロックに小型のトラックボール(Sanwa Supply 400-MA018)を組み込むことによりペイントアプリケーション向けにデザインされた同様のデバイスを作成した(図7e)。このデバイスは5種類の入力ツールを把持状態によって選択することができる(図8)。例えばユーザは先端の黄色いブロックをつかむことにより、ルーベツールが選択される。



図 9 携帯情報端末における3段階の把持力。

Fig. 9 Three hand postures for pressure recognition.



図 10 iPhone 4S と作成したケース。

Fig. 10 iPhone 4S and a grasp sensitive case.



図 11 携帯情報端末における 6 種類の把持状態。  
Fig. 11 Six hand postures for posture recognition.

### 5.3 携帯情報端末への適用

多くの研究者が把持状態を認識可能な携帯情報端末を開発し、それによってどのようにユーザインタフェースが改善されるかを探求している [4], [9], [20], [29], [33]. ほとんどの研究者は端末の把持姿勢や把持力を認識するために複数のタッチセンサを用いている. GripSense[9] は端末に予め内蔵されたセンサのみを用いて把持姿勢や把持力を認識可能であるが、把持姿勢を認識する際にユーザに画面のスイープ操作を要求する. 本手法は携帯情報端末の把持姿勢や把持力の認識にも応用可能である. 我々は iPhone4S 用のハードケースに 2 つの圧電素子を取り付けることにより、把持認識可能なケースを作成した (図 10). 2 つの素子とケーブルの厚さにより、ケースを本体に取付た際に不安定な構造になるため、6 つのプラスチックビーズをケース底面に取り付けることにより、底面の厚さを均一化した. 以上の構成を用いて、把持姿勢と把持力の認識を行った.

#### 5.3.1 把持姿勢の認識

近年の携帯情報端末のアプリケーションを使用する際に独特の把持姿勢を伴うものがある. 例えば、電話をする間、ユーザの親指は端末の側面に沿い、他の 4 本指が反対側の側面に触れるような把持姿勢になる. カメラを使うときは端末の 4 隅を摘むような把持姿勢になる. したがって、もしこれらの把持姿勢を認識できれば、持ち方に応じてアプリケーションを自動的に切り替えるといったことが可能になる [20], [33]. 加えて、異なる把持姿勢において、タッチの届く範囲等が異なるため、端末の操作性や使い方に影響する [1]. もしデバイスがユーザの把持姿勢を認識できれば、その把持姿勢において最も使いやすくなるように UI を調節することが可能になる (例: キーのレイアウト [1] やスクリーンローテーション [4] の最適化). 本手法はそのようなインタラクションを探求したい研究者に対して、デバイスのプロトタイピングを行うための簡単なハードウェア構成を提供する. 我々は 3 種類のアプリケーション (Call, SMS, Camera) と 3 種類の持ち手 (Left handed, Right handed, Two-handed) の組み合わせから構成される 6 種類の把持姿勢 (図 11) を認識するようにシステムを学習させた.

#### 5.3.2 把持力の認識

タッチスクリーンを持つ携帯情報端末はいくつかのマル

チタッチジェスチャが使用可能であるが、多くのユーザは片手による操作を望んでいるため [19], 親指を用いたシングルタッチ操作が未だ良く用いられる. しかしながら、片手操作中はシングルタッチ操作しか行えないため、入力が制限されている. 一方で、もし、端末を握る強さを認識できれば、それを新たな操作チャネルとして活用することにより、片手操作時の操作性の改善が期待できる. 把持力も端末の音響特性に変化を与えるため、本手法を用いて異なる力を加えた状態の音響特性を学習することにより、把持力を認識することも可能である. 我々は 3 種類の把持力を認識するようにシステムを学習させた.

## 6. 評価実験

本手法の認識率と学習回数のトレードオフを調べる評価実験を行った. 対象のアプリケーションとして、5.1.2 節で述べたプラスチック製の玩具における 5 種類のタッチジェスチャと、5.3.1 節で述べた携帯情報端末における 6 種類の把持状態認識を選択した.

### 6.1 被験者

この実験は 22 歳から 32 歳の 10 人の被験者 (男性 7 人, 女性 3 人) に協力していただいた. すべての被験者はスマートフォンの利用者であり、1 週間から 3 年の利用歴があった.

### 6.2 実験手順

実験では、各被験者が対象となる 2 つのアプリケーションについて独立に次のタスクを行う.

まず、各ジェスチャセット (すなわち、図 5 の 5 種類のタッチジェスチャと図 11 の 6 種類の把持姿勢) の中から 1 つの写真がランダムに選択され、被験者に 7 秒間提示され、被験者はそれに従って同様のジェスチャを行う. ジェスチャを行う間、その姿勢を微動させるように指示した. これは、音響特性が物体に対する触れ方の微妙な変化に対して敏感に変化するため、それらの変動に対してロバストな識別結果を得るためである. 最後の 1 秒間に共振スペクトルデータ (400 点の特徴ベクトル) を 40 個分取得する. 写真が提示されてから 7 秒経過すると、その後異なるジェスチャの写真が提示される. 以上の手順はすべてのジェス

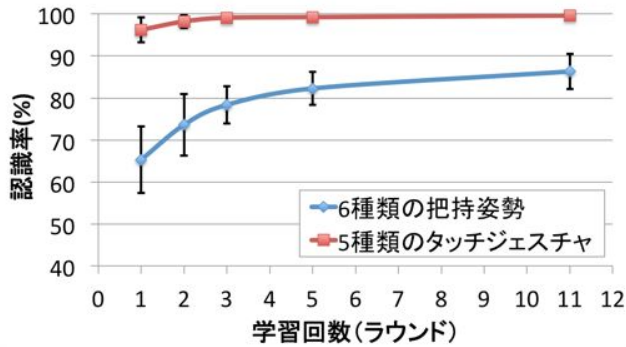


図 12 Per-user classifier における認識率.

図 13 Accuracy using a per-user classifier.

チャの写真が提示されるまで繰り返される（以降これを1ラウンドとして扱う）。各被験者は12ラウンド分のジェスチャを行う。これによって各被験者、各ジェスチャあたり480個のデータが得られることになる。

本実験からは、プラスチック製の玩具について、24000個のデータ（10被験者×5タッチジェスチャ×12ラウンド×40データ）を、携帯情報端末について、28800個のデータ（10被験者×6把持姿勢×12ラウンド×40データ）がそれぞれ得られた。実験は一人あたり約25分の時間を要した。

## 7. 実験結果

以上の手順によって得られたデータを用いて、per-user classifier と general classifier からなる2種類の識別器の評価を行った。識別モデルの構築、評価にはLIBSVM (RBF kernel,  $c=1.0$ ,  $g=0.0025$ ) を用いた。

### 7.1 Per-User Classifiers

各ユーザのデータごとに識別モデルを構築し、個人利用環境で本手法を使用した際の認識率の評価を行った。各被験者のデータについて、11ラウンド分のデータを訓練データとして利用し、残りの1ラウンド分のデータをテストデータとして認識率を求めた。これをすべてのラウンドの組み合わせについて行い（12分割交差検定）、各被験者の結果から平均値を求めた。結果として、プラスチック製の玩具における5種類のタッチジェスチャの認識率は99.6% (SD=0.36%) となった。一方で、携帯情報端末における6種類の把持姿勢の認識率は86.3% (SD=4.2%) となった。

認識率と学習回数のトレードオフを調べるために、異なる学習回数で学習を行った際の認識率を求めた。そのために、各被験者のデータを2,3,4,6つのデータセットに分割し、それぞれのデータに対して6,4,3,2分割交差検定を行い、各被験者の結果から平均値を求めた（図13）。結果として、プラスチック製の玩具については1ラウンド分の学

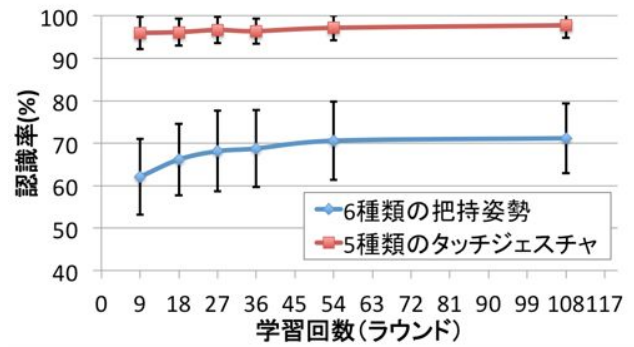


図 14 General classifier における認識率.

Fig. 14 Accuracy using a general classifier.

習データにおいても96.2% (SD=7.9%) の認識率が得られた。携帯情報端末については1ラウンド分の学習データにおいて65.3% (SD=7.9%) の認識率にとどまった。また、5ラウンド分の学習データを用いることにより、認識率は82.3% (SD=3.0%) に改善した。

### 7.2 General User Classifiers

すべてのユーザのデータから識別モデルを構築し、複数人利用環境で本手法を使用した際の認識率の評価を行った。すべての被験者のデータから9人分の被験者のデータを訓練データとして使用し、残り1人分のデータをテストデータとして認識率を求めた。これをすべての被験者の組み合わせについて行い（10分割交差検定）。結果としてプラスチック製の玩具における5種類のタッチジェスチャの認識率は97.8% (SD=2.9%) となった。一方で、携帯情報端末における6種類の把持姿勢の認識率は71.2% (SD=8.2%) にとどまった。

Per-User Classifiers と同様に認識率と学習回数のトレードオフ曲線を求めた（図14）。結果として、プラスチック製の玩具は全被験者の1ラウンド分のデータを用いた学習において96.0% (SD=3.8%) と高い認識率を保った。一方で携帯情報端末については62.1% (SD=8.9%) の認識率にとどまった。また、全被験者の5ラウンド分のデータを用いた学習付近で認識率の改善は見られなくなった。

## 8. 議論

評価実験の結果からタッチジェスチャや把持姿勢の認識について期待できる認識率が得られたことから、本手法はタッチ入力を扱うインタラクティブオブジェクトのプロトタイピングに適すると考えられる。特にプラスチック製の玩具については、図13が示すように、1ラウンドの学習を行うだけで、96.2%の認識率が得られた。このことから、開発者が自身のアイデアを即座に試したいときにおいても十分に有用なプロトタイピング手法となると考えられる。また、携帯情報端末については86.3%の認識率を得るため

には 11 ラウンドの学習を行う必要があったが、それでも所要時間は 7.7 分 (7 秒×6 把持姿勢×11 ラウンド) となり、大量のセンサを用いてハードウェアを構築するよりも十分に早い。

一方で、本手法は様々な物体に適用できるものの、適用対象に対していくつかの制約がある。まず、本手法は振動を利用するため、物体の材質は振動をよく伝達するもの(木、金属、プラスチック、セラミック等)でなければならず、振動を吸収する柔らかい材質(ゲル、粘土、ゴム)等に対する適用は難しい。また、対象となる物体が大きすぎる場合、振動の減衰によって有効な特徴量を得られる領域が制限される。この問題に対する対策として、高出力のアンプを使用するか、センサの数を増やす必要がある。

## 9. まとめ

本稿では、アクティブ音響センシングを用いたタッチセンシング技術について述べた。これは 1 組のコンタクトスピーカとコンタクトマイクをセンサとして既存の物体に貼り付けるだけで、その物体上のタッチや把持を認識する。これにより、タッチ入力を扱うインタラクティブオブジェクトをプロトタイプングする際のハードウェア構築を容易にする。いくつかの応用アプリケーションを紹介し、本手法の幅広いセンシング能力を示した。本手法の認識率と学習回数のトレードオフを調べる評価実験を行った。結果として、期待できる認識率が得られたことから、本手法はタッチ入力を扱うインタラクティブオブジェクトのプロトタイプングに適すると考えられる。

## 参考文献

- [1] Azenkot, S., and Zhai, S. Touch behavior with different postures on soft smartphone keyboards. In *MobileHCI '12*, 251–260.
- [2] Brenner, M.C., and Fitzgibbon, J.J. Surface acoustic wave touch panel system, Feb.17 1987. US Patent 4,644,100.
- [3] Chang, C.-C., Lin, C.-J. LIBSVM: a library for support vector machines., <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [4] Cheng, L.-P., Hsiao, F.-I., Liu, Y.-T., and Chen, M.Y. iRotate grasp: automatic screen rotation based on grasp of mobile devices. In *UIST '12 Adjunct*, 15–16.
- [5] Christopher, D. Frequency Range of Human Hearing. In *The Physics Factbook*. 2003.
- [6] Collins, T. Active acoustic touch interface. *Electronics Letters* 45, 20 (24 2009), 1055–1056.
- [7] Cralg, L. Right-rapid-rough. *Ask Magazine*, 13 (2003), 12–15.
- [8] Ewins, D.J. *Modal testing: theory, practice and application*, vol.2. Research studies press Baldock, 2000.
- [9] Goel, M., Wobbrock, J., and Patel, S. GripSense: using built-in sensors to detect hand posture and pressure on commodity mobile phones. In *UIST '12*, 545–554.
- [10] Gupta, S., Morris, D., Patel, S., and Tan, D. Sound-Wave: using the Doppler effect to sense gestures. In *CHI '12*, 1911–1914.
- [11] Han, J.Y. Low-cost multi-touch sensing through frustrated total internal reflection. In *UIST '05*, 115–118.
- [12] Harrison, C., Benko, H., and Wilson, A.D. OmniTouch: wearable multitouch interaction everywhere. In *UIST '11*, 441–450.
- [13] Harrison, C., Sato, M., and Poupyrev, I. : exploring user differentiation by sensing electrical properties of the human body. In *UIST '12*, 537–544.
- [14] Hartmann, B., Abdulla, L., Mittal, M., and Klemmer, S.R. Authoring sensor-based interactions by demonstration with direct manipulation and pattern recognition. In *CHI '07*, 145–154.
- [15] Hartmann, B., Klemmer, S.R., Bernstein, M., Abdulla, L., Burr, B., Robinson-Mosher, A., and Gee, J. Reflective physical prototyping through integrated design, test, and analysis. In *UIST '06*, 299–308.
- [16] Hoggan, E., Stewart, C., Haverinen, L., Jacucci, G., and Lantz, V. Pressages: augmenting phone calls with non-verbal messages. In *UIST '12*, 555–562.
- [17] Holman, D., and Vertegaal, R. TactileTape: low-cost touch sensing on curved surfaces. In *UIST '11 Adjunct*, 17–18.
- [18] Hudson, S.E., and Mankoff, J. Rapid construction of functioning physical interfaces from cardboard, thumb-tacks, tin foil and masking tape. In *UIST '06*, 289–298.
- [19] Karlson, A.K., and Bederson, B.B. Understanding single-handed mobile device interaction. Tech. rep., Dep. of Computer Science, Univ. of Maryland, 2006.
- [20] Kim, K.-E., Chang, W., Cho, S.-J., Shim, J., Lee, H., Park, J., Lee, Y., and Kim, S. Hand grip pattern recognition for mobile user interfaces. In *IAAI '06*, 1789–1794.
- [21] Miyaki, T., and Rekimoto, J. Graspzoom: zooming and scrolling control model for single-handed mobile interaction. In *MobileHCI '09*, 11.
- [22] Rekimoto, J., and Sciammarella, E. ToolStone: effective use of the physical manipulation vocabularies of input devices. In *UIST '00*, 109–117.
- [23] Sato, M., Poupyrev, I., and Harrison, C. Touché: enhancing touch interaction on humans, screens, liquids, and everyday objects. In *CHI '12*, 483–492.
- [24] Savage, V., Zhang, X., and Hartmann, B. Midas: fabricating custom capacitive touch sensors to prototype interactive objects. In *UIST '12*, 579–588.
- [25] Schwarz, B.J., and Richardson, M.H. Experimental modal analysis. *CSI Reliability week* (1999).
- [26] Song, H., Benko, H., Guimbretière, F., Izadi, S., Cao, X., and Hinckley, K. Grips and gestures on a multi-touch pen. In *CHI '11*, 1323–1332.
- [27] Stewart, J., and Westerfield, E. A theory of active sonar detection. *Proceedings of the IRE* 47, 5 (1959), 872–881.
- [28] Takemura, K., Ito, A., Takamatsu, J., and Ogasawara, T. Active bone-conducted sound sensing for wearable interfaces. In *UIST '11 Adjunct*, 53–54.
- [29] Taylor, B.T., and Bove, Jr., V.M. Graspables: grasp-recognition as a user interface. In *CHI '09*, 917–926.
- [30] Vogel, D., and Casiez, G. Conté: multimodal input inspired by an artist's crayon. In *UIST '11*, 357–366.
- [31] Wilson, G., Stewart, C., and Brewster, S.A. Pressure-based menu selection for mobile devices. In *MobileHCI '10*, 181–190.
- [32] Wimmer, R., and Baudisch, P. Modular and deformable touch-sensitive surfaces based on time domain reflectometry. In *UIST '11*, 517–526.
- [33] Wimmer, R., and Boring, S. HandSense: discriminating different ways of grasping and holding a tangible user interface. In *TEI '09*, 359–362.