

SenseSurface：アクティブ音響センシングによる 物体識別と位置推定

岩瀬 大輝^{1,a)} 伊藤 雄一^{1,b)} 秦 秀彦¹ 尾上 孝雄¹

受付日 2018年12月25日, 採録日 2019年7月3日

概要：実空間内の物体の種類や位置のような情報を取得できると、実生活において様々な支援が可能となる。そこで本研究では、そのような物体情報を取得するため、アクティブ音響センシングを用いた物体情報識別手法を提案し、検討を進めている。本稿では、特にどんな物体がどこに置かれているかという物体識別および位置情報の取得手法を提案し、評価実験を通じて提案手法の有効性を論じる。システムのプロトタイプを実装し、装置上の単一物体に関してその種類と位置をそれぞれ 98.9%、89.2%の精度で取得できることを確認した。次に、装置上の複数の物体に関して評価実験したところ、最大 98.3%の精度でそれぞれの物体の種類と位置を取得することができた。

キーワード：アクティブ音響センシング, 周波数応答, 物体識別, 位置推定

SenseSurface: Using Active Acoustic Sensing to Detect What is Where

DAIKI IWASE^{1,a)} YUICHI ITOH^{1,b)} HIDEHIKO SHIN¹ TAKAO ONOYE¹

Received: December 25, 2018, Accepted: July 3, 2019

Abstract: To acquire various information such as types of objects and their positions in the real world holds a great deal of potential in realizing calm computing. Therefore, we propose SenseSurface, which recognizes types of objects and their positions on a surface by using active acoustic sensing. By applying acoustic sensing to the flat plate and analyzing the frequency response which varies depending on the type and position of the object on the flat plate, we tried to recognize the type of object and its position. We developed a prototype of the SenseSurface system and conducted experiments. As a result, SenseSurface could classify 26 objects with 98.9% accuracy and recognize their positions with 89.2% accuracy. In addition, even if multiple objects are placed on the surface at the same time, SenseSurface could recognize the objects and their positions with at the maximum 98.3% accuracy.

Keywords: active acoustic sensing, frequency response, object recognition, position detection

1. はじめに

日常生活空間内には道具や機器といった様々な物体が存在し、人々はそれらを用いて活動している。そのため人間の行動や状態、周囲の環境などをコンピュータが認識し、人間の活動を支援するためには、実空間内に存在する物体が、どこでどのように使用されているかといった、物体の

状態や状況を認識する必要がある。実空間内の物体を認識する技術では、特に画像処理技術を用いた日常空間に存在する物体位置の認識などが進展してきた。一方で、近年のセンサ技術の発達により、加速度センサや光センサ、ジャイロセンサや圧力センサといった様々なセンサを用いて、物体の種類だけでなくその位置や使用方法までコンピュータに認識させる技術に注目が集まっている。そういった物体の状態や状況をコンピュータが認識することができれば、人が何をどこでどのように使用したかのような詳細な活動情報に基づいた行動推薦システムなど、人間の生活をより豊かにする情報技術の発展が期待できる。物体の状態や状

¹ 大阪大学
Osaka University, Suita, Osaka 565-0871, Japan
a) iwase.daiki@ist.osaka-u.ac.jp
b) itoh@ist.osaka-u.ac.jp

況は、その種類や位置、加速度など、多角的な物体情報の組み合わせで表され、認識には物体の様々な情報が必要となる。たとえば、どの物体がどこに配置されているかを認識するためには、物体の種類とその位置をそれぞれ取得する必要がある、物体の使用状態を認識するためには、物体の種類とその位置や姿勢、加速度などの情報を取得する必要がある。したがって、物体の状態・状況認識において、物体の多角的な情報を同時取得するシステムは非常に有用であると考えられる。

そこで本研究では、多角的な物体情報を取得するために、アクティブ音響センシングを用いた物体情報識別手法を提案し、検討する。アクティブ音響センシングとは、物体表面にマイクとスピーカを配し、特定の音響信号を物体内部に伝搬させ、その応答の周波数解析をすることで状態を推定する手法である。物体の周波数特性には、その材質や形状だけでなく、その物体への接触状況が反映される。そのため机や棚など、その上に物体を置くことのできる表面を持つ道具や家具の周波数特性は、その上に置かれた物体の位置や形状、向き、底面の凹凸、重量をはじめとした様々な要因によって変化すると考えられる。そこで周波数特性から要因ごとに变化する特徴を抽出することができれば、置かれている物体の情報を多角的に取得できると考えられる。多角的な物体情報の取得には、単一の情報を取得するための測定装置を複数組み合わせる手法も考えられるが、アクティブ音響センシング装置は、実装が容易かつ安価であるため、普及時のコストを抑えることができるという利点がある。

本稿では、アクティブ音響センシングを用いた多角的な物体情報取得のなかでも、特に単一物体と複数物体の物体識別と位置推定について検討する。アクティブ音響センシングで音響信号を伝搬させる物体としてアクリル製平板を用意し、その上に置かれた物体の情報を取得する。システムのプロトタイプを実装し、実装したプロトタイプを用いて物体の種類や位置を推定できるかを確認する。次に、複数物体の同時識別について検討を行い、推定式を用いた複数物体の同時識別に関する評価実験を通して提案手法を評価する。

2. 関連研究

2.1 物体情報識別

物体情報を識別する研究として、Yeo らの RadarCat は、物体に放射した電波の跳ね返りをマルチチャンネルで観測し、random forest classifier を用いて物体識別を行った [1]。Laput らは、電子機器の電磁ノイズを測定、学習することで、触れた物体を瞬時に識別するシステムを提案した [2]。しかしこれらの研究では、得られる情報はその物体が何であるかといった単一なものに限られ、物体の位置や形状、重量といった複数の情報を同時に取得することができない。一方、物体の位置、形状を認識する研究としては、

Baudisch らの Lumino がある [3]。これは、光ファイバとマーカを組み合わせたブロックを用いることで、ディスプレイの下からカメラで光ファイバ越しにマーカを認識し、ブロックの位置認識や、重なったブロックの識別からブロックの形状を認識できるシステムである。しかし、このシステムでは、位置や形状の認識をするために物体に特殊な機構を持たせたり、物体にマーカを搭載する必要がある、様々な物体を識別するシステムとしての汎用性を持ち合わせていない。物体やその位置を認識する手法として Villar らの Zanzibar [4] などで用いられる近距離無線通信 (NFC) もあるが、認識する物体にタグを持たせる必要がある、Lumino と同様の問題点を有する。このように、物体情報識別手法は多く提案されているが、得られる情報が単一なものに限られることや、様々な物体の識別には適用しづらいといった問題がある。これらの問題に対して我々の手法は、物体情報を識別したい物体が配置された表面に対してセンシングを行うため、配置物体に制限がない点で優れている、また、音響の伝搬はその接触状況の影響を受け、配置物体にかかわるいくつもの要因で変化するという特徴を持つため、様々な物体に対して、多角的な物体情報の取得を実現できる可能性があると考えられる。

2.2 音響センシング

音響を用いて動作や物体情報をセンシングする研究はさかんになされており、それらは、スピーカなどで音響信号を印加しないパッシブなもの、音響信号を印加するアクティブなものに分類することができる。

パッシブ音響センシングは、何らかのアクションによって引き起こされる音響を受信し、信号を解析する手法である。Amento らの The Sound of One Hand は、指先で叩いたり、撫でたりといった動作を、その動作によって起こる骨伝導音響を解析し認識するシステムである [5]。Murray-Smith らは、質感のある表面に対し、ユーザのスクラッチ動作や撫でるといったインタラクションによって起こる音響から、そのインタラクションが何であるかを識別した [6]。Harrison らは、質感のある物体の表面をスクラッチすることで生まれる独特な音響を用いた入力手法に関して検討した [7]。また、Harrison らは、物体に彫られたバーコード状の切れ目を、爪などの硬いものが横切ったときに生まれる音響から、そのバーコードの単一の ID を認識するシステムである Acoustic Barcodes を提案した [8]。Paradiso らは、音響の到達時間差を用いて、多数のマイクで、ガラスシート上のタップ位置を推測した [9]。Harrison らの Skininput は、生体音響センサを用いることで、肌上のタップ位置を推測する [10]。物体表面へのタップを通じた物体の識別としては、Harrison らや Lopes らが行ったものがある。Harrison らは、タップ音からタップに使われた物体のタイプを特定し [11]、Lopes らは、どの指でタップし

たかを特定した [12]. Hwang らは、ペン先のスクラッチノイズを取得することにより、感圧ペンを実現した [13].

一方、アクティブ音響センシングは、スピーカなどを用いて音響を与え、その応答を受信し、解析する手法である。Wylie らの Active Sonar は、音響を発生させ、その応答が返ってくるまでの時間を測定することで周囲の物体との距離を測定する [14]. Brenner らの Surface Acoustic Wave (SAW) touch screens は、ガラス板とトランスミッタ、レシーバ、リフレクタで構成され、ガラス板上のタッチ位置を検出する [15]. SAW は弾性のある物体の表面を進む波であり、これをトランスミッタから送信し、ガラスのエッジに並べられたリフレクタで反射する。この状態でタッチが起きたときに発生する波の吸収をレシーバで観測し、タッチ位置を測定する。Gupta らは、発生させた音響のドップラー効果から空中でのジェスチャを認識した [16]. Takemura らはスピーカと2つのマイクを用いて肘の骨伝導音響から肘の角度を認識した [17]. Collins は、ピエゾトランスミッタとコンタクトマイクを用いてガラスなどの平面をタッチパネルにする手法を提案した [18]. Laput らの Acoustruments は、アクティブ音響システムを利用し、携帯端末などに物理インタフェースを追加する手法を提案した [19].

パッシブ音響センシングに対するアクティブ音響センシングの利点として、振動をとまなわないソフトタッチやタッチ圧力の変動を取得できることがあげられる。Ono らの Touch&Activate は、この点を利用し、アクティブ音響センシングを用いて物体の触れ方を認識するシステムである [20]. これは、認識したい把持状態をラベルとし、アクティブ音響センシングで得られる周波数応答を特徴量とした機械学習を用いて、状態を推定する。また、Liu らの VibeWrite はアクティブ音響センシングを用いて指先の触れ方を認識し、新しい入力インタフェースや、触れ方の属人性を利用したセキュリティとしての可能性を提案している [21]. これらの研究では人間による物体への触れ方の認識、すなわちインタラクションの取得にアクティブ音響センシングを用いているが、我々の手法ではアクティブ音響センシングを、盤面上に置かれる物体の状態認識に用いる点で異なる。また、複数の物体を同時に認識するといった、状態の組合せ認識を実現している点でも異なる。本研究では、アクティブ音響センシングを用いて多元的な物体情報の取得を目指す。

3. アクティブ音響センシングによる物体識別と位置推定

3.1 センシング原理

物体の周波数特性は、振動モード、固有振動数、モード減衰比によって表現され、物体の形状、材質、物体への荷重や拘束による物体表面の状態（境界条件）の3つの要因によって変化する。これらの周波数特性は建築や機械の分

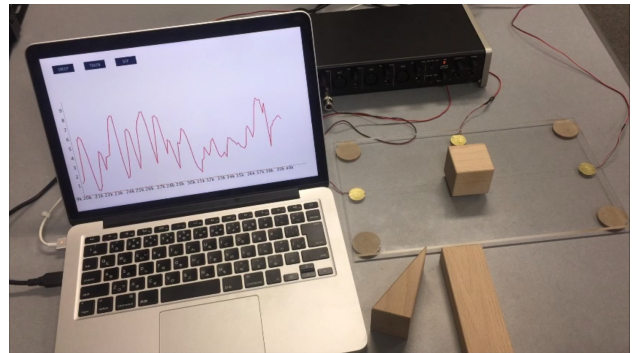


図 1 実装した SenseSurface のプロトタイプ全体図

Fig. 1 A prototype of SenseSurface.

野において、構造物の耐久度の評価や制振技術としても応用されている [20]. ここで境界条件は、物体の接触面の状態や接触位置、接触面に対する荷重などから変化し、周波数特性に影響を与える。そのため、物体配置可能表面の周波数特性の変化から配置物体の位置情報ならびに接触状況を取得できる。ここで接触状況は配置物体の底面材質や形状といった特徴を反映しているので、ここから物体認識ができると考えられる。

周波数特性の取得・解析には、前述のアクティブ音響センシングを用いる。本稿では、スピーカからアクリル製平板にスイープ信号を与え、その上に置かれた物体の状態や状況によって変化する応答をピエゾマイクで取得し、Fast Fourier Transform (FFT) を用いて周波数領域に変換することで周波数特性を解析する。スイープ信号とは、時間経過に応じてその周波数が変化する信号を指す。スイープ信号は様々な周波数を持っているので、応答から周波数特性を解析することができる [20]. 得られた周波数特性から機械学習を用いて配置物体の種類や位置を識別する。

3.2 実装

実装したシステムの全体図を図 1 に、システムの概要を図 2 にそれぞれ示す。物体を配置する表面として縦 180 mm、横 320 mm、厚さ 5 mm の直方体型のアクリル製平板を用いた。アクリルを用いた理由として、加工して実際に机に用いられる材料であること、丈夫で振動を伝搬させやすいことがあげられる。このアクリル製平板の四隅に、外部からの振動が板に伝わるのを防止するために、直径 32 mm、厚さ 5 mm の家具足用のフェルトクッションを、板の辺がクッションの接線になる位置に取り付けた。また、スピーカおよびマイクの位置としては、四隅が固定された状態でもよく振動する図 3 に示す A, B, C, D, E の 5 地点を考えた。今回はスピーカで与えた振動が確実に装置上の物体の影響を受けるように、予備実験を通じてスピーカの位置には A 地点を、マイクの位置には C 地点をそれぞれ用い、それに加え、マイクの位置と数の影響を考察するため B 地点を追加のマイク位置として用いた。スピーカ A お

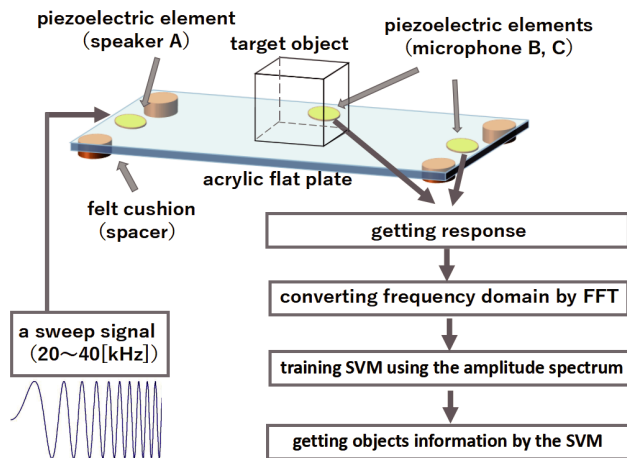


図 2 システム概要
Fig. 2 System over view.

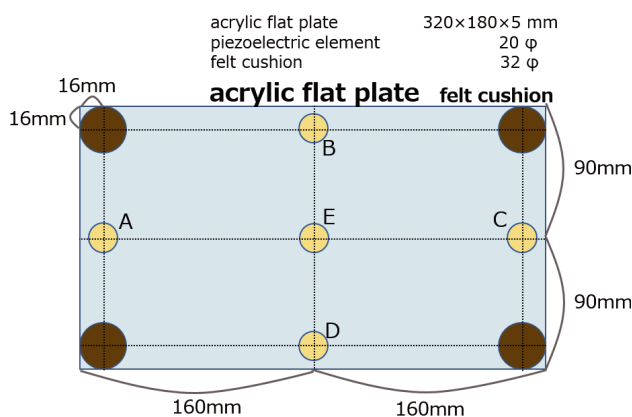


図 3 よく振動すると考えられるアクリル製平板上の位置 (A~E)
Fig. 3 The positions vibrating well (A-E).

よびマイク B, C としてこの A 地点, B 地点, C 地点に直径 20 mm, 厚さ 0.3 mm の円盤型のユニモルフ式 piezo 素子を瞬間接着剤を用いて取り付けました。スイープ信号としては, 1.0 秒間に周波数が 20,000 Hz から 40,000 Hz まで一定スピードで増加する正弦波を用いる。今回用いたこの周波数帯域は, 音声人が聞こえないこと, 環境ノイズに強いことから [20] と同様のものを採用した。信号の送受信は, TASCAM 社製のオーディオインタフェース (US-4x4-SC) を経由して, サンプル周波数 96,000 Hz で行う。このとき, オーディオインタフェースを介して piezo 素子と送受信できる信号に含まれる最大周波数はサンプル定理より 48,000 Hz となる。オーディオインタフェースを介して piezo 素子に印可される信号の振幅は 0.26 V であった。piezo マイクから取得した振動応答からフレームサイズが 512 点のハミング窓を用いて順次データを抽出し, FFT を施す。ここでハミング窓の重複幅は 496 点とする。得られた周波数スペクトルのうち 20,000 Hz から 40,000 Hz の各周波数 (107 種類) の振幅スペクトルの最大値をそれぞれ取得し, これを物体の周波数特性として用いる。スイープ信号の長さおよび FFT の窓長, 信号の振幅は予備実験よ

り決定した。なお提案手法ではアクリル製平板に振動を伝搬させるが, 実験全体を通して振動によって装置上の物体の移動, 接触面の摩耗, 物体の振動音といった現象は観測されなかった。

また板の検討として, 厚みによる物体認識への影響を確認する予備実験を行った。1 mm, 3 mm, 5 mm の厚みのアクリル製平板を用いてそれぞれシステムを作成し, 装置中央に 1g のおもりを置いた際の周波数特性の変化を評価したが, どの厚みにおいても有意に周波数特性が変化することが確認された。この結果から今回行う物体識別において板の厚みの影響は小さいと考えられ, より丈夫な厚さ 5 mm のものを用いた。

4. 物体の識別と位置推定

4.1 単一物体の識別と位置推定

アクリル製平板上に物体が 1 つだけ置かれている場合に, 提案手法によってその物体の種類および位置を取得できるか確認する。

4.1.1 単一物体識別

本実験では, アクティブ音響センシングを用いて, [1] を参考に選択した図 4 に示す 26 種類の日常的に使用する物体のうち 1 つをアクリル製平板に置き, その物体を識別し, その精度を評価する。識別は, 機械学習の 1 つである Support Vector Machine (SVM) で行う。アクリル製平板に置かれた単一物体の種類をラベルとし, そのときの周波数特性をそれぞれ SVM で学習させ, 交差検証の結果からその識別精度を評価する。ここで SVM は線形カーネル ($C = 100$) のものを用い, 以後実験で用いる SVM はすべて同様のものである。

まず, アクリル製平板上に何も置かれていない状態で, スピーカ A よりスイープ信号を送信し, マイク B およびマイク C でその応答を取得する。次に, 26 種類の各日常物体について, アクリル製平板の中心 (図 3 中 E 地点) に物体の中心が一致するように置き, それぞれ応答を取得する。これを 12 回繰り返す測定を 1 セッションとし, セッションごとに気温, 湿度といった環境条件が異なるように 1 日 2 度計 3 日間行い, 合計 1,944 個の周波数特性を取得する。セッションごとの環境パラメータを表 1 に示す。得られた周波数特性とラベルの組をランダムに 6 つのグループに分割し, 交差検証を行う。6 つのグループの中から 1 グループを選択し, 残り 5 グループを訓練データとして SVM に学習させる。学習させた SVM を用いて, 選択したグループを識別し, その識別率を算出する。この試行を全グループに対して行い, 識別率の平均を求める。また, 周りの環境条件に対するシステムのロバスト性を評価するため, ランダムに 6 つのグループに分割するのではなくセッション単位で分割した場合でも交差検証を行い, その識別率を評価する。ここで識別率はマイク B の結果を用いる場合, マ

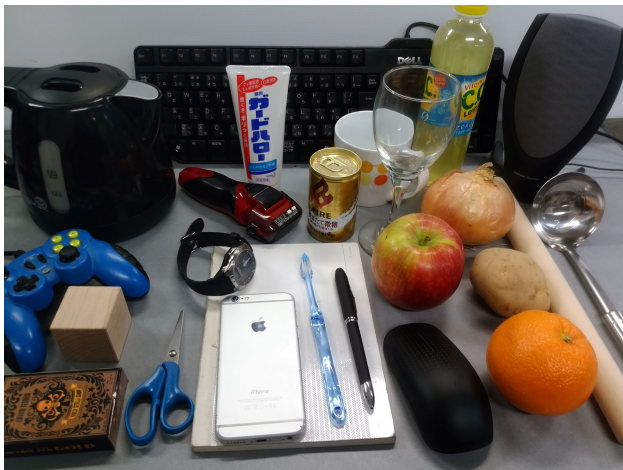


図 4 識別に用いた 26 種類の日常的に使用する物体. a deck of trump, apple, building block, canned drinks, electric kettle, electric razor, gamepad, glass, glass (water 100 ml), iPhone6, keyboard, ladle, mouse, mug, notebook, onion, orange, pen, PET bottled drinks, rolling pin, scissors, speaker, potato, toothbrush, toothpaste, wrist watch

Fig. 4 26 kinds of everyday objects used in our experiments.

表 1 セッションごとの環境パラメータ

Table 1 Environmental parameters in each session.

セッション No.	気温 [°C]	湿度 [%]
1	26.2	61
2	26.2	64
3	24.9	65
4	22.5	61
5	22.6	63
6	23.5	56

表 2 単一物体識別の結果

Table 2 Results of accuracy about everyday objects.

	microphone B	microphone C	microphone B + microphone C
random session	95.5 %	98.2 %	98.9 %
session	74.7 %	87.9 %	87.1 %

イク C の結果を用いる場合、マイク B の結果およびマイク C の結果を用いる場合の 3 種類でそれぞれ求める。

結果と考察

各交差検証で求めた識別率を表 2 に示す。また、詳細な識別結果として、マイク B およびマイク C から得られた周波数特性を用いて行った交差検証において発生した誤識別を表 3 に示す。表 3 に掲載されていない物体については識別率は 100%であった。表 3 では、たとえば、TrueLabel 列に示す何も置かれていない状態 (air) を識別させた際の識別率が 98.6%で、その際 FalseLabel 列に示すように a deck of trump に 1.4%誤識別されたことを示している。本実験において最も精度よく識別できたのは notebook と iPhone6 の 2 種類であり、これらは他のラベルに誤識別さ

れることも他のラベルがこのラベルであると誤識別されることもなかった。

セッション単位での分割ではランダムな分割と比較して識別精度が低いことから、同じ物体であってもセッション間で測定値に差がある、つまり測定環境が影響を及ぼしていることが確認された。この差を生み出すものとして外気温があげられる。一般に物体はその温度によって剛性が変化するため、この剛性の変化が周波数特性の変化となって観測されたと考えられる。したがって温度変化を観測し、それを補正することで精度を改善することができると考えられる。用いるマイクによる識別率への影響としては、マイク C を用いる場合とマイク B およびマイク C を用いる場合が同等で、マイク B を用いる場合はそれらより劣る結果となった。マイク B から得られる周波数特性は、それ単体である程度識別を行うことが可能だが、マイク C から得られる周波数特性と組み合わせても物体識別において情報量が増えないことが示唆された。ただし、これは今回物体を配置した場所がスピーカ A とマイク C の中心 (図 3 中の E 地点) であったことが影響していると考えられ、今後アクリル製平板全面における物体識別実験を通してマイク位置を検討する必要がある。notebook と iPhone6 が精度良く識別できた原因としては、アクリル製平板に接触する面の形状が平らであったことが考えられる。接触面の形状が平らな場合、アクリル製平板に触れる面積が大きくなるので、その周波数特性に与える影響も大きくなると考えられる。表 3 より、空のグラス (glass) と水の入ったグラス (glass (water 100 ml)) は、ランダムな分割ではともに 95.8%、セッション別の分割ではそれぞれ 79.2%、76.4%の識別率であり、重量のみが異なる場合においても識別できる可能性が示唆された。しかしセッション単位の分割において水の入ったグラスを空のグラスとする誤識別が 22.2%起っており、この対策としては、両方とも glass のラベルで学習させて識別を行い、glass と識別された後、周波数特性から細かく水量を取得するような変更が考えられる。今回は配置された物体ごとに装置の周波数特性を SVM を用いて学習させたのみだが、より詳しく接触面積や重量が与える周波数特性への影響を観測して数理モデルを作成することで、物体の識別だけでなく接触面積や重量の取得が可能になると考えられる。本実験を通して提案手法によって物体識別が可能であることが示されたが、1 品目あたり 1 個でしか評価していないことは、物体識別を実現するうえで限定的な条件であったと考えられる。特に apple や onion といった農作物では、品目に対して形状や大きさが固定ではなく、今回の実装では同一品目の別の物体を識別することは困難である。今後は 1 品目あたりの物体個数を増やして実験を行い、得られた結果から学習アルゴリズムを検討する必要がある。

表 3 単一物体識別の誤識別一覧
Table 3 Misrecognition list of everyday objects.

	True Label	[%]	False Label	[%]		
ランダム分割 識別率=98.9%	air	98.6	a deck of trump	1.4		
	electric kettle	95.8	PET bottled drinks	2.8	mouse	1.4
	electric razor	95.8	apple	1.4	onion	1.4
	glass	95.8	canned drinks	1.4	glass (water 100 ml)	1.4
	glass (water 100 ml)	95.8	glass	2.8	canned drinks	1.4
	mouse	98.6	rolling pin	1.4		
	mug	98.6	glass (water 100 ml)	1.4		
	orange	98.6	onion	1.4		
	rolling pin	95.8	apple	1.4	toothbrush	1.4
	scissors	98.6	wrist watch	1.4		
セッション単位分割 識別率=87.1%	air	98.6	a deck of trump	1.4		
	apple	95.8	toothpaste	4.2		
	canned drinks	84.7	building block	11.1	apple	2.8
	electric kettle	90.3	PET bottled drinks	4.2	apple	1.4
			rolling pin	1.4	toothbrush	1.4
	electric razor	70.8	apple	16.7	mouse	4.2
			rolling pin	2.8	toothpaste	1.4
	glass	79.2	glass (water 100 ml)	11.1	canned drinks	8.3
	glass (water 100 ml)	76.4	glass	22.2	canned drinks	1.4
	keyboard	98.6	apple	1.4		
	ladle	86.1	rolling pin	13.9		
	mouse	98.6	rolling pin	1.4		
	mug	90.3	canned drinks	4.2	glass (water 100 ml)	2.8
			electric razor	1.4		
	onion	66.7	apple	16.7	toothpaste	16.7
	orange	97.2	electric kettle	1.4	gamepad	1.4
	pen	97.2	air	2.8		
	PET bottled drinks	98.6	apple	1.4		
	potato	76.4	apple	13.9	onion	5.6
			pen	1.4	wrist watch	1.4
rolling pin	70.8	electric razor	16.7	air	5.6	
		onion	2.8	toothbrush	1.4	
scissors	70.8	air	16.7	wrist watch	11.1	
speaker	83.3	apple	6.9	electric razor	4.2	
		electric kettle	1.4	ladle	1.4	
toothbrush	79.2	air	18.1	scissors	2.8	
toothpaste	66.7	apple	33.3			
wrist watch	79.2	air	12.5	electric razor	6.9	
				toothbrush	1.4	

4.1.2 位置推定

本項では、アクリル製平板上に置かれている物体の位置を 5mm 単位で移動させ、アクティブ音響センシングで位置を取得できるかを検証する。配置物体の位置による周波数特性の変化の再現性を評価するため、位置の取得は 4.1.1 項同様、SVM による識別で行う。実験に用いる物体としては、4.1.1 項の物体識別に用いた物体のうち、最も構造がシンプルである building block (木製で一辺 44mm の立方体) を採用した。アクリル製平板上に置かれた物体の位置をラベルとし、そのときの周波数特性をそれぞれ SVM で学習させ、交差検証の結果からその識別精度を評価する。

板上に何も置かれていない状態で、スピーカ A からスイープ信号を送信し、マイク B およびマイク C でその応答を取得する。図 3 で左側に示されるアクリル製平板の端から 30mm の位置に中心が一致するように building block を置き、同様に応答を測定する。そこから 5mm 刻みで、端から 35mm, 40mm, ..., 285mm, 290mm の計 53 カ所の地点にそれぞれ building block を置いて、それぞれの位置に対して同様に周波数特性を測定する。アクリル製平板上に何も置かれていない状態と building block を各地点に置いた状態での測定を各 12 回行う。得られた 648 個の周波数特性と状態の組をランダムに 6 つのグループに分割す

表 4 位置推定の結果

Table 4 Results of position detection.

	microphone B	microphone C	microphone B + microphone C
Accuracy [%]	85.5	85.5	89.2
Absolute error [mm]	0.825	2.78	0.605
Standard deviation [mm]	0.534	4.81	0.409

る。その中から1グループを選択し、残り5グループを訓練データとしてSVMに学習させる。学習させたSVMを用いて、選択したグループを識別し、その識別率を算出する。この試行を全グループに対して行い、識別率の平均を求める。ここで識別率はマイクBの結果を用いる場合、マイクCの結果を用いる場合、マイクBの結果およびマイクCの結果を用いる場合の3種類でそれぞれ求める。また、得られた識別結果を用いて絶対誤差の平均と分散を求める。**結果と考察**

各交差検証で得られた結果を表4に示す。何も置かれていない状態の識別率はすべての交差検証において100%であった。特にマイクBおよびマイクCから得られる周波数特性を用いた場合の平均絶対誤差が0.605mmと用いた物体サイズ44mm四方と比較して非常に小さいことから、本実験を通して提案手法による位置推定の可能性が示唆された。マイク位置による物体位置推定精度への影響としては、絶対誤差に与える影響が大きく、マイクBを用いる場合とマイクBおよびマイクCを用いる場合が同等で、マイクCを用いる場合はそれらより劣る結果となった。マイクCを用いる場合の位置取得精度が低い理由として、85mm地点を235mm地点に、215mm地点を110mm地点に誤識別するなど、位置を大きく誤る誤識別がいくつか見られたことがあげられる。この誤識別の原因として周波数特性の対称性があげられる。図5にマイクBおよびマイクCから得られた周波数特性と物体位置の関係を示す。縦軸が周波数、横軸は図3で左側に示されるアクリル製平板の端から見たbuilding blockの位置であり、色の濃さは周波数ごとに最大最小値で正規化した振幅の大きさを表す。マイクCから得られた振幅スペクトルは、building blockの位置がアクリル製平板の中心である160mmの位置を軸に対称になっている様子が見られた。このため、アクリル製平板の中心に関して対称の位置への誤識別が発生し、位置の識別を大きく誤ることが考えられる。

また、位置情報は連続値であるため、SVMなどのclassifierを用いると、取得する位置の最小単位に応じて事前に学習させておく手間が増大するといった問題が発生する。そのため今後は、配置物体の位置による周波数特性の変化を詳細に観察するとともに、振動工学の見地より周波数特性から位置を算出できるような定量的なモデルを検討する必要がある、今後の課題である。

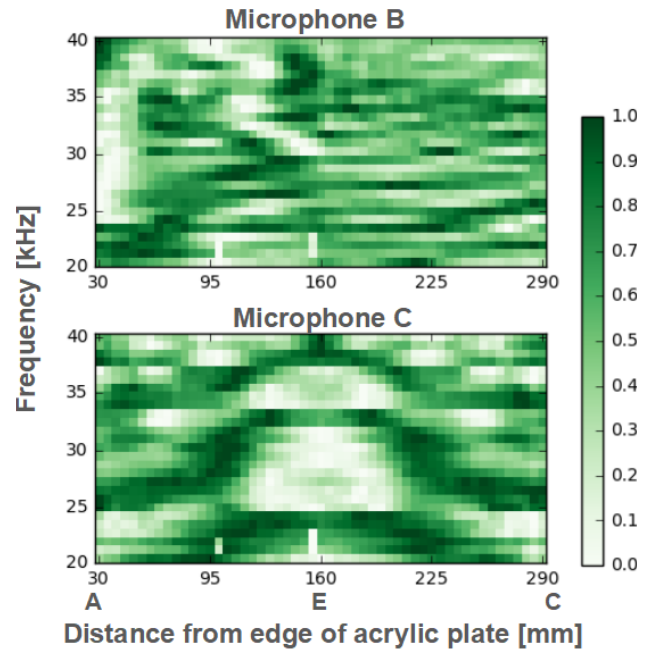


図 5 周波数特性と building block の位置の関係

Fig. 5 The relationship between frequency characteristics and the block position.

4.2 複数物体識別

4.2.1 単一物体識別データによる複数物体データの推定

これまでに述べたように、提案システムによって単一物体の種類およびその位置は、それぞれスイープ信号に影響を与え、その特徴量を用いて種類とその位置を識別できることを示した。ここでは複数の物体を同時に識別する手法について検討する。ここで、特定の特徴量を用いて複数の物体が置かれている状態を学習させようとする場合、物体の組合せに対し学習データを用意する必要があるが、識別する物体の種類が増えると、物体が置かれる組合せの数は指数関数的に増加し、学習データを用意することが非常に困難になるという問題が発生する。そこで振動工学的な観点から複数物体が各位置に同時に置かれている状態の周波数特性を、何も置かれていない状態で得られる周波数特性と個々の物体が1つだけ置かれた状態で得られる周波数特性から推定する手法について検討する。

まず次のように状態を定める。

状態0：何も置かれていない状態。

状態1：物体1がある位置に置かれた状態。

状態2：物体2がある位置に置かれた状態。

状態 3 : 状態 1, 2 と同じ位置に物体 1, 2 が同時に置かれた状態.

一般に, 外力によって物体を振動させる場合, その様子は次の運動方程式で記述できる [22].

$$MU'' + CU' + KU = P \quad (1)$$

ここで M, C, K はそれぞれ質量マトリックス, 減衰マトリックス, 剛性マトリックスであり, U は変位ベクトル, P は外力ベクトルである.

上記各状態 i ($i = 0, 1, 2, 3$) においてパラメータ M_i, C_i, K_i は変化しないと考えられるので, 式 (1) を用いて各状態に対し運動方程式を求めると以下ようになる.

$$MU_0'' + CU_0' + KU_0 = I \quad (2)$$

$$MU_1'' + CU_1' + KU_1 = I + W_1 \quad (3)$$

$$MU_2'' + CU_2' + KU_2 = I + W_2 \quad (4)$$

$$MU_3'' + CU_3' + KU_3 = I + W_1 + W_2 \quad (5)$$

ここで U_i は状態 i においてスイープ信号を与えた際の物体の変位ベクトルを示す. スイープ信号は時間的に周波数が増える信号なのでスイープ信号を与えた際の変位は直接周波数特性と対応する. I はスイープ信号を与える外力, W_i は状態 i における物体 i による外力を示し, $W_3 = W_2 + W_1$ である. (式 (4)) + (式 (3)) - (式 (2)) において微分の線形性を考慮すると

$$M(U_2 + U_1 - U_0)'' + C(U_2 + U_1 - U_0)' + K(U_2 + U_1 - U_0) = I + W_1 + W_2 \quad (6)$$

となり, 式 (5) と式 (6) を比較すると

$$U_3 = U_2 + U_1 - U_0 \quad (7)$$

となる. つまり物体 1 と 2 が特定の位置に置かれた状態のスイープ信号による変位は, 物体 1 が置かれた状態のスイープ信号による変位に, 物体 2 が置かれた状態のスイープ信号による変位を加えた値から, 何も置かれていない状態のスイープ信号による変位を減らすことで求められる. 物体が 3 個以上の場合も同様にする事で, n 個の物体 (物体 $n, n-1, \dots, 1$) がある位置に同時に置かれた場合の変位 U_{n+1} は

$$U_{n+1} = \sum_{i=1}^n U_i - (n-1)U_0 \quad (8)$$

となる. したがって式 (8) を用いることで, 複数の物体が置かれた状態のデータを個々の物体が置かれた状態のデータから算出できると考えられる.

4.2.2 複数物体の識別評価

評価実験として, ここでは, 式 (8) より算出される複数の物体が特定位置に置かれた状態の推定値を用いて識別器を作成し, 実際に複数の物体が特定位置に置かれた状態の実測値を用いて識別を行い, その精度を評価する. 識別は, 4.1.1 項同様, SVM で行う. 識別に用いる位置は図 3 で左側に示されるアクリル製平板の端から 30 mm, 160 mm, 290 mm の 3 種類とし, 識別に用いる物体の組は 4.1.1 項で用いた物体のうち, 学習机に置かれる実物体を想定した mug, notebook, pen と, 冷蔵庫に格納される実物体を想定した apple, onion, orange の 3 種類 2 パターンとした. マイクは 4.1.1 項, 4.1.2 項の結果をふまえてマイク B とマイク C の両方を用いる.

まず, アクリル製平板に何も置かれていない状態で, スピーカ A からスイープ信号を送信し, マイク B およびマイク C でその応答を取得する. 次にそれぞれのパターンの 3 種類の物体を 3 種類の位置に配置する組合せ 33 通り ($= {}_3C_1 \times {}_3P_1 + {}_3C_2 \times {}_3P_2 + {}_3C_3 \times {}_3P_3$) について, 各状態でそれぞれ応答を取得する. この試行を 12 回繰り返す. 取得したデータに対し, 図 6 に概要を示した次の 4 種類の解析を行う. 解析 1 として, 物体が 1 つだけ置かれた状態のデータに対して 6 分割の交差検証を行い単体での識別精度を求める. これは 4.1.1 項で述べた単一の物体識別と同じ実験である. 解析 2 として, 取得したデータのうち, 何も

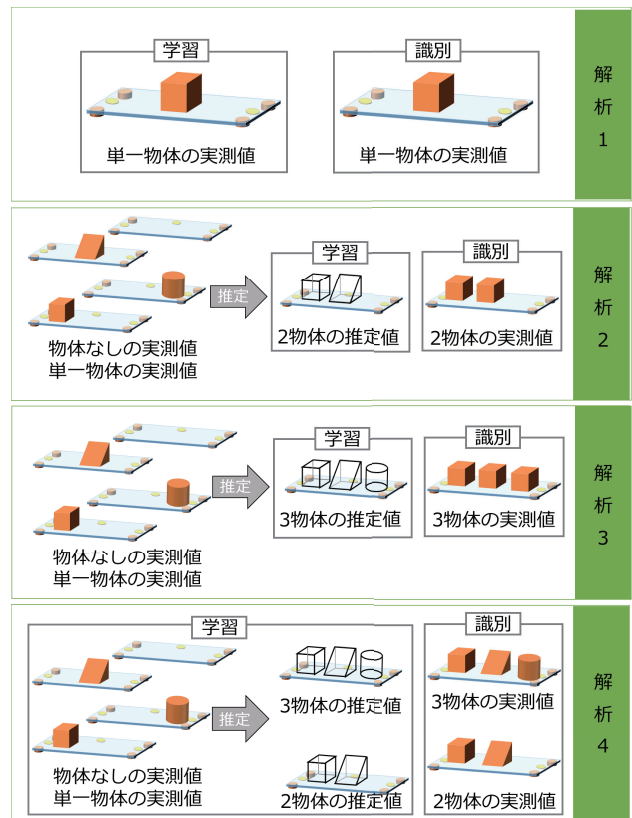


図 6 複数物体識別における解析概要

Fig. 6 Summary of analysis in multiple objects recognition.

表 5 学習機パターンにおける各解析の誤識別一覧

Table 5 Misrecognition list of multiple objects in the desk pattern.

	True Label	[%]	False Label	[%]	
解析 1 識別率=98.1%	mug — —	91.7	— — mug	8.3	
	notebook — —	91.7	— pen —	8.3	
解析 2 識別率=96.8%	mug notebook —	91.7	pen notebook —	8.3	
	notebook — mug	91.7	notebook mug —	8.3	
	— notebook mug	83.3	— notebook pen	16.7	
	notebook — pen	91.7	notebook pen —	8.3	
	— notebook pen	83.3	pen notebook —	16.7	
解析 3 識別率=100%					
解析 4 識別率=91.7%	mug notebook —	91.7	pen notebook —	8.3	
	— mug notebook	91.7	— — notebook	8.3	
	notebook — mug	75.0	notebook pen mug	16.7	notebook mug — 8.3
	— notebook mug	83.4	pen notebook mug	8.3	— notebook pen 8.3
	notebook — pen	91.7	notebook pen —	8.3	
	— notebook pen	83.4	pen notebook —	8.3	— notebook — 8.3
	mug notebook pen	58.4	mug notebook —	33.3	— notebook pen 8.3
	mug pen notebook	83.3	mug — notebook	16.7	
	notebook pen mug	83.3	notebook — mug	16.7	
	notebook mug pen	83.4	notebook — pen	8.3	notebook mug — 8.3
	pen mug notebook	91.7	pen — notebook	8.3	
	pen notebook mug	83.3	pen notebook —	16.7	

表 6 冷蔵庫パターンにおける各解析の誤識別一覧

Table 6 Misrecognition list of multiple objects in the refrigerator pattern.

	True Label	[%]	False Label	[%]
解析 1 識別率=100%				
解析 2 識別率=98.6%	— onion orange	91.7	— apple orange	8.3
	orange apple —	91.7	apple — orange	8.3
	orange — apple	91.7	apple — orange	8.3
解析 3 識別率=98.6%	orange onion apple	91.7	orange apple onion	8.3
解析 4 識別率=98.3%	— onion orange	91.7	— apple orange	8.3
	orange apple —	91.7	apple — orange	8.3
	orange — apple	91.7	apple — orange	8.3
	onion orange apple	91.7	onion orange —	8.3
	orange onion apple	91.7	orange apple onion	8.3

置かれていない状態のデータと物体が1つだけ置かれた状態のデータを用いて物体が2個置かれた状態18通りの推定値を式(8)に基づいて算出し、これをSVMに学習させ、取得した各状態の実測値の識別精度を求める。解析3として、解析2と同様の解析を物体が3個置かれた状態6通りに対して行う。解析4として、解析2および解析3で求めた推定値、何も置かれていない状態、物体が1つだけ置かれた状態のデータを用いて識別器を作成し、物体が2個および3個置かれた状態の実測値を識別し、識別精度を求める。この実験を学習機パターンおよび冷蔵庫パターンに対して行う。

結果と考察

各解析において発生した誤識別をまとめたものを表5および表6に示す。表の見方は表3と同様である。ラベルは左から順に位置と対応しており、たとえば、notebook mug —はアクリル製平板の端から30mmの位置にnotebookが置かれ、160mmの位置にmugが置かれ、290mmの位置には何も置かれていないことを表す。学習機パターンでの識別率はそれぞれ、解析1の場合98.1%、解析2の場合96.8%、解析3の場合100%、解析4の場合91.7%であり、冷蔵庫パターンでの識別率はそれぞれ、解析1の場合100%、解析2の場合98.6%、解析3の場合98.6%、解析4

の場合 98.3%であった。

この実験において、解析 4 で得られた識別率は学習機パターンで 91.7%，冷蔵庫パターンで 98.3%と用いた物体によって差があることが分かる。学習機パターンで最も識別率が低かった組合せは mug notebook pen であり、その 33.3%が mug notebook に誤識別されている。このような、notebook より右側に置いた物体の識別を誤った結果はいくつか見られ、この原因としてはスピーカから出た信号が、接触面積の大きい notebook の接触により減衰し、それより右の範囲における物体の影響が弱まることが考えられる。今回は用いた振動の強度に関して十分な議論がなされていないため、今後強度と識別率の関係を調べる必要がある。また、学習機パターンではすべての誤識別は notebook と他の物体が共に置かれた場合に発生しており、その原因として、notebook が装置の周波数特性に与える影響が大きいことが考えられる。notebook は pen や mug と比べ、接触面積が大きく、実際 4.1.1 項においても notebook は識別しやすいクラスとしてあげられている。そのため notebook の測定誤差が pen や mug による周波数特性の変化と比べて大きくなりやすくなり、誤識別が多く発生したと考えられる。notebook と pen のような、装置の周波数特性への影響に差がある 2 物体を同時に識別する際の誤識別を解決するためには、周波数特性の測定誤差を減少させることが重要であり、これは今後物体位置の数理モデル化を行うことで解決できると考えられる。今回発生した測定誤差の多くは、装置に物体を置く際の位置の誤差が原因だと考えられるため、モデルから位置の影響を補正することで測定誤差が減少し、識別の精度が向上すると考えられる。本実験を通して、複数物体の種類と位置の識別が単一物体識別データを用いて、可能であることが示された。位置については 3 カ所でしか試行していないが、4.1.2 項の結果から 5 mm 間隔であっても位置の差によって周波数特性が変化することが確認できているため、同程度までは識別可能であると考えられる。ただし、物体どうしが接触したり、重なったりした場合の識別については今後の課題である。

5. 応用と今後の課題

提案手法は暗所や物体どうしのオクルージョンに強く、環境変化の小さい場所における物体識別に適しており、適用例としては冷蔵庫があげられる。冷蔵庫の内部は往々にして密集度が高いためオクルージョンが発生しやすく、普段ライトが消えていることもあり、カメラによる物体認識には不向きである。提案手法によって棚板に置かれた食材を認識することができれば庫内の自動在庫管理が実現できる。加えて、食材の重量を認識することができれば、食品の残量が分かるとともに食品の消費量が分かるので、家庭全体の健康管理なども実現できる。またこれらの認識情報と位置情報と組み合わせることで、外出時に足りない食材

の買い足しを提案するなどの行動推薦システムを構築することも可能になると考えられる。一方で、その実用化のためには、冷蔵庫に入れた物体が冷やされていく際の温度変化による応答への影響や、野菜や果物など形や重さが不均一な物体をどのように認識するのかなど、解決すべき問題も存在する。

一方今後の課題としては、まず 4.1.1 項で述べたように、アクリル製平板全面における物体識別実験を通じたマイク位置の検討や、ガラス内の水量認識用のアルゴリズムの構築が必要である。また、1 品目あたりの物体個数を増やして実験を行い、得られた結果から学習アルゴリズムを検討する必要がある。その後、アクリル製平板の周波数特性に影響を与えると考えられる、配置物体の位置や重さ、底面積、底面形状、角度といった物体パラメータや外気温といった環境パラメータの影響を個々に観測し、数理モデルを構築することで、それらの連続値での推定や識別の際の補正に用いることで精度の向上を目指す。また複数物体については、物体や位置の種類を増やした際の評価や、物体どうしが接触したり、重なったりした場合の識別手法の検討が今後の課題となっている。

6. おわりに

本稿では、アクティブ音響センシングを用いた物体情報識別手法を提案した。また、手法の検討のため、アクリル製平板、 piezo 素子、オーディオインタフェースなどからなるセンシング装置のプロトタイプを実装し、実装したプロトタイプを用いて単一物体の種類やその位置がそれぞれ 98.9%，89.2%の精度で取得できることを確認した。その後複数物体の同時識別について検討を行い、推定式を用いた複数物体の同時識別に関する評価実験を行った。実験の結果、学習機の上を想定した物体の組および冷蔵庫内を想定した物体の組を配置する組合せに関して、それぞれ 91.7%および 98.3%の精度で識別できた。本実験を通して単一の物体識別データを用いた複数物体の同時識別可能性が示された。今後はアクリル製平板の周波数特性に影響を与えると考えられる配置物体のパラメータや環境パラメータの影響を個々に観測し、数理モデルを構築することで、それらの連続値での推定や識別の際の補正に用いることで精度の向上を目指す。

謝辞 本研究の一部は JSPS 科研費 JP16H02891 の助成を受けたものである。

参考文献

- [1] Yeo, H.-S., Flamich, G., Schrempf, P., Harris-Birtill, D. and Quigley, A.: RadarCat: Radar categorization for input & interaction, *Proc. 29th Annual Symposium on User Interface Software and Technology, UIST '16*, pp.833–841, ACM (online), DOI: 10.1145/2984511.2984515 (2016).

- [2] Laput, G., Brockmeyer, E., Hudson, S.E. and Harrison, C.: Acoustruments: Passive, acoustically-driven, interactive controls for handheld devices, *Proc. 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, pp.2161–2170, ACM (online), DOI: 10.1145/2702123.2702414 (2015).
- [3] Baudisch, P., Becker, T. and Rudeck, F.: Lumino: Tangible Building Blocks Based on Glass Fiber Bundles: Invited Demo, *ACM International Conference on Interactive Tabletops and Surfaces, ITS '10*, pp.299–299, ACM (online), DOI: 10.1145/1936652.1936727 (2010).
- [4] Villar, N., Cletheroe, D., Saul, G., Holz, C., Regan, T., Salandin, O., Sra, M., Yeo, H.-S.Y., Field, W. and Zhang, H.: Project Zanzibar: A portable and flexible tangible interaction platform, *Proc. 2018 CHI Conference on Human Factors in Computing Systems*, p.515, ACM (2018).
- [5] Amento, B., Hill, W. and Terveen, L.: The Sound of One Hand: A wrist-mounted bio-acoustic fingertip gesture interface, *CHI '02 Extended Abstracts on Human Factors in Computing Systems, CHI EA '02*, pp.724–725, ACM (online), DOI: 10.1145/506443.506566 (2002).
- [6] Murray-Smith, R., Williamson, J., Hughes, S. and Quaade, T.: Stane: Synthesized surfaces for tactile input, *Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI '08*, pp.1299–1302, ACM (online), DOI: 10.1145/1357054.1357257 (2008).
- [7] Harrison, C. and Hudson, S.E.: Scratch Input: Creating large, inexpensive, unpowered and mobile finger input surfaces, *Proc. 21st Annual ACM Symposium on User Interface Software and Technology, UIST '08*, pp.205–208, ACM (online), DOI: 10.1145/1449715.1449747 (2008).
- [8] Harrison, C., Xiao, R. and Hudson, S.: Acoustic Barcodes: Passive, durable and inexpensive notched identification tags, *Proc. 25th Annual ACM Symposium on User Interface Software and Technology, UIST '12*, pp.563–568, ACM (online), DOI: 10.1145/2380116.2380187 (2012).
- [9] Paradiso, J.A., Leo, C.K., Checka, N. and Hsiao, K.: Passive acoustic knock tracking for interactive windows, *CHI '02 Extended Abstracts on Human Factors in Computing Systems, CHI EA '02*, pp.732–733, ACM (online), DOI: 10.1145/506443.506570 (2002).
- [10] Harrison, C., Tan, D. and Morris, D.: Skininput: Appropriating the skin as an interactive canvas, *Comm. ACM*, Vol.54, No.8, pp.111–118 (online), DOI: 10.1145/1978542.1978564 (2011).
- [11] Harrison, C., Schwarz, J. and Hudson, S.E.: TapSense: Enhancing finger interaction on touch surfaces, *Proc. 24th Annual ACM Symposium on User Interface Software and Technology, UIST '11*, pp.627–636, ACM (online), DOI: 10.1145/2047196.2047279 (2011).
- [12] Lopes, P., Jota, R. and Jorge, J.A.: Augmenting touch interaction through acoustic sensing, *Proc. ACM International Conference on Interactive Tabletops and Surfaces, ITS '11*, pp.53–56, ACM (online), DOI: 10.1145/2076354.2076364 (2011).
- [13] Hwang, S., Bianchi, A. and Wohn, K.: MicPen: Pressure-sensitive pen interaction using microphone with standard touchscreen, *CHI '12 Extended Abstracts on Human Factors in Computing Systems, CHI EA '12*, pp.1847–1852, ACM (online), DOI: 10.1145/2212776.2223717 (2012).
- [14] Wylie, C. and Mackie, R.R.: Active sonar target detection and reporting: Perceived consequences and their effects on performance, Technical Report, DTIC Document (1972).
- [15] Brenner, M.C. and Fitzgibbon, J.J.: Surface acoustic wave touch panel system, *The Journal of the Acoustical Society of America*, Vol.84, No.4, pp.1578–1578 (1988).
- [16] Gupta, S., Morris, D., Patel, S. and Tan, D.: SoundWave: using the doppler effect to sense gestures, *Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI '12*, pp.1911–1914, ACM (online), DOI: 10.1145/2207676.2208331 (2012).
- [17] Takemura, K., Ito, A., Takamatsu, J. and Ogasawara, T.: Active bone-conducted sound sensing for wearable interfaces, *Proc. 24th Annual ACM Symposium Adjunct on User Interface Software and Technology, UIST '11 Adjunct*, pp.53–54, ACM (online), DOI: 10.1145/2046396.2046419 (2011).
- [18] Collins, T.: Active acoustic touch interface, *Electronics letters*, Vol.45, No.20, pp.1055–1056 (2009).
- [19] Laput, G., Brockmeyer, E., Mahler, M., Hudson, S.E. and Harrison, C.: Acoustruments: Passive, acoustically-driven, interactive controls for handheld devices, *ACM SIGGRAPH 2015 Emerging Technologies, SIGGRAPH '15*, p.3:1, ACM (online), DOI: 10.1145/2782782.2792490 (2015).
- [20] Ono, M., Shizuki, B. and Tanaka, J.: Touch & Activate: adding interactivity to existing objects using active acoustic sensing, *Proc. 26th Annual ACM Symposium on User Interface Software and Technology, UIST '13*, pp.31–40, ACM (online), DOI: 10.1145/2501988.2501989 (2013).
- [21] Liu, J., Wang, C., Chen, Y. and Saxena, N.: VibWrite: Towards finger-input authentication on ubiquitous surfaces via physical vibration, *Proc. 2017 ACM SIGSAC Conference on Computer and Communications Security*, pp.73–87, ACM (2017).
- [22] Agar, T.: The analysis of aerodynamic flutter of suspension bridges, *Computers & Structures*, Vol.30, No.3, pp.593–600 (1988).



岩瀬 大輝

2017年大阪大学工学部電子情報工学科卒業。現在、同大学大学院情報科学研究科修士課程在学中。主に、タンジブルユーザインタフェースの研究に従事。IEEE学生会員。



伊藤 雄一 (正会員)

1998年大阪大学工学部卒業。2000年同大学大学院工学研究科博士前期課程修了。2002年同研究科博士後期課程中退と同時に同研究科助手。情報科学研究科助教を経て、2008年より大阪大学クリエイティブユニット准教授。

主に、ヒューマンコンピュータインタラクションの研究に従事。



秦 秀彦

Panasonic 株式会社イノベーション推進部門要素技術開発センターに勤務。2015 年より大阪大学大学院情報科学研究科博士後期課程在学中。



尾上 孝雄

1993 年大阪大学大学院工学研究科電子工学専攻博士前期課程修了。同大学の助手、講師、京都大学の助教授、大阪大学の助教授を経て、現在、同大学院情報科学研究科教授。画像符号化システム、音声処理システム等のシステムオンチップ設計、組み込みシステムの低消費電力化の研究に従事。