

# アクティブ音響センシングによる日常物体識別と位置推定

岩瀬 大輝<sup>1</sup> 伊藤 雄一<sup>1</sup> 秦 秀彦<sup>1</sup> 山下 真由<sup>1</sup> 尾上 孝雄<sup>1</sup>

**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 98.2%, 85.5%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 93.8%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。

## 1. はじめに

日常生活空間内には様々な物体が存在し、人々は多くの道具や機器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。例えば、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、その上に乗せら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変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、一元的な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中でも、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の取得を行う。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の推定が行えるかを確認する。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。

<sup>1</sup> 大阪大学 大学院情報科学研究科  
Graduate School of Information Science and Technology, Osaka University

## 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なった [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。また、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もたせたり、物体にマーカを搭載する必要があり、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。このように、物体情報識別手法は多く提案されているが、得られる情報が一元的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、配置物体に制限はなく、音響の伝搬はその接触状況の影響を受け、配置物体に関わるいくつもの要因で変化するという特徴をもつため、様々な物体に対して、多元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盛んになされており、それらは、スピーカ等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。Ament 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 [4]. 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 [5]. 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 [6]. 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

時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 [7]. 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 [8]. Harrison らの Skinput 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 [9]. 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 [10], Lopes らは、どの指でタップしたかを特定した [11]. 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 [12].

一方、アクティブ音響センシングは、スピーカ等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。Stewart 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 [13]. 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 [14]. 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時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 [15]. 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 [16]. 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等の平面をタッチパネルにする手法を提案した [17]. Laput らの Acoustruments は、アクティブ音響システムを利用し、携帯端末等に物理インタフェースを追加する手法を提案した [18].

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、振動を伴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挙げられる。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 [19]. 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。この研究では人間による物体の把持状態の認識のみ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システムを施した表面の上に置かれる物体の状態認識に用いる点で異なる。本研究では、アクティブ音響センシングを用いて多角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 日常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 [19]。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並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、ここから物体認識を行うことができると考えられる。

周波数特性の取得・解析には、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増加する信号を指す。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 [19]。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

システムの概要を図1に示す。物体を配置する表面として縦180mm、横320mm、厚さ5mm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、丈夫で振動を伝播させやすいことが挙げられる。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32mm、厚さ5mm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、スピーカ及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2に示すA,B,C,D,Eの5地点が考えられるが、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、予備実験を通じて今回は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た。スピーカ及びマイクとしてこのA地点、C地点に直径20mm、厚さ0.3mmの円盤型のユニモルフ式ピエゾ素子を瞬間接着剤を用いて取り付けた。

スイープ信号としては、1.0秒間に周波数が20000Hzから40000Hzまで一定スピードで増加する正弦波を用いる。この帯域の周波数を用いる利点として、音声に聞こえないこと、環境ノイズに強いことが挙げられる [19]。信

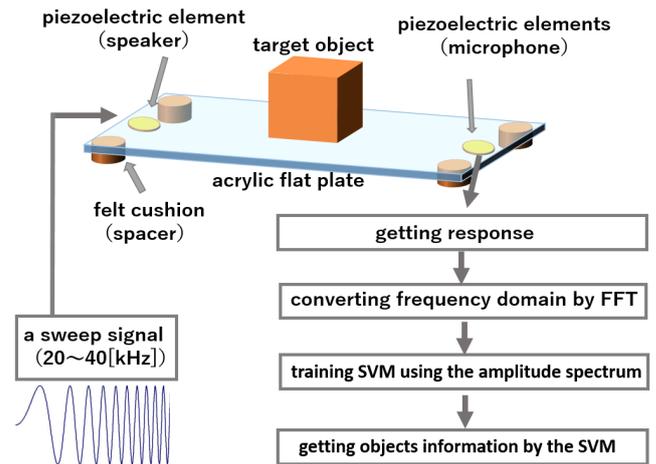


図1 システム概要

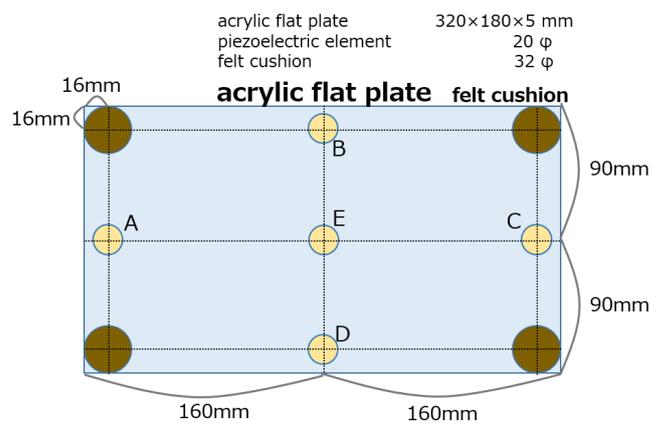


図2 よく振動すると考えられるアクリル製平板上の位置 (A~E)

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、サンプリング周波数96000 Hzで行う。このとき、オーディオインタフェースを介してピエゾ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リング定理より48000 Hzとなる。オーディオインタフェースを介してピエゾ素子に印可される信号の振幅は0.26Vであった。ピエゾ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20000Hzから40000Hz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。

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。1mm、3mm、5mm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のおもりを乗せ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5mmのものを用いた。

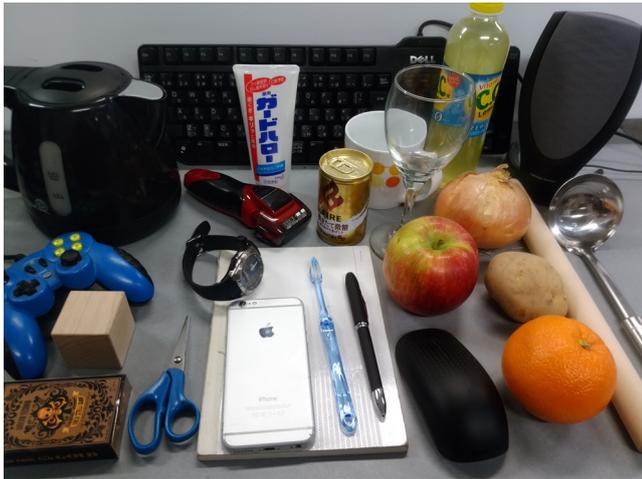


図 3 識別に用いた 26 種類の日常物体  
 a deck of trump, apple, building block,  
 canned drinks, electric kettle, electric razor,  
 gamepad, glass, glass(water 100ml), iPhone6,  
 keyboard, ladle, mouse, mug, notebook,  
 onion, orange, pen, PET bottled drinks,  
 rolling pin, scissors, speaker, potato,  
 toothbrush, toothpaste, wrist watch

## 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 つだけ乗っている場合に、提案手法によってその物体の種類及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、図 3 に示す 26 種類の日常物体のうち 1 つをアクリル製平板に乗せ、その物体を識別し、その精度を評価する。識別は、機械学習のひとつである Support Vector Machine (SVM) で行う。アクリル製平板に乗っている単一物体の種類をラベルとし、その時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。

まず、アクリル製平板に何も乗っていない状態で、図 2 の A 地点に取り付けたピエゾ素子からスイープ信号を送信し、図 2 の C 地点に取り付けたマイク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 2 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、セッション単位の測定を気温、湿度といった環境条件が異なるように 1 日 2 度 3 日間行い、合計 1944 個の周波数特性を取得する。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。

### 結果と考察

ランダムな分割による識別率は全体で 98.2%、セッション単位での分割による識別率は全体で 87.9%であった。各分割で行った交差検証において発生した誤識別を表 1 にまとめる。表 1 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 1 より、例えば、何も乗っていない状態 (air) を識別させた際の識別率が 98.6%で、その際 1.4%が a deck of trump に誤識別されたことを示している。両分割において共通して最も精度よく識別できたのは building block, notebook, iPhone6 の 3 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挙げられる。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。また、building block や notebook, iPhone6 などが精度よく識別できた原因として接触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、アクリル製平板の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。同様の特徴を有する a deck of trump も比較的高精度で識別できており、セッション別の分割においてこのクラスが誤識別されることはなかった (ただし air の 1.4%がこのクラスへの誤識別されている)。空のグラス (glass) と水の入ったグラス (glass(water 100ml)) は、ランダムな分割ではそれぞれ 97.2%, 93.1%, セッション別の分割ではそれぞれ 90.3%, 83.3%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 16.7%起こっており、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。今回は装置の周波数特性を配置された物体ごとに SVM に学習させたのみだが、更に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。

表 1 日常物体の誤識別一覧

	True Label	[%]	False Label	[%]				
ランダム分割 識別率=98.2%	air	98.6	a deck of trump	1.4				
	a deck of trump	98.6	air	1.4				
	canned drinks	97.2	electric kettle	1.4	mug	1.4		
	electric kettle	93.1	apple	1.4	keyboard	1.4	mouse	1.4
			rolling pin	1.4	speaker	1.4		
	electric razor	94.4	apple	2.8	potato	1.4	toothpaste	1.4
	glass	97.2	electric kettle	1.4	glass(water 100ml)	1.4		
	glass(water 100ml)	93.1	mug	4.2	glass	2.8		
	mouse	98.6	rolling pin	1.4				
	mug	94.4	glass(water 100ml)	2.8	canned drinks	1.4	electric razor	1.4
	onion	98.6	apple	1.4				
	pen	98.6	toothbrush	1.4				
	rolling pin	94.4	canned drinks	2.8	scissors	1.4	toothbrush	1.4
	scissors	98.6	toothbrush	1.4				
toothpaste	98.6	potato	1.4					
wrist watch	97.2	air	1.4	scissors	1.4			
セッション単位分割 識別率=87.9%	air	95.8	a deck of trump	1.4	rolling pin	1.4	toothbrush	1.4
	apple	90.3	toothpaste	5.6	onion	2.8	potato	1.4
	canned drinks	73.6	mug	12.5	glass	11.1	apple	1.4
			rolling pin	1.4				
	electric kettle	84.7	PET bottled drinks	4.2	apple	2.8	keyboard	2.8
			mouse	2.8	ladle	1.4	mug	1.4
	electric razor	81.9	rolling pin	6.9	apple	5.6	potato	5.6
	glass	90.3	canned drinks	4.2	glass(water 100ml)	2.8	gamepad	1.4
			PET bottled drinks	1.4				
	glass(water 100ml)	83.3	glass	16.7				
	ladle	91.7	rolling pin	8.3				
	mouse	87.5	electric razor	9.7	electric kettle	1.4	rolling pin	1.4
	mug	95.8	electric razor	2.8	glass(water 100ml)	1.4		
	onion	70.8	apple	20.8	potato	8.3		
	orange	98.6	apple	1.4				
	rolling pin	65.3	electric razor	23.6	air	5.6	apple	1.4
			canned drinks	1.4	scissors	1.4	toothbrush	1.4
	scissors	86.1	air	5.6	wrist watch	5.6	electric razor	2.8
	speaker	88.9	gamepad	2.8	PET bottled drinks	2.8	electric kettle	1.4
			electric razor	1.4	onion	1.4	rolling pin	1.4
potato	81.9	air	15.3	apple	1.4	electric razor	1.4	
toothbrush	80.6	air	16.7	scissors	2.8			
toothpaste	66.7	apple	33.3					
wrist watch	59.7	scissors	30.6	air	6.9	electric razor	2.8	

#### 4.1.2 位置推定

本節では、アクリル製平板に乗っ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識別できるかを検証する。識別に用いる物体としては、4.1.1節の物体識別に用いた物体のうち、もっとも構造がシンプルである building block を採用した。識別は、4.1.1節同様、SVMで行う。アクリル製平板に乗っている物体の位置をラベルとし、その時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。

まず、アクリル製平板に何も乗っていない状態で、図 2

の A 地点に取り付けたピエゾ素子からスイープ信号を送信し、図 2 の C 地点に取り付けたマイクでその応答を取得する。図 2 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 箇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。アクリル製平板に何も乗っていない状態と 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

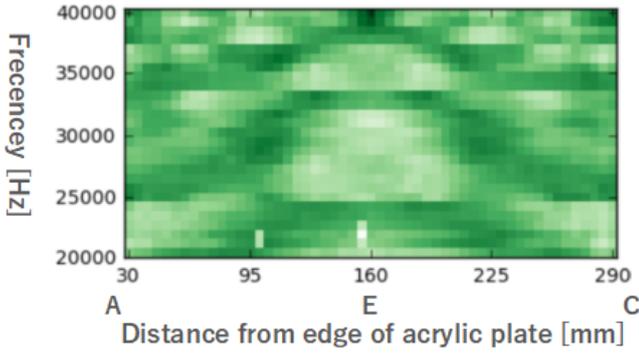


図 4 周波数特性と building block の位置の関係

と状態の組をランダムに 6 つのグループに分割する。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。

#### 結果と考察

交差検証の結果、識別率は 85.5 % であった。何も乗っていない状態の識別率は 100% であり、識別した位置の平均絶対誤差は 2.35mm (SD=4.73mm) であった。本実験を通して概ね位置推定が可能であると言える。誤識別のほとんどは 5mm 離れた地点への識別であり、誤識別の 81.8% がこの種類の誤識別であった。また 85mm 地点を 235mm 地点に、215mm 地点を 110mm 地点に誤識別するなど、位置を大きく誤ることもあり、誤識別の 14.8% がこの種類の誤識別であった。

位置の取得精度を低下させた要因として、位置を大きく誤る誤識別がいくつか見られたことが考えられる。この誤識別の原因として周波数特性の対称性が挙げられる。図 4 に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 2 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、色の濃さが振幅スペクトルの大きさを表す。多くの周波数で building block の位置がアクリル製平板の中心である 160mm の位置を軸にその振幅スペクトルが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。実際 15mm 以上の誤識別はすべて中心に関して対称の位置への誤識別であった。この対策としては用いる板の形状を非対称にすることやマイクの位置を図 2 の B 地点や D 地点などに変更す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発生する。そのため今後は、他の物体に対しても同様の実験を実施することで配置物体の位置による装置の周波数特性の変化を詳細に観察するとともに、振動工学の知見より周波数特性か

ら位置を算出できるような定量的なモデルを検討しなければならないと考えられる。

## 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

これまでに単一物体の種類及びその位置を取得できることが確認できた。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。特定の特徴量を用いて複数の物体が乗っている状態を学習させようとする場合、物体の組み合わ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多いと、物体を乗せる組み合わ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同時に乗っている状態の周波数特性を、何も乗っていない状態で得られる周波数特性と個々の物体が 1 つだけ乗っている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: 何も乗っていない状態。

状態 1: 物体 1 が乗っている状態。

状態 2: 物体 2 が乗っている状態。

状態 3: 物体 1 と物体 2 が乗っている状態。

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 [20]。

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

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加す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

$$\begin{aligned} & M(U_2 + U_1 - U_0)'' \\ & + C(U_2 + U_1 - U_0)' \\ & + K(U_2 + U_1 - U_0) = I + W_1 + W_2 \end{aligned} \quad (6)$$

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 が乗った状態のスweep信号による変位は、物体 1 が乗った状態のスweep信号による変位に、物体 2 が乗った状態のスweep信号による変位を加えた値から、何も乗っていない状態のスweep信号による変位を減らすことで求められる。すなわち式 (7) を用いることで、複数物体が乗っている状態のデータを個々の物体が乗っている状態のデータから推定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 (7) より算出される複数の物体が乗っている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乗っている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 節同様、SVM で行う。識別に用いる位置は図 2 で左側に示されるアクリル製平板の端から 30mm, 160mm, 290mm の 3 種類とし、識別に用いる物体の組は 4.1.1 節で用いた物体のうち、学習機に乗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。

まず、アクリル製平板に何も乗っていない状態で、図 2 の A 地点に取り付けたピエゾ素子からスweep信号を送信し、図 2 の C 地点に取り付けたマイク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み合わせ 33 通り ( $= {}_3C_1 \times {}_3P_1 + {}_3C_2 \times {}_3P_2 + {}_3C_3 \times {}_3P_3$ ) 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、次の 4 種類の解析を行う。解析 1 として、物体が 1 つだけ乗っている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 節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、何も乗っていない状態のデータと物体が 1 つだけ乗っている状態のデータを用いて物体が 2 個乗っている状態 18 通りの推定値を式 (7) 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 個乗っている状態 6 通りに対して行う。解析 4 として、解析 2 及び解析 3 で求めた推定値、何も乗っていない状態、物体が 1 つだけ乗っている状態のデータを用いて識別器を作成し、物体が 2 個及び 3 個乗っている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及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 2 及び表 3 に示す。表の見方は表 1 と同様である。ラベルは左から順に位置と対応しており、例えば、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 86.1%、解析 3 の場合 97.2%、解析 4 の場合 75.7% 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 99.2%、解析 2 の場合 95.8%、解析 3 の場合 91.7%、解析 4 の場合 93.8% 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 75.7%、冷蔵庫パターンで 93.8% と用いた物体による差が大きい。学習機パターンでは notebook と同時に pen や mug が置かれた際に識別を誤ることが多く、特に mug notebook pen はその 41.7% が mug notebook ---- に誤識別されている。この原因として、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 節においても notebook は識別しやすいクラスとして挙がっている。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。また、アクリル製平板の中心に関して対称な位置への誤識別もいくつか発生しており、最大で pen notebook ---- の 66.7% を ---- notebook pen に誤識別している。これは学習機パターンの解析 2 でも同様の誤識別が発生しており、前述の notebook の測定誤差の影響の大きさに加え、4.1.2 節で見られた配置物体の位置に対する周波数特性の対象性が原因と考えられる。一方で冷蔵庫パターンでは 4 つすべての解析において 90% 以上の識別率であり、全体的に識別率がよかったが、解析 3, 4 において apple onion orange の 33.3% を onion apple orange に誤識別するなど、apple と onion を相互に誤識別するケースがいくつか見られた。この傾向は 4.1.1 節の識別結果にも表れており、apple と onion が与える装置の周波数特性への影響に類似性が見られることが示唆される。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や apple と onion のような、装置の周波数特性への影響が似ている 2 物体の相互誤識別の問題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を示された。

## 5. 応用と今後の課題

本システムの具体的な応用先として冷蔵庫の棚板のセンサ化を現在検討している。棚板に乗っている食材とその位置を認識することができれば、温度などその食材に合った

表 2 学習機パターンにおける各解析の誤識別一覧

	True Label	[%]	False Label	[%]		
解析 1 識別率=98.1%	mug ----	91.7	---- mug	8.3		
	notebook ----	91.7	---- notebook	8.3		
解析 2 識別率=86.1%	mug notebook ----	75.0	pen notebook ----	8.3	---- notebook pen	8.3
	mug ---- notebook	83.3	notebook ---- pen	8.3	pen ---- notebook	8.3
	notebook ---- mug	66.7	notebook pen ----	16.7	notebook mug ----	16.7
	---- notebook mug	66.7	---- notebook pen	25.0	pen notebook ----	8.3
	notebook pen ----	91.7	notebook ---- pen	8.3		
	notebook ---- pen	75.0	notebook pen ----	25.0		
	---- notebook pen	91.7	pen notebook ----	8.3		
	pen notebook ----	33.3	---- notebook pen	66.7		
	pen ---- mug	91.7	---- mug pen	8.3		
	---- pen mug	83.3	notebook mug ----	8.3	pen ---- mug	8.3
	---- mug pen	91.7	pen mug ----	8.3		
解析 3 識別率=97.2%	mug notebook pen	91.7	pen notebook mug	8.3		
	notebook mug pen	91.7	notebook pen mug	8.3		
解析 4 識別率=75.7%	mug notebook ----	66.7	mug notebook pen	16.7	---- notebook pen	8.3
	mug ---- notebook	50.0	mug pen notebook	33.3	notebook ---- pen	8.3
	---- mug notebook	75.0	pen mug notebook	25.0		
	notebook ---- mug	41.7	notebook pen mug	33.3	notebook pen ----	16.7
	---- notebook mug	33.3	pen notebook mug	41.7	---- notebook pen	16.7
	notebook pen ----	91.7	notebook ---- pen	8.3	notebook mug ----	8.3
	notebook ---- pen	75.0	notebook ----	16.7	notebook pen ----	8.3
	---- notebook pen	91.7	pen notebook ----	8.3		
	pen notebook ----	25.0	---- notebook pen	66.7	---- notebook ----	8.3
	pen ---- notebook	91.7	---- notebook	8.3		
	pen mug ----	91.7	---- mug ----	8.3		
	pen ---- mug	91.7	---- mug pen	8.3		
	---- pen mug	75.0	---- pen ----	8.3	---- mug ----	8.3
	---- mug pen	91.7	pen mug ----	8.3	pen ---- mug	8.3
	mug notebook pen	50.0	mug notebook ----	41.7	---- notebook pen	8.3
	mug pen notebook	83.3	pen ---- notebook	16.7		
	notebook pen mug	66.7	notebook ---- mug	33.3		
	notebook mug pen	58.3	notebook mug ----	33.3	notebook pen mug	8.3
	pen mug notebook	75.0	---- mug notebook	16.7	---- pen notebook	8.3
pen notebook mug	91.7	pen notebook ----	8.3			

保存環境を提供することができる。また、加えてその重量を認識することができれば、食品の残量がわかるとともに食品の消費量がわかるので、健康管理なども実現できる。物体認識はカメラを用いるものが主流であるが、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をあり、カメラによって認識するには不向きである。

今後はまず 4.1.2 節で述べた周波数特性の対称性の問題を解決するため、マイクの位置の検討を行う必要がある。その後、アクリル製平板の周波数特性に影響を与えられ、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。

## 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、ピエゾ素子、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 98.2%、85.5%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及び冷蔵庫内を想定した物体の組を配置する組み合わせに関して、それぞれ 75.7%及び 93.8%の精度で識別できた。本実験を通して単

表 3 冷蔵庫パターンにおける各解析の誤識別一覧

	True Label	[%]	False Label	[%]		
解析 1 識別率=99.2%	---- onion	91.7	---- apple	8.3		
解析 2 識別率=95.8%	---- onion orange	91.7	---- apple orange	8.3		
	---- orange onion	91.7	---- orange apple	8.3		
	orange apple ----	83.3	orange onion ----	8.3	apple ---- orange	8.3
	orange ---- apple	91.7	apple ---- orange	8.3		
	apple orange ----	91.7	onion orange ----	8.3		
	apple ---- orange	75.0	onion ---- orange	25.0		
解析 3 識別率=91.7%	apple onion orange	66.6	onion apple orange	33.3		
	orange onion apple	83.3	orange apple onion	8.3	onion apple orange	8.3
解析 4 識別率=93.8%	---- onion orange	83.3	---- apple orange	8.3	apple onion orange	8.3
	---- orange onion	91.7	---- orange apple	8.3		
	orange apple ----	83.3	orange onion apple	8.3	apple onion orange	8.3
	orange ---- apple	75.0	orange onion apple	8.3	apple ---- orange	8.3
	apple ---- orange	75.0	onion ---- orange	25.0		orange ---- onion 8.3
	apple onion orange	66.7	onion apple orange	33.3		
	orange onion apple	75.0	orange apple onion	16.7	onion apple orange	8.3

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, Proceedings of the 29th Annual Symposium on User Interface Software and Technology, UIST '16, New York, NY, USA, ACM, pp. 833–841 (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, Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, New York, NY, USA, ACM, pp. 2161–2170 (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, New York, NY, USA, ACM, pp. 299–299 (online), DOI: 10.1145/1936652.1936727 (2010).

[4] 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, New York, NY, USA, ACM, pp. 724–725 (online), DOI:

10.1145/506443.506566 (2002).

[5] Murray-Smith, R., Williamson, J., Hughes, S. and Quaade, T.: Stane: Synthesized Surfaces for Tactile Input, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08, New York, NY, USA, ACM, pp. 1299–1302 (online), DOI: 10.1145/1357054.1357257 (2008).

[6] Harrison, C. and Hudson, S. E.: Scratch Input: Creating Large, Inexpensive, Unpowered and Mobile Finger Input Surfaces, Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology, UIST '08, New York, NY, USA, ACM, pp. 205–208 (online), DOI: 10.1145/1449715.1449747 (2008).

[7] Harrison, C., Xiao, R. and Hudson, S.: Acoustic Barcodes: Passive, Durable and Inexpensive Notched Identification Tags, Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology, UIST '12, New York, NY, USA, ACM, pp. 563–568 (online), DOI: 10.1145/2380116.2380187 (2012).

[8] 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, New York, NY, USA, ACM, pp. 732–733 (online), DOI: 10.1145/506443.506570 (2002).

[9] Harrison, C., Tan, D. and Morris, D.: Skininput: Appropriating the Skin As an Interactive Canvas, Commun. ACM, Vol. 54, No. 8, pp. 111–118 (online), DOI: 10.1145/1978542.1978564 (2011).

[10] Harrison, C., Schwarz, J. and Hudson, S. E.: TapSense: Enhancing Finger Interaction on Touch Surfaces, Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, UIST '11, New York, NY, USA, ACM, pp. 627–636 (online), DOI: 10.1145/2047196.2047279 (2011).

[11] Lopes, P., Jota, R. and Jorge, J. A.: Augmenting Touch Interaction Through Acoustic Sensing, Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, ITS '11, New York, NY, USA, ACM, pp. 53–56 (online), DOI: 10.1145/2076354.2076364 (2011).

- [12] 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, New York, NY, USA, ACM, pp. 1847–1852 (online), DOI: 10.1145/2212776.2223717 (2012).
- [13] Wylie, C. and Mackie, R. R.: Active Sonar Target Detection and Reporting: Perceived Consequences and their Effects on Performance, Technical report, *DTIC Document* (1972).
- [14] 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).
- [15] Gupta, S., Morris, D., Patel, S. and Tan, D.: SoundWave: Using the Doppler Effect to Sense Gestures, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, New York, NY, USA, ACM, pp. 1911–1914 (online), DOI: 10.1145/2207676.2208331 (2012).
- [16] Takemura, K., Ito, A., Takamatsu, J. and Ogasawara, T.: Active Bone-conducted Sound Sensing for Wearable Interfaces, Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology, UIST '11 Adjunct, New York, NY, USA, ACM, pp. 53–54 (online), DOI: 10.1145/2046396.2046419 (2011).
- [17] Collins, T.: Active acoustic touch interface, Electronics letters, Vol. 45, No. 20, pp. 1055–1056 (2009).
- [18] 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, New York, NY, USA, ACM, pp. 3:1–3:1 (online), DOI: 10.1145/2782782.2792490 (2015).
- [19] Ono, M., Shizuki, B. and Tanaka, J.: Touch & Activate: Adding Interactivity to Existing Objects Using Active Acoustic Sensing, Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, UIST '13, New York, NY, USA, ACM, pp. 31–40 (online), DOI: 10.1145/2501988.2501989 (2013).
- [20] 塩尻弘雄, 中村秀治 : 構造解析における動的応答解析の一方方法について, 土木学会論文報告集, Vol. 1976, No. 246, pp. 21–33 (1976).