

音の反響を用いたスマートフォン置き場の材質推定

長谷川 達人¹ 平橋 智史² 越野 亮³

概要: 本研究では、スマートフォン置き場の材質を推定するシステムを提案する。スマートフォン置き場の材質がわかることで、材質に応じて適切な音量設定やバイブレーション設定などを自動で変更することが可能になる。これまで、歩行時におけるスマートフォンの所持位置推定の研究は行われていたが、加速度センサ値の変動の少ない停止時においては推定が難しいとされてきていた。本研究では、スマートフォンから音を発したとき、スマートフォンが置かれている場所によって反響音が変わるという仮定のもと、音の反響を用いたスマートフォン置き場の材質推定手法を開発した。評価実験の結果、6種類に分類した材質を89.2%の精度で識別に成功し、環境や置き方による影響が大きくないことも明らかにした。

Material estimation of smartphone place using sound echo

Tatsuhito Hasegawa¹ Satoshi Hirahashi² Makoto Koshino³

1. はじめに

コンテキストウェアネスはスマートフォンにおいて重要な技術である。スマートフォンが利用者の状況や周囲の環境を認識できることで、より便利なサービスが実現できる。例えば、行動認識の研究 [1,2] は盛んに行われており、観測されるセンサデータを分析し、歩行、走行、歯磨きや皿洗いなど、様々な利用者の行動を認識している。利用者の状況が認識できることで、ライフログとしての記録、エクササイズ支援など多くのサービスへの応用が実現できる。

本研究では、コンテキストウェアネスの一部として、スマートフォン置き場の材質を推定するシステムを提案する。Lopez-Tovar らの調査 [3] 内で示されるように、スマートフォンがポケットにあるときにはバイブレーションで通知してほしいと思う人が多いが、机の上にあるときにはLED点灯で通知してほしいと思う人が多くなる。このように、利用者の予定やスマートフォンの所持位置によっ

て通知方法を適切に変更することが望まれている。また、机のようなかたい材質の上にスマートフォンが置かれた場合と、布団のようなやわらかい材質の上に置かれた場合では、バイブレーションの振動による反動音や通知の間聞こえ方が変わるため、スマートフォンが置かれている物体の材質によっても、通知設定が変えられることが望ましい。

スマートフォンが置かれている状況では加速度がほぼ変動しないため、加速度センサを用いたとしても、主に端末の傾きしか特徴量として利用することができない。更に、スマートフォンの置かれ方は材質によらず様々である可能性があることから、加速度センサを用いた識別は本研究には適さないと考えられる。

そこで本研究では、スマートフォンから発した音の反響を用いて、どのような材質の上に置かれているのかを推定する手法を開発する。推定対象は、スマートフォンが停止時に置かれることが多いと考えられる、鞆の中、ポケットの中、机の上、布団の上、スマートフォンスタンドの上を中心に、様々な材質にて識別可能性を調査する。

2. 関連研究

2.1 歩行時におけるスマートフォンの所持位置推定

スマートフォンの所持位置を推定する研究は行われているが、歩行時に限定して加速度センサを用いた所持位置推定を行っているものが多い [4-8]。

¹ 東京医療保健大学医療保健学部
Division of Healthcare Informatics,
Faculty of Healthcare, Tokyo Healthcare University

² 金沢大学大学院自然科学研究科
Graduate School of Natural Science & Technology,
Kanazawa University

³ 石川工業高等専門学校電子情報工学科
Department of Electronics and Information Engineering,
Ishikawa National College of Technology

これらの研究では、歩行のような一定の動作を行っている際、スマートフォンの所持位置に応じて加速度などのセンサ値に特徴が現れるということから、前ポケットや後ポケット、手に持っているなどのスマートフォンの所持位置を推定している。これらは加速度などのセンサ値を一定の間隔で特徴量に変換し、機械学習にて所持位置の推定を行っている。しかし、加速度などのセンサを用いた場合、スマートフォンが停止している際にどのような場所で停止しているのかを推定することは難しいとされてきた。

2.2 停止時におけるスマートフォン置き場の推定

独自センサを用いた研究として、Harrison らの研究 [9] では、独自センサを対象物に接地させることでスマートフォン置き場を識別している。センサは赤外線 LED、紫外線 LED、RGB LED などを併用して開発されている。リュックサックやリュックサックのポケット、プラスチック、ブランケット、カーペットなど日常で置かれうる 27 材質の識別を行っている。

スマートフォン標準機能を用いた研究として、バイブレーションを用いた研究がある。Cho らの研究 [10] では、バイブレーションを動作させ、その反動を加速度センサで計測しスマートフォン置き場を識別している。ソファや木の机、手、カバン、ズボンのポケットを 85% 程度の精度で識別している。また、Hwang らの研究 [11] では、バイブレーションを動作させ、その反響音を計測してスマートフォン置き場を識別している。ベッドや机、ポケット、携帯ホルダーを 92% 程度の精度で識別している。

同様に、スマートフォンの標準機能を用いた研究として、音を用いた研究がある。Diaconita らの一連の研究 [12-14] では、スマートフォンから音を発して、その反響音を計測しスマートフォン置き場を識別している。識別対象は共通して鞆、机、手、ポケットとしている。手法は、ホワイトノイズの反響音を用いる手法 [12]、メールなどの通知音の反響音を用いる手法 [13]、通知音の手法とバイブレーションの反動を併用する手法 [14] を提案している。

2.3 本研究の位置づけ

関連研究でも示されている通り、スマートフォンが置かれている場所の材質を識別することは、コンテキストウェアネス研究の分野で以前から望まれている。本研究のように音の反響を用いることで、外部センサを用いずともスマートフォン置き場の識別が実施可能となる。また、バイブレーションを用いた手法は置き場の識別が高精度に実現できるものの、日常生活で利用する場合バイブレーションを電話やメールの通知と誤認してしまう恐れがある。それに対して音を発する場合、ホワイトノイズやビーブ音のように利用者が通知と誤認しづらい音を用いることでそれを回避することができる。

Diaconita らの一連の研究 [12-14] との大きな違いは、ビーブ音を用いる点と、識別する材質の種類である。着信時の通知音の反響を用いた研究 [13,14] は、高精度でスマートフォン置き場が推定でき、かつ実際の通知を利用することで不要な音を発する必要がなくなるという利点があるが、着信が起こるタイミングでしか置き場が推定できなくなるという不便さもある。本研究に最も近い研究 [12] では、ホワイトノイズを用いて周波数成分ごとの減衰から MFCC などの特徴を抽出して置き場の推定を行っている。

本研究では、単一周波数音の反響による高調波成分の変化に着目して特長を抽出する手法を提案する。単一周波数を用いるため、スペクトルの外形を示す MFCC よりも、反響による高調波成分を直接特徴として扱うことで情報を劣化させることなく特徴表現ができると考えられる。既存とは異なる手法を実現することで、今まで識別が困難であった対象物の識別が可能になったり、既存手法との組み合わせによって全体の推定精度が向上したりする可能性がある。

既存研究では、ポケット、鞆、机、手の 4 種類の識別しか行われていないため、識別しやすい材質の考察が難しい。したがって、本研究の位置づけは、ビーブ音を用いたスマートフォン置き場推定手法を新たに開発し、置き場の材質をどの程度の種類まで識別できるのかという点を考察する。音の反響を用いることから、周囲の環境や、音の大きさ、ビーブ音の周波数など様々な要因が精度に影響すると考えられるが、本稿では第一歩として、どの程度の種類の材質を識別できるのかという点に着目する。

3. 提案手法

3.1 識別原理

本研究では正弦波によるビーブ音の反響を用いてスマートフォン置き場を識別する手法を提案する。秋田らの基礎研究 [15] にて検証されているように、超音波を用いる非接触計測において反射の波形がもつ情報を解析することで、対象物の材質など高次の情報が識別可能であると考えられている。したがって、図 1 のように、スマートフォンから対象物にビーブ音を当て、同時に周囲の音を記録し、その音の波形を解析することで、対象物の特性を識別することができる可能性がある。

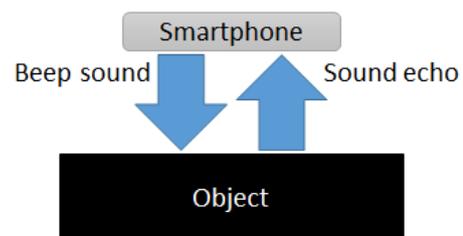


図 1 音の反響を用いたスマートフォン置き場の材質推定

Fig. 1 Material estimation of smartphone place using sound echo.

本研究においては、図1のスマートフォンは対象物の上に置く形となるため、両者の距離はほぼなくなる。観測する反響音は、対象物にかぎらず壁や周囲の物体からの反響を含んでいるが、最も近くにある対象物の影響を大きく受けていると考えられる。秋田らの研究では40000Hzの超音波を用いているが、スマートフォンの取音特性上サンプリング周波数44100Hzが上限となるため、本研究では超音波でなく可聴域のビーブ音を採用する。

提案する処理手順は図2のとおりである。はじめに、対象物の上に置かれたスマートフォンは特定の周波数の音を鳴らし、同時に周囲の音を記録する。記録された音の波形を一定のウィンドウに区切って抽出し、フーリエ変換を実施する。変換後のデータから対象物の特性を示すであろう特徴量を抽出し、機械学習によって対象物を識別する。

本手法の利点は、一定のウィンドウ分だけのデータがあれば識別が可能であるため、音を鳴らす時間がごくわずかでかまわないという点である。例えば、サンプリング周波数44100Hzで4096サンプルを抽出する場合、理論上は約0.01秒音を鳴らすだけでよいということになる。観測時間が短いことで、利用者に不要な音を感じさせづらくするだけでなく、置き場推定のタイムラグが小さくなる。

3.2 特徴量の抽出

提案手法ではFFT処理を行った観測音から特徴量を抽出する。一般的に超音波を用いた対象物との関係推定には、送信波と観測波の時間差を計測することが多いが、提案手法では振幅や周波数が一定なビーブ音を用い、周波数特性の変動を計測する。スマートフォンで観測する音は反響音を含むため、周囲に置かれている物質が観測波の周波数特性に影響を与えると考えられる。

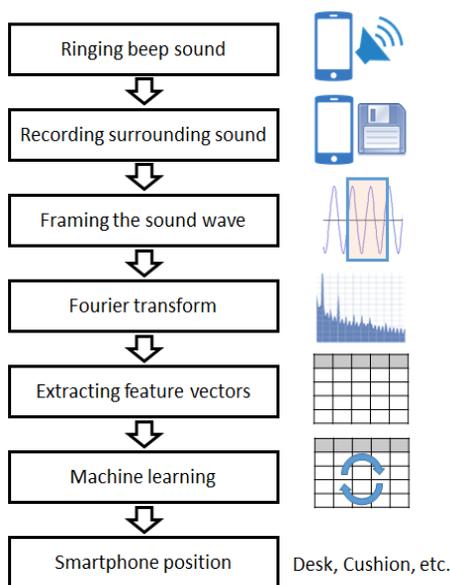


図2 提案手法の処理概要

Fig. 2 Process outline of our proposed method.

図3は鋼の板、木の板、前ポケット、クッションにスマートフォンを置いた際の観測音のFFT結果である。基音(発信音)は1024Hzであるが、観測音には1024Hz以外にも、その整数倍の周波数にいくつかのピークが見られる。これらは高調波成分(あるいは倍音)と呼ばれる成分で、これらの振幅比で音色が変わると言われている。図を見てわかるように、対象物の種類によって高調波成分のピーク大きさが変わることが見て取れる。したがって、高調波成分のピークを表現する特徴量を抽出することで、対象物を識別できると考えられる。

周波数スペクトルから抽出する特徴は色々提案されているが、本研究で調査した結果、中でも高調波成分の比率を特徴とするいくつかの指標を採用する。Martinらの研究[16]では、楽器の識別課題に対する倍音振幅比関連の特徴として、Spectral Centroid (SC)や奇数倍音*1と偶数倍音*2の振幅比の特徴を用いている。SCはスペクトルの形状を示す特徴であり、周波数成分の重み付き平均として式(1)で計算される。ここで C_t 、 $X_t[n]$ はそれぞれ時刻 t におけるSC、 n 番目の周波数成分のパワーである。

$$C_t = \frac{\sum_{n=1}^N nX_t[n]}{\sum_{n=1}^N X_t[n]} \quad (1)$$

更に、秋田らの研究[15]にて、有用性が検討されている指標である、Total Harmonic Distortion (THD)、THD+N、THD_N、Primary Signal to Noise Ratio (S₁NR)を特徴として採用する。THDは基本周波数と高調波成分の比であり、式(2)で計算される。THD+NはTHDに似た指標で、基本周波数とノイズを含む全高調波の比であり、式(3)で計算される。THD_NはTHDとTHD+Nの差であり、式(4)で計算される。S₁NRは全体のパワーに対する基本周波数のパワーであり、式(5)で計算される。これらは基本

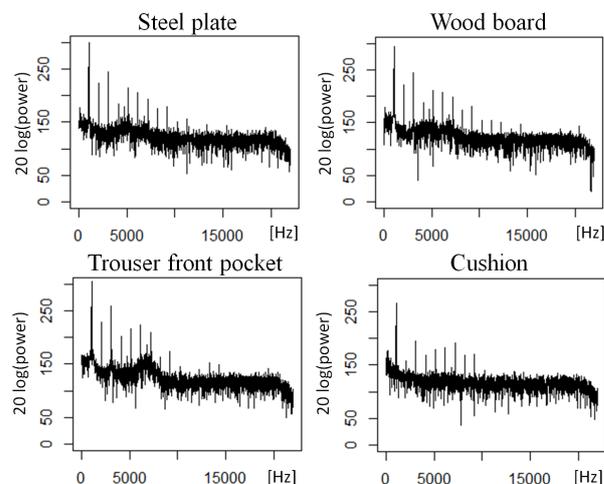


図3 観測音のFFT結果

Fig. 3 FFT result of recorded sounds.

*1 基音の奇数倍となる倍音

*2 基音の偶数倍となる倍音

周波数に対して観測波がどれほど歪んでいるかを示す指標であり、スマートフォン置き場の材質や表面の形状によって歪み方が変わると考えられる。

$$\text{THD} = \frac{\sqrt{\sum_{i=2}^n V_i^2}}{V_1} \quad (2)$$

$$\text{THD} + \text{N} = \frac{\sqrt{V_{total}^2 - V_1^2}}{V_1} \quad (3)$$

$$\text{THD}_N = (\text{THD}) - (\text{THD} + \text{N}) \quad (4)$$

$$S_1\text{NR} = \frac{V_1^2}{V_{total}^2} \quad (5)$$

ここで、 V_1 は基本周波数成分のパワーであり、 V_i はその i 倍の周波数成分のパワーである。 V_{total} は基本周波数を含むパワースペクトル全体の和である。

他にも、パワースペクトルの二乗平均平方根、元波形のゼロ交差率、高調波成分間の比、高調波成分のパワーなど、表 1 の 30 次元を特徴として採用している。

表 1 採用した特長量の一覧

Table 1 List of features which our proposed method used.

Index	Name	Note
1	SC	Spectral Centroid
2	OER	Odd / even ratio
3	THD	Total Harmonic Distortion
4	THD+N	Total Harmonic Distortion + Noise
5	THD _N	Difference between the THD and THD+N
6	S ₁ NR	Primary Signal to Noise Ratio
7	RMS	Root Mean Square
8	ZCR	Zero Crossing Rate
9	PR12	Power Rate (V_2/V_1)
10	PR13	Power Rate (V_3/V_1)
11	PR14	Power Rate (V_4/V_1)
12	PR15	Power Rate (V_5/V_1)
13	PR16	Power Rate (V_6/V_1)
14	PR17	Power Rate (V_7/V_1)
15	PR18	Power Rate (V_8/V_1)
16	PR19	Power Rate (V_9/V_1)
17	PR23	Power Rate (V_3/V_2)
18	PR24	Power Rate (V_4/V_2)
19	PR34	Power Rate (V_4/V_3)
20	PR35	Power Rate (V_5/V_3)
21	P1	Power (V_1)
22	P2	Power (V_2)
23	P3	Power (V_3)
24	P4	Power (V_4)
25	P5	Power (V_5)
26	P6	Power (V_6)
27	P7	Power (V_7)
28	P8	Power (V_8)
29	P9	Power (V_9)
30	P10	Power (V_{10})

4. 評価実験と考察

4.1 実験環境とデータセット

提案手法の有効性を示すため、様々な物体の上にスマートフォンを置き、ピープ音を発すると同時に周囲の音を記録する実験を行い、提案手法の推定精度評価を行った。今回、実験を行った対象物は、スマートフォンが置かれそうな場所を対象に、図 4 の 18 種類を用いた。実験手順は、まずピープ音の発信と周囲の音の記録を開始し、対象物の上もしくは中にスマートフォンを置く、5 秒間の音を記録した後、次の実験を始める。観測した 5 秒間のデータから 1~4 秒地点をトリミングし、50 分割した後それぞれ 4096 サンプル切り出して 1 つのインスタンスを生成する。すなわち、一回の観測実験にて、1 データセット (50 インスタンス) を生成する。



図 4 対象としたスマートフォン置き場 18 種類 (a) 胸ポケット, (b) ジャケットポケット, (c) 前ポケット, (d) 後ポケット, (e) 鞆, (f) クッション, (g) 車のマット, (h) ブランケット, (i) タオル, (j) ゴムマット, (k) 木の机, (l) 木の板, (m) 鋼の板, (n) 石のタイル, (o) スマホスタンド A, (p) スマホスタンド B, (q) スマホスタンド C, (r) 手

Fig. 4 18 smartphone places which our system classifies. (a)chest pocket, (b)jacket pocket, (c)trouser front pocket, (d)trouser back pocket, (e)bag, (f)cushion, (g)car mat, (h)blanket, (i)towel, (j)gom mat, (k)wood desk, (l)wood board, (m)steel plate, (n)stone tile, (o)smartphone stand A, (p)smartphone stand B, (q)smartphone stand C, and (r)hand

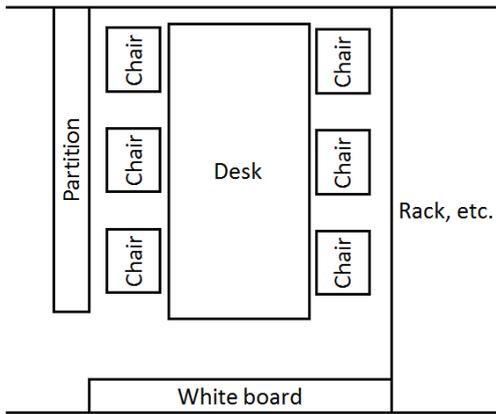


図 5 実験を実施した部屋の間取り

Fig. 5 Layout of the room where our experiments performed.

実験に使用したスマートフォンは Samsung 社製の GALAXY S2 である。GALAXY S2 は取扱説明書^{*3}P46 によると、端末背面下部に音を発するスピーカがあり、端末の上下に音を記録するマイクがある。スマートフォンを置く向きによっては図 1 のように対象物にスピーカが直接接しない向きになることがある。したがって、本研究ではスマートフォン（のスクリーン）を上向きに置く場合と、下向きに置く場合の 2 種類の状態において実験を行い、その差を考察する。

実験を行った部屋は図 5 の間取りであり、机の上に対象物を置いて、その上にスマートフォンを置いて記録を行った。図中左に壁がないのは、左側には部屋が続いていることを示す。実験は同部屋に実験者以外が在室しない状態で実施され、不要な物音は立てないようにしているが、廊下での足音や会話音、外で車が走る音などの雑音が聞こえる環境である。また、反響音は周囲に置かれているものの影響も受けると予測されることから、図 6 の 3 種類の環境を構築し、記録を行った。図中の仮想壁 (Temporary wall) は、高さ約 0.5m、奥行き約 1m、幅約 0.5m のダンボールを 2 つ積んで壁に見立てたものである。環境 1 は周囲に何も無い状況を、環境 2 は物に囲まれている状況を、環境 3 はデスク上で起こりうる環境を再現している。ただし、ポケット、鞆に関しては一部例外があり、環境 C ではポケットに格納する実験者は椅子に座って「人」の位置におり、鞆は足元に置いている。

以上より、実験に用いるデータは一回の観測実験 5 秒間から生成されるデータセット (50 インスタンス) を基準として、18 種類の対象物に対して、スマートフォンを上下向き 2 種類置き、かつ、図 6 の 3 環境で記録を行っている。すなわち、各対象物に対して、6 データセット (300 インスタンス) を生成している。

^{*3} NTT docomo: GALAXY S II SC-02C 取扱説明書
https://www.nttdocomo.co.jp/binary/pdf/support/trouble/manual/download/SC-02C_J_OP_01.pdf
 (2016/5/8 Accessed)

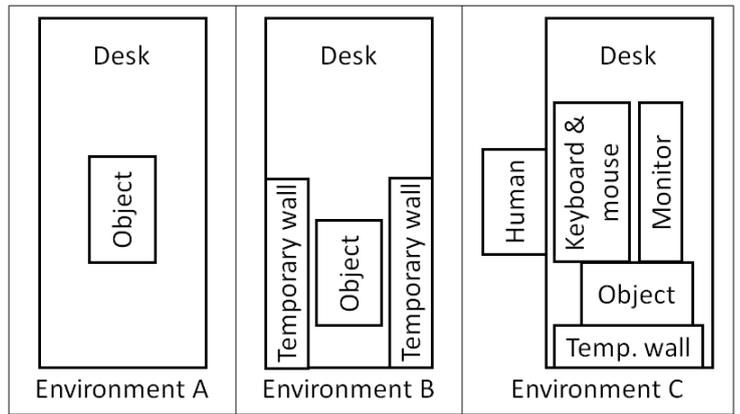


図 6 実験環境

Fig. 6 Experimental environments.

今回実験では、ピープ音に 1024Hz の正弦波を用いた。音量は GALAXY S2 のレベル 7 (最大音量の半分程度) で統一している。今回用いた機器は、サンプリング周波数 44100Hz が上限であり、高調波成分を特徴として用いる上で最適な周波数を選択した結果、1024Hz を採用した。音量は小さいほうが実環境で利用した際に利用者がわずらわしく感じないと考えられるが、今回の実験では周囲の雑音の影響を減らすため、相対的に観測音の音量を高めた。

4.2 評価方法

生成したデータセットに対して手動でラベル付けを行い、機械学習でスマートフォン置き場を推定する。機械学習アルゴリズムには今回 Random Forest [17] を採用した。パラメータは標準パラメータ (特徴量次元数の平方根) のままとしている。

ただし、本実験にて一般的に採用される 10-fold cross-validation を用いた精度評価を行うと、ほぼ 100% の予測精度を記録する。これは 1 つのデータセットから 50 インスタンスを生成しているため、10 分割した際に同じデータセット内のデータを学習データに含むことから発生する。しかし、実環境ではまったく同じ環境でのデータが学習データに含まれていることは珍しい。

そこで本研究では、1 つのデータセット (50 インスタンス) をテストデータに、残りを学習データとする評価手法として、Leave-one-dataset-out cross-validation (LODO-CV) にて予測精度の評価を行った。

4.3 スマートフォン置き場 18 種類の識別結果

LODO-CV にて 18 種類のスマートフォン置き場を識別した結果が表 2 である。全体の Accuracy は 44.6% と高くないものの、いくつか特徴的な結果が見受けられる。

誤りパターンに着目すると、グレーで囲んだ範囲でグループ化してみることができる。まず、(a)~(e) はポケットと鞆に格納するケースであり、相互に予測を誤るケース

表 2 スマートフォン置き場 18 種類の識別結果

Table 2 Classification result with 18 kinds of smartphone places.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)	(o)	(p)	(q)	(r)	Precision
(a)chest pocket	161	31		21	28														66.8%
(b)jacket pocket	47	68	5	79														65	25.8%
(c)trouser front pocket	11	18	186	44	50													7	58.9%
(d)trouser back pocket	30	103		56														24	26.3%
(e)bag	1		64	76	117		3			7									43.7%
(f)cushion						297		26	25									1	85.1%
(g)car mat		3	43	18	47		203	57			1	37					45	21	42.7%
(h)blanket				6	6	1	48	194	27	2								41	59.7%
(i)towel	50	30						22	225									50	52.7%
(j)gom mat		1			50	2				251	2	56							69.3%
(k)wood desk										3	1	104	66	87					0.4%
(l)wood board							38			36	102		7	84	8				0.0%
(m)steel plate				2		2					42	40	8	129					3.6%
(n)stone tile								1			152	63	219						0.0%
(o)smartphone stand A								1	2	1					292			76	78.5%
(p)smartphone stand B																	300	100	75.0%
(q)smartphone stand C		4					5		17									52	38.8%
(r)hand		42	2						4									12	0.0%
Recall	53.7%	22.7%	62.0%	18.7%	39.0%	99.0%	67.7%	64.7%	75.0%	83.7%	0.3%	0.0%	2.7%	0.0%	97.3%	100.0%	17.3%	0.0%	44.6%

が多く、これらはスマートフォンが何かに包まれている状態という共通点がある。(f)~(g)はクッション、車のマット、ブランケット、タオルであり、(f)クッションは他に比べて誤りが少ないものの、(g)-(h)、(h)-(i)間での誤りが多く、これらは対象物がやわらかい材質という共通点がある。(k)~(n)は木の机、木の板、鋼の板、石のタイルであり、相互に予測を誤るケースが多く、これらは対象物がかたい材質かつ表面に凹凸が少ないという共通点がある。ただし木の板に関しては木独特のざらざらした肌触りである。そして(p)~(q)はスマホスタンド B、C であり、相互に誤るケースが多く、これらは同じ形状のホールド部分を持つスマホスタンドであるという共通点がある。

ただし、(r)手に着目すると、正しく予測できているものが少なく、更に誤りパターンもまばらである。手持ちの場合は、持ち方によってスピーカ部に直接手が触れたり、触れなかったりしたことが原因と考えられる。従って、本提案手法は手持ちの識別精度が低いという結果になった。

単独で識別精度が良かったのは、(j)ゴムのマットと、(o)スマホスタンド A である。(j)は業務デスクに一般的に敷かれるマットであり、(o)はプラスチック製の立てかけるタイプのスマホスタンドである。どちらも他にグループ化できない材質であり、特徴的な感触の材質だったことから識別精度が高くなったと考えられる。

4.4 スマートフォン置き場 6 種類の識別結果

前節の結果より、提案手法は次の 6 グループの識別に有用と考えられる。(a)~(e)をまとめて (s) ポケット&鞆、(f)~(g)をまとめて (t) ソフト、(j)はそのまま (u) ゴムのマット、(k)~(n)をまとめて (v) ハード、(o)はそのまま (w) スタンド A、(p)~(q)をまとめて (x) スタンド B&C と再定義し、再度 LODO-CV にて識別精度の評価を行った。

表 3 スマートフォン置き場 6 種類の識別結果

Table 3 Classification result with 6 kinds of smartphone places.

	(s)	(t)	(u)	(v)	(w)	(x)	Precision
(s)pocket&bag	1330	15	24		3		96.9%
(t)soft	116	1134	6	33	1	134	79.6%
(u)gom	21	2	214	50		1	74.3%
(v)hard		44	56	1117	8		91.2%
(w)standA		2			288		99.3%
(x)standB&C	33	3				465	92.8%
Recall	88.7%	94.5%	71.3%	93.1%	96.0%	77.5%	89.2%

スマートフォン置き場を 6 種類にまとめて再評価した結果が表 3 である。ただし今回、手は除外している。全体の推定精度は 89.2%となり、提案手法は大まかに分類した 6 種類の材質を高い精度で識別可能であることが確認された。特に何かに (s) 包まれている状態、(t) やわらかい材質、(u) ゴムのマット、(v) かたい材質の上に置かれていることを識別できたという結果は、今後通知手法の自動変更に応用する際に役立つと考えられる。

4.5 環境の影響に関する考察

周囲の環境やスマートフォンの置き方による影響を考察する。表 4 は表 3 の結果に対して、上向きのデータのみ、下向きのデータのみ、環境 A~C のみの結果にわけてまとめたものであり、各置き場の列には F 値を、右端に Accuracy を記載している。

まず、スマートフォンを置く向きに着目する。一般的にはスマートフォンを上向きに置くであろうことを考慮すると、Accuracy が 90.9%と高精度である。下向き、すなわちスピーカが空中に向いている状態のほうが、上向きよりも Accuracy が 3%程度低下することが確認されたが、スピーカが対象物に向いていない状況でも Accuracy が 87.4%と

表 4 6 種類の識別結果の条件別比較

Table 4 Comparison of classification result with 6 kinds of smartphone places between each environment.

	F-measure						Accuracy
	(s)pocket&bag	(t)soft	(u)gom mat	(v)hard	(w)stand A	(x)stand B&C	
All	92.6%	86.4%	72.8%	92.1%	97.6%	84.5%	89.2%
Upward	96.0%	85.9%	85.4%	93.1%	95.5%	83.7%	90.9%
Downward	88.8%	86.9%	60.1%	91.1%	99.7%	85.1%	87.4%
Environment A	94.4%	92.2%	74.8%	91.9%	97.5%	100.0%	92.8%
Environment B	89.8%	78.7%	69.8%	90.3%	99.5%	59.5%	83.4%
Environment C	93.7%	88.9%	73.3%	94.1%	95.8%	90.4%	91.4%

なり、スピーカが表面にあるか背面にあるかという点は影響が大きいという予想外な結果が得られた。ただし、(u) ゴムのマットに対してだけは、下向きの場合 20%以上の大幅な F 値の低下が確認された。

次に、スマートフォン周辺の環境に着目する。環境 A～C では、環境 B のみ Accuracy が 10%程度低下している。環境 B は周囲にダンボールの仮想壁を作ったことから、近い距離で、周囲を物体で囲まれると推定精度が低下する可能性があることが確認された。特に (t) やわらかい材質と、(x) ホールドタイプのスマートフォンスタンド B と C は強く環境の影響を受け F 値が変動している。

以上を踏まえると、スマートフォンの置き方や、周囲のもの置かれ方によって、推定精度に 5～10%の変動は見られるものの、全体の推定精度は 80%以上を保っており、環境の影響は大きくないことが確認された。

4.6 特徴量の重要度

LODO-CV にて 6 種類の識別精度評価を行った際、Random Forest ではジニ係数による特徴量の重要度評価を行っている。表 5 に本研究で採用した特長量と重要度 (MeanDecreaseGini) を重要度降順で示す。Index 列は表 1 に対応している。

表 5 を見ると、基音のパワー (P1) をはじめとして、各高調波成分のパワーの重要度が高いことがわかる。また、THD など音の歪みを示す特徴よりも、高調波間のパワー比率などの直接的な特徴のほうが重要であり、提案手法では、各ピークのパワーを用いるだけのシンプルな実装ができる可能性が見受けられた。

5. おわりに

本研究では、スマートフォンがどのような材質の上に置かれているかというコンテキストを推定する手法を開発した。スマートフォン置き場の材質がわかることで、材質に応じて通知方法を適切に自動変更するシステムなどへの応用が可能となる。例えば、ポケットに格納されているときにはバイブレーションで、かたい机の上に置かれているときには、バイブレーションなしで LED 点灯のみで通知す

表 5 特長量の重要度一覧

Table 5 List of importance of features.

Order	Index	Name	MeanDecreaseGini
1	21	P1	437.7
2	22	P2	396.8
3	23	P3	365.2
4	25	P5	325.9
5	26	P6	320.6
6	10	PR13	249.2
7	12	PR15	244.1
8	24	P4	214.3
9	17	PR23	208.4
10	13	PR16	194.2
11	20	PR35	186.0
12	9	PR12	182.7
13	19	PR34	119.4
14	11	PR14	101.8
15	27	P7	80.8
16	14	PR17	78.6
17	18	PR24	76.6
18	7	RMS	43.6
19	28	P8	22.4
20	30	P10	18.1
21	3	THD	16.1
22	29	P9	10.7
23	6	S ₁ NR	9.6
24	4	THD+N	9.4
25	5	THD _N	9.3
26	1	SC	8.3
27	2	OER	6.6
28	15	PR18	6.1
29	16	PR19	4.5
30	8	ZCR	1.2

るなどのように、利用者が気づきやすい通知を実現することが可能になる。

提案手法では、スマートフォンから 1024Hz のピーブ音を発し、反響音を解析することでスマートフォン置き場の推定を実現した。18 種類の材質にて提案手法を評価した結果、いくつかのグループに分類できる可能性が見受けられ、最終的に、ポケット&鞆、やわらかい材質、ゴムのマッ

ト、かたい材質、スマートフォンスタンド（置き型）、スマートフォンスタンド（ホルド型）の6種類を89.2%の精度で識別した。

更に、スマートフォンを上向きや下向きに置いた場合、周囲に色々なものが置かれていた場合に関しても、影響は大きくないことも、実験を経て明らかにした。

今後の発展として3点を挙げる。1つ目は、周囲の雑音の影響を考察するため、周囲のノイズレベルを変動させたときの識別精度の変動を実験していくことである。また、雑音の音量に対して、提案手法が発する音量をどの程度まで下げても識別精度に影響がでないかを考察する。

2つ目は、発信音の周波数特性の最適化である。今回、単一周波数1024Hzの正弦波を発信音に採用したが、単一周波数でも更に最適な周波数がある可能性がある。また、今回採用しなかった高調波成分間の比率や、複数の周波数を混成することで得られる相互変調歪や、混変調歪などの特徴を用いることで、更に識別精度が向上できる可能性がある。

3つ目は、学習アルゴリズムの調整である。今回は学習アルゴリズムに関する最適化は意識せず、提案手法の有効性評価を優先した。今後は、様々な学習アルゴリズムにおいて提案手法を比較実験し、最適なものを探すと、ならびに、最適なパラメータを模索することで、更に識別精度が向上できる可能性がある。

参考文献

- [1] Kobayashi, A., Muramatsu, S., Kamisaka, D., Watanabe, T., Minamikawa, A., Iwamoto, T. and Yokoyama, H.: Shaka: User Movement Estimation Considering Reliability, Power Saving, and Latency Using Mobile Phone, *IEICE TRANSACTIONS on Information and Systems*, Vol. E94-D, No. 6, pp. 1153–1163 (2011).
- [2] Ouchi, K. and Doi, M.: Indoor-outdoor Activity Recognition by a Smartphone, *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, New York, NY, USA, ACM, pp. 600–601 (2012).
- [3] Lopez-Tovar, H., Charalambous, A. and Dowell, J.: Managing Smartphone Interruptions Through Adaptive Modes and Modulation of Notifications, *Proceedings of the 20th International Conference on Intelligent User Interfaces*, IUI '15, pp. 296–299 (2015).
- [4] Yue Shi, Yuanchun Shi, J. L.: A rotation based method for detecting on-body positions of mobile devices, *Proceedings of the 13th international conference on Ubiquitous computing(UbiComp2011)*, pp. 559–560 (2011).
- [5] Fujinami, K. and Kouchi, S.: Recognizing a Mobile Phone's Storing Position as a Context of a Device and a User, *In Proc. of the 9th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous2012)*, pp. 76–88 (2012).
- [6] Stephen A. Antosa, Mark V. Albertb, K. P. K.: Hand, belt, pocket or bag: Practical activity tracking with mobile phones, *Journal of Neuroscience Methods*, Vol. 231, No. 15, pp. 22–30 (2014).
- [7] Alanezi, K. and Mishra, S.: Design, implementation and evaluation of a smartphone position discovery service for accurate context sensing, Vol. 44, pp. 307–323 (2015).
- [8] Hoseinitabatabaei, S. A., Gluhak, A. and Tafazolli, R.: Towards a position and orientation independent approach for pervasive observation of user direction with mobile phones, *Pervasive and Mobile Computing*, Vol. 17, No. A, pp. 23–42 (2015).
- [9] Harrison, C. and Hudson, S. E.: Lightweight Material Detection for Placement-aware Mobile Computing, *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, pp. 279–282 (2008).
- [10] Cho, J., Hwang, I. and Oh, S.: Vibration-Based Surface Recognition for Smartphones, *Proceedings of the 2012 IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*, pp. 459–464 (2012).
- [11] Hwang, S. and Wohn, K.: VibroTactor: Low-cost Placement-aware Technique Using Vibration Echoes on Mobile Devices, *Proceedings of the Companion Publication of the 2013 International Conference on Intelligent User Interfaces Companion*, pp. 73–74 (2013).
- [12] Diaconita, I., Reinhardt, A., Englert, F., Christin, D. and Steinmetz, R.: Do you hear what I hear? Using acoustic probing to detect smartphone locations, *Proceedings of the Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pp. 1–9 (2014).
- [13] Diaconita, I., Reinhardt, A., Christin, D. and Rensing, C.: Bleep Bleep!: Determining Smartphone Locations by Opportunistically Recording Notification Sounds, *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pp. 110–119 (2014).
- [14] Diaconita, I., Reinhardt, A., Christin, D. and Rensing, C.: Inferring Smartphone Positions Based on Collecting the Environment's Response to Vibration Motor Actuation, *Proceedings of the 11th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, pp. 99–106 (2015).
- [15] 秋田純一, 小松孝徳: 超音波の反射波情報から対象物の材質を推測する手法の基礎的検証, 研究報告ユビキタスコンピューティングシステム (UBI), Vol. 45, No. 12, pp. 1–6 (2015).
- [16] Martin, K. D.: Sound-Source Recognition: A Theory and Computational Model, PhD Thesis, Massachusetts Institute of Technology (MIT) (1999).
- [17] Breiman, L.: Random forests, *Machine Learning*, Vol. 45, pp. 5–32 (2001).