

宅内でのスマートフォン検索に向けた抽象度別接触物体推定手法の初期的検討

西陽也[†] 石田繁巳[†] 村上友規[‡] 大槻信也[‡]

[†] 公立はこだて未来大学 [‡] NTT アクセスサービスシステム研究所

1 はじめに

宅内でスマートフォンを紛失したユーザは電話をかけるなどしてスマートフォンから音を鳴らし、その音を頼りに探すことがある。しかし、音が反響する宅内でユーザの聴覚のみに頼って探すことから、発見までに時間と労力が必要になる。

この課題に対して著者らはスマートスピーカを用いた音響センシングによるスマートフォン検索支援システムを提案している [1]。本システムではスマートフォンから発せられる音をスマートスピーカで解析することで、スマートフォンが存在する部屋、接触物体、被覆状態を同時に推定し、最終的にスマートフォンの正確な位置を特定する。本稿では提案システムの実現に向けて、スマートフォンから発せられる音の中でもバイブレーション音を用いた接触物体推定手法を提案する。宅内に存在する数千種類の物体を対象に推定を行うことは困難である。そこで本研究では物体推定モデルの他に材質、剛柔推定モデルをそれぞれ用意し、物体推定における推定確信度が低い場合に、より抽象度の高い「材質」や「剛柔」を推定するモデルの推定結果を出力する。実環境で収集したデータを用いて初期的評価を行った結果、抽象度別推定を行うことで推定の誤りを低減できることを確認した。

2 関連研究

スマートフォン内蔵センサを用いたスマートフォン接触物体推定手法として、スマートフォン内蔵の加速度センサ [2]、マイクロフォン [3] を用いた手法などが報告されている。Choら [2] はスマートフォン内蔵加速度センサを用いて接触物体によるスマートフォンの滑りやすさから6種類の物体を85%の精度で推定している。Hwangら [3] はスマートフォン内蔵マイクを用いて接触物体によるバイブレーション音の変動から12種類の物体を91%の精度で推定している。

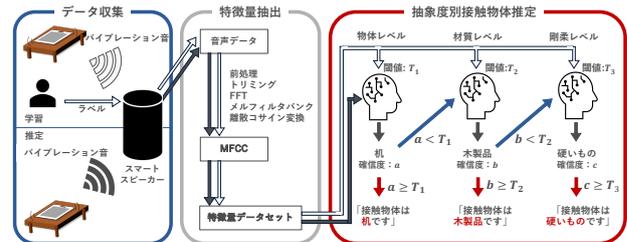


図 1: 抽象度別接触物体推定手法の概要

しかしながら、これらの手法では学習モデルに含まれる物体のみが推定対象であり、宅内に存在する物体を全て学習に用いる必要がある。宅内に様々な物体が存在することを考慮するとこのような学習は困難であり、未知の物体に対する推定も考慮する必要がある。

スマートフォンから発せられた振動や音を別のデバイスで取得し、スマートフォンの接触物体を推定する研究は著者らが調べた限り報告されていない。

3 抽象度別接触物体推定手法

図 1 に抽象度別接触物体推定手法の概要を示す。本手法はデータ収集ブロック、特徴量抽出ブロック、抽象度別接触物体推定ブロックの3つで構成されている。

データ収集ブロックでは、スマートフォンから発せられるバイブレーション音をスマートスピーカ内蔵のマイクで収集する。モデルの学習用データはスマートフォンに通知が届く際のバイブレーション音を収集し、ユーザにそのとき接触している物体をたずねることで正解ラベルをつける。推定時にはスマートフォンをバイブレーションさせてその音声データを収集する。

特徴量抽出ブロックでは、収集した音声データを 1500 ms の長さにトリミングし、MFCC (Mel-frequency cepstrum coefficient) を算出する。

抽象度別接触物体推定ブロックでは、抽出した特徴量を用いて学習モデルの構築と抽象度レベルごとの推定を行う。抽象度レベルとして、物体レベル、材質レベル、剛柔レベルの3つを定義し、レベルごとに学習モデルを構築する。はじめに物体レベルの学習モデルを用いて推定を行う。このとき推定確信度が閾値を上回っていれば推定した物体名を推定結果として出力する。推定確信度が閾値を下回っている場合は材質レベルでの推定を行う。同様にして、材質レベルにおいて

Initial Study of Contact-object Estimation Method by Abstraction Level for Indoor Smartphone Search

Haruya Nishi[†], Shigemi Ishida[†], Tomoki Murakami[‡], Shinya Otsuki[‡]

[†]Future University Hakodate, Japan

[‡]Nippon Telegraph and Telephone Corporation, Japan

[†]{g2123046, ish}@fun.ac.jp

[‡]{tomoki.murakami, shinya.otsuki}@ntt.com



図 2: 実験環境

表 1: 評価に使用した接触物体

物体	材質	剛柔
1 机	金属	硬い
2 iPad	金属	硬い
3 木の板	木	硬い
4 小物入れ	プラスチック	硬い
5 本	紙	硬い
6 布団	衣類	柔らかい
7 ジーパン	衣類	柔らかい
8 スウェット	衣類	柔らかい
9 毛布	衣類	柔らかい
10 トートバッグ	衣類	柔らかい

表 2: 抽象度レベルごとの正解率

抽象度レベル	10-fold CV	LOSO CV
物体	0.942	0.718
材質	0.990	0.918
剛柔	0.966	0.920

表 3: 提案手法の性能評価結果

抽象度レベル	正解率	
	正解率	推定終了した試行割合
物体	0.852	0.408
材質	0.994	0.312
剛柔	0.952	0.126

推定確信度が閾値を上回っていれば推定した材質を推定結果として出力し、下回っていれば剛柔レベルでの推定を行う。全てのレベルで閾値に満たない場合はもう一度データ収集を行い、それでも閾値に満たなければ接触物体推定の結果は出力しない。

本手法ではスマートフォンに通知が届く度にバイブレーション音が収集されるため、データが収集される度に推定モデルの再学習を行う。これにより、日常的にスマートフォンが接触する物体を推定モデルに学習させ、実環境での利用に適した推定モデルを構築できる。

4 評価

実際の宅内環境で収集したデータを用いて初期的評価を行った。

4.1 評価環境

図 2 に実験環境を示す。宅内環境にスマートフォンとマイクロフォンを 1m 離して設置した。スマートフォンを 200 ms と 1000 ms の長さで 1 回ずつ振動させ、そのときのバイブレーション音をオーディオレコーダで記録した。マイクロフォンは audio-technica AT2050、スマートフォンは ASUS Zenfone 8、オーディオレコーダは ZOOM H6 である。音声データはサンプリング周波数 44100 Hz、量子化ビット数 16 bit で記録した。

音声データは表 1 に示す 10 種類の接触物体のそれぞれについて収集した。スマートフォンが接触する物体を変更しながら、各接触物体について音声データを 50 回収集した。このデータ収集を 1 セットとして、時間を空けて 2 セットのデータ収集を行った。

評価では、各抽象度レベルのモデルの性能と、推定確信度に基づいて抽象度レベルを切り替える提案手法の性能を検証した。各抽象度レベルのモデルの評価では、10 分割層化 (10-fold) 交差検証および Leave-One-Session-Out (LOSO) 交差検証により正解率を算出した。提案手法の評価は LOSO 交差検証で行った。提案手法では推定確信度が閾値を超えた場合にその抽象度レベルで推定を終了して結果を出力することから、正解率に加えて各抽象度レベルで推定を終了した試行の割合を算出した。抽象度レベルを変更する推定確信度の閾値は、各レベルのモデルの LOSO 交差検証評価における推定確信度の平均値とした。

4.2 評価結果

表 2 に、抽象度レベルごとの正解率を示す。表 2 に示すように 10-fold 交差検証の正解率は全てのレベルで 0.94 を超えた。

一方、LOSO 交差検証では材質、剛柔レベルでは 0.90 を超えているが、物体レベルでは 0.718 と正解率が大幅に低下した。抽象度レベルが高いほど汎化性能が高く、学習データにない物体や状況に対しても安定して推定可能であることが分かった。

表 3 に、提案手法を用いた場合の正解率と、各抽象度レベルで推定を終了した試行の割合を示す。表 2 の LOSO 交差検証と比較すると、すべてのレベルで正解率が向上していることが分かる。特に、物体レベルにおいては正解率が 0.852 と大幅に向上した。推定終了した試行の割合は物体レベルで 0.408、材質レベルで 0.312、剛柔レベルで 0.126 であり、評価データの約 85% を推定した。推定確信度を用いた抽象度別接触物体推定により高い正解率で推定可能な抽象度レベルを選択することで推定の誤りを大幅に改善できることを確認した。

5 おわりに

本稿では宅内でのスマートフォン検索に向けた抽象度別接触物体推定手法を提案し、その実現可能性を示した。実環境で収集したデータを用いて、各抽象度における推定性能を評価し、提案手法によって推定の誤りを大幅に低減できることを確認した。今後は学習用データ収集機能を実装し、実利用に向けて堅牢な学習モデルの構築を行って提案システムの実現を目指す。

謝辞

本研究の一部は東北大学電気通信研究所における共同プロジェクト研究の助成で行われた。

参考文献

- [1] 西 陽也 他: 音響センシングによる宅内スマートフォン被覆状態分類の精度向上に向けた改善, DPSWS, pp. 144–151 (2023).
- [2] Cho, J., Hwang et al.: Vibration-Based Surface Recognition for Smartphones, *Proc. IEEE RTCSA*, pp. 459–464 (2012).
- [3] Hwang, S. et al.: VibroFactor: Low-Cost Placement-Aware Technique Using Vibration Echoes on Mobile Devices, *Proc. ACM IUI*, pp. 73–74 (2013).