

スマートフォンセンサを用いた学習環境への過剰適合を考慮した屋内位置セマンティクス推定手法

立川 雅也¹ 前川 卓也^{1,2,3} 松下 康之¹

概要：近年スマートフォンを用いた屋内位置推定に関する研究が多く行われており、推定される屋内位置情報のうち、「会議室」や「食堂」といった屋内位置セマンティクス情報はライフログや高齢者の見守り等の様々なアプリケーションのための重要な情報である。一般的な屋内位置のセマンティクスを推定する手法では、対象とする環境ごとに学習用のセンサデータを収集する必要があり導入コストが大きいという問題がある。そこで本研究では、異なる環境において得られた学習データを利用することで、対象とする環境において新たに学習データを収集する必要のない、屋内位置のセマンティクスを推定する手法を提案する。提案手法では、一般的な屋内位置推定で利用されるパッシブセンシングに加えて音声のアクティブセンシングを組み合わせることにより、環境に依存しない特徴量を取得し、他環境から得た学習データを用いて屋内位置セマンティクスを推定する。提案手法では、他環境の学習データに過剰適合しない分類器をランダムフォレスト分類器を基に構築する。評価実験では複数の環境でセンサ情報の取得を行い、提案手法の有効性を検証した。

1. 序論

近年のセンサの小型化、低価格化により日常を映像・音声・位置情報などのデジタルデータとして記録するライフログが注目されており、スマートフォンを用いたライフログに関する研究が多く行われている。一般的な屋内位置推定では位置座標を推定する手法が盛んに研究されているが、ライフログ等への応用を考えると、位置の意味情報（セマンティクス）を取得することが重要である。しかし、屋内位置のセマンティクスを推定する手法の多くは、学習用のセンサデータを環境ごとに収集する必要があるため、導入コストが大きいという問題がある。

そこで本研究では、異なる環境において収集した情報を学習データとし対象とする推定環境に適用することで、対象とする環境において新たに学習データを収集する必要のない、情報収集コストを抑えた屋内位置セマンティクスの推定を行う。提案手法では、一般的な屋内位置推定で利用されるパッシブセンシングに加え、対象に対してセンサ側から何らかの働きかけを行い、その結果対象から得られ

る反応や応答をセンシングするアクティブセンシングを併せて利用することで高精度な推定を目指す。このとき、Wi-Fi 電波を用いてあらかじめユーザのが停留している地点をクラスタリングしたあと、そのクラスタのセマンティクスの推定を行う。これにより、Wi-Fi 電波とクラスタセマンティクスの紐付けができるため、以降はWi-Fi 電波のみでセマンティクスの推定が可能になる。さらに提案手法では、他環境の学習データに過剰適合しない分類器をランダムフォレスト分類器を基に構築する。パッシブセンシングではWi-Fi 電波情報・加速度情報・磁気情報・気圧情報を取得し、アクティブセンシングではスマートフォンから音声を能動的に発生させて得られるインパルス応答を取得する。

2. 関連研究

2.1 Wi-Fi 電波を用いた一般的な屋内位置推定

屋内位置推定に関する研究は多く行われているが、その中でも多くの研究者がWi-Fiのアクセスポイント(AP)を利用した屋内位置推定モデルの構築を試みている。Wi-Fi電波を用いた一般的な屋内位置推定手法の中で最も実用的な手法の1つにフィンガープリンティングによる位置推定がある。フィンガープリンティングの学習フェーズではユーザの位置座標とともにユーザが持つスマートフォン等によって観測されるWi-Fi電波情報(AP固有のMACア

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

² 国際電気通信基礎技術研究所
Advanced Telecommunications Research Institute International (ATR)

³ JST CREST

ドレスと AP からの電波強度) を用いる。AP の組み合わせとそれらの電波強度が位置座標固有の電波情報 (フィンガープリント) となる。収集されたフィンガープリントは位置推定モデルを構築する学習データとして用いる。

2.2 音声のアクティブセンシングを用いた屋内位置セマンティクス推定

音声情報のアクティブセンシングを用いた屋内位置セマンティクス推定手法として、環境音やインパルス応答を利用したものがある。環境音を利用した手法では、事前に部屋ごとの環境音の特徴を学習することで、建物内において部屋単位のセマンティクスの推定を行うが、環境音がほとんど存在しない場合は部屋ごとの特徴が得られず正しい分類が困難になる問題がある [5]。インパルス応答を利用する手法では、屋内環境で得られたインパルス応答は室内音響を表現しているため [2]、事前に推定環境でインパルス応答を測定し、その特徴を学習させることによって部屋単位でのセマンティクスの推定を行う。この手法は、インパルス応答を測定するために測定者が持つ端末から音声を発生させ録音した音声情報を利用するアクティブセンシングを利用するため、環境に関わらず屋内位置セマンティクスの推定が可能である。一方、推定環境ごとに事前学習をする必要があるため学習コストが大きい [4]。

3. 提案手法

3.1 提案手法の概要

提案手法の概要を図 1 に示す。本研究ではユーザが日常生活で長期間停留する地点を対象とし、それらの屋内位置セマンティクスを推定する。学習フェーズでは想定環境とは異なる環境において、パッシブセンシングとアクティブセンシングを利用し磁場情報・気圧情報・音声情報を収集し、それぞれ特徴量を抽出してセマンティクス分類器を学習する。推定フェーズでは想定環境において磁場情報・気圧情報・音声情報に加え加速度情報・Wi-Fi 電波情報を収集し、加速度情報と Wi-Fi 電波情報を用いて停留地点を推定しクラスタリングを行う。さらに、各停留地点で収集した磁場情報・気圧情報・音声情報から学習フェーズと同様に特徴量を抽出しセマンティクス分類器を用いてセマンティクスを推定する。以降の節で学習フェーズと推定フェーズについて詳しく説明する。

3.2 学習フェーズ

本研究では、環境から受動的に情報を取得するパッシブセンシングと、環境に働きかけて情報を得るアクティブセンシングを組み合わせたセンシングを行う。想定環境とは異なる環境の場所クラスが既知である地点においてパッシブセンシングによって磁気情報・気圧情報を、スマートフォンから音声を発生させ録音するアクティブセンシング

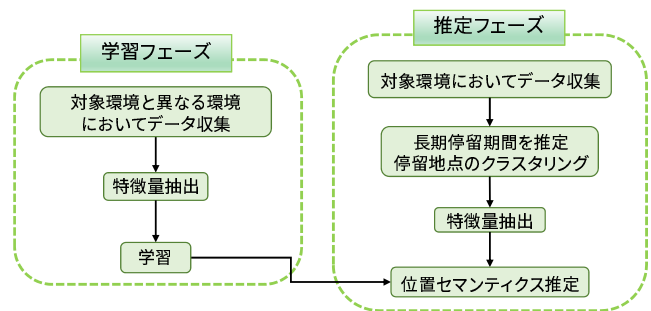


図 1 提案手法の概要

表 1 スマートフォンによるセンシング

センサ	サンプリング周波数	利用目的
加速度	30Hz	停留地点推定
Wi-Fi	0.1Hz	停留地点クラスタリング
磁気	30Hz	位置セマンティクス推定
気圧	30Hz	位置セマンティクス推定
マイク	44.1kHz	位置セマンティクス推定

によって音声情報を収集し、それぞれ特徴量を抽出、結合した特徴ベクトルを用いて分類器を学習する。本研究では分類器にランダムフォレストを基に構築した学習環境への過剰適合を考慮した分類器を用いる。

3.2.1 パッシブセンシングによる特徴抽出

パッシブセンシングにより取得した磁気情報・気圧情報は、それぞれ周辺の電子機器による磁場、エレベータや階段等の高度の変化する移動によって影響を受けると考えられるため、時系列データに対してオーバーラップ 50%、窓幅 0.5 秒の時間窓で求めた移動分散を求め学習する。

3.2.2 アクティブセンシングによる特徴抽出

環境に依存しない普遍的な特徴量を得るため、スマートフォンから音声を発生させそのインパルス応答を測定する。インパルス応答は音響性質を表すものであり、特に同じマイク・スピーカーを用いて様々な環境で測定すると環境ごとの室内音響が表現される。例えばトイレ内の鏡や便器のように、同じ用途の部屋には共通するオブジェクトが存在することが多いため、室内音響を得ることでセマンティクスの推定に有用であると考えられる。インパルス応答はインパルス信号を発生させて得られるが、実世界において直接インパルス信号を発生させると、エネルギーが小さいため信号と雑音の比率である SN 比の小さいインパルス応答が得られてしまう。そのためインパルス信号を利用せずにインパルス応答を測定する手法として、ランダム信号を利用する手法 [1], [6] や時間とともに周波数が変化するサインスイープ信号を利用する手法 [3] などが提案されているが、本研究では SweptSine 法 [3] と呼ばれるサインスイープ信号とサインスイープ信号を時間反転させた信号で畳み込みを行うことによってインパルス応答を得る手法を用いる。本研究では、図 2 のような 20Hz から 20kHz まで周波数が

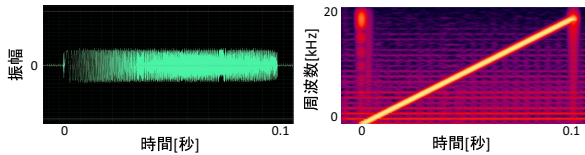


図 2 スweep信号波形

図 3 時間周波数表示

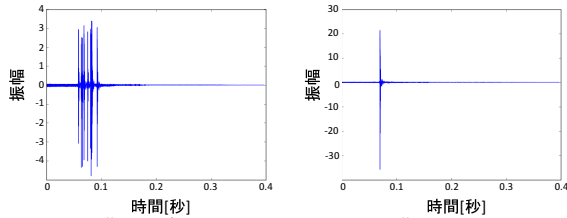


図 4 同期処理無しで得られたインパルス応答

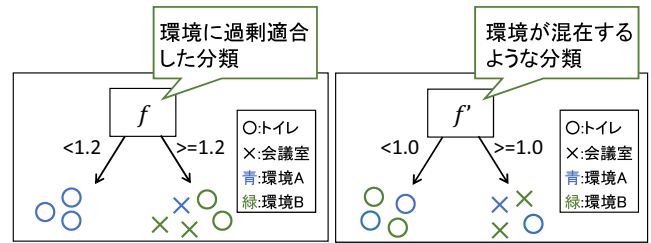
図 5 同期処理ありで得られたインパルス応答

変化する 0.1 秒間のスweepサイン波をスマートフォンから発生させ録音する。スマートフォンから信号を発生させるたびにその時刻を記録する。

周辺の人の声やドアの開閉音などの時間に周期的でないノイズの影響を軽減するため、スweepサイン波は 0.3 秒の間隔で連続再生し、録音された音声を再生したスweepサイン波の時間である 0.1 秒に連続再生の間隔の 0.3 秒を加えた合計 0.4 秒ごとに切り出し加算平均を行う。しかし、スマートフォンが OS に音声を発生するよう命令してから実際に音声が発生するまでに遅延が生じるため、単純に 0.4 秒ごとに切り出した受信信号を加算平均すると図 4 のような不正なインパルス応答が得られてしまう。そこで信号を発生させるときに記録した時刻から 0.4 秒ごとに単純に切り出した波形を時間軸で前後にスライドさせ 1 周期目の波形との差の絶対値を計算し、その値が最も小さくなる波形を加算平均に利用する同期処理を行うことによって正しいインパルス応答を測定する (図 5)。その後、測定したインパルス応答から音声特徴量としてよく用いられるメル周波数ケプストラム係数 (Mel Frequency Cepstral Coefficient: MFCC) による特徴抽出を行い学習する。本研究では、16 のスweep信号に同期加算を適用し得られたインパルス応答から MFCC を抽出する。

3.3 学習環境への過剰適合を考慮した学習

本研究では、ランダムフォレストを基に特定の学習環境への過剰適合をしないような分類器を構築する。ランダムフォレストにおける決定木学習では、データのクラスラベルを基に計算した情報エントロピーを $H(S)$ とすると式 (1)、式 (2) のように表されるクラスラベルのみを考慮したエントロピーの情報利得 $I(S)$ が最大となるように分岐条件 f を決定する。例えば、図 6 の左図の通常のランダムフォレストの分岐条件の情報利得は 1.2 となっており、情報利得は大きい環境 A に過剰適合してしまい、学習環境と異なる環境のテストデータを与えられたときクラスラベルの分類が正しく行われない。



通常のランダムフォレスト

提案手法の分類器

図 6 提案する分岐関数を用いた分類

$$I(S) = H(S) - \sum_{i \in \{L, R\}} \frac{|S^i|}{|S|} H(S^i) \quad (1)$$

$$f = \arg \max I(S) \quad (2)$$

そこで提案手法では、データが得られた環境のラベルを基に計算した情報エントロピーを $H(E)$ とおくと式 (3) のように分岐関数 f' を決定することによって図 6 の右図のような多くの環境に共通する場所クラスの分類ルールを反映するモデルを作成する。

$$f' = \arg \max (I(S) - I(E)) \quad (3)$$

右図の提案手法の情報利得は左図の通常のランダムフォレストの情報利得より小さくなっているが、環境に過剰適合していない。環境ラベルから計算した情報利得 $I(E)$ は、分岐後に同じ環境ラベルを持つデータが多数存在していると大きくなるため、クラスラベルから計算した情報利得 $I(S)$ に $I(E)$ をペナルティとして与えることによって、同じ環境ラベルを持つデータが一方の分岐に偏らないようなクラス分類を行うことができる。

3.4 推定フェーズ

推定フェーズでは対象とする環境の様々な場所クラスにおいて、パッシブセンシングによって磁気情報・気圧情報・加速度情報・Wi-Fi 電波情報、アクティブセンシングによって音声情報 (インパルス応答) を取得する。取得した加速度情報から長期停留期間を推定しその期間内に得られた Wi-Fi 電波情報から停留地点のクラスタリングを行う。また同時に取得した磁気情報・気圧情報・インパルス応答から学習フェーズと同様に特徴量を抽出し、学習フェーズによって作成された分類器を用いて各停留地点の位置セマンティクスの推定を行う。

3.4.1 長期停留期間の推定

取得した加速度情報を用いて長期停留期間を推定する。加速度の時系列データに対してオーバーラップ 75%、窓幅 2 秒の時間窓で移動分散を求め、事前に設定した閾値より分散が小さく、事前に設定した停留地点とみなす期間より長い期間を長期停留期間として推定する。閾値は実際に歩行と静止を繰り返した時の加速度の移動分散の値を考慮して定めた。



図 7 6つの場所クラスの例

表 2 実験環境

環境	建物	竣工年
1	情報科学棟	2004
2	情報科学棟	2015
3	工学棟	1992
4	理学棟	1996

3.4.2 停留地点のクラスタリング

加速度情報によって推定された各長期停留期間ごとに、全ての長期停留期間で観測されたアクセスポイントについてそれぞれ得られた Wi-Fi 電波強度のベクトルを作成する。強度を取得できていない場合はその強度を形式的に -100dBm とする。各長期停留期間ごとに作成されたベクトルを用いて階層的クラスタリングを行う。階層的クラスタリングとは全てのデータをそれぞれ 1 つのクラスと設定しクラス間距離が最も近いクラス同士を併合する処理をクラス数が 1 になるまで繰り返す手法である。本研究では各クラスを各長期停留期間で得られた Wi-Fi 電波強度のベクトルとし、クラス間距離の計算はアクセスポイントごとに強度の差の絶対値を取る絶対平均誤差で求める。本研究では、閾値以下の距離を持つクラスが見つからなくなるまでクラス間の併合を行う、閾値は事前実験の結果から定めた。

3.4.3 位置セマンティクス推定

加速度情報によって推定された各停留期間ごとに取得した磁気情報・気圧情報・インパルス応答から学習フェーズと同様に磁気情報・気圧情報からは移動分散、インパルス応答からは MFCC を特徴量として抽出する。このとき、インパルス応答は 16 周期の信号に対して同期加算を適用し求める。その後、学習環境への過剰適合を考慮した分類器を用いて位置セマンティクスの推定を行う。

4. 評価実験

提案手法の有効性を示すために、図 7 のようなデスク、エレベータ、食堂、休憩所 (屋外喫煙所)、会議室、トイレの 6 つを場所クラスとする評価実験を行った。表 2 に示すような 4 つの異なる環境を実験環境として用いた。

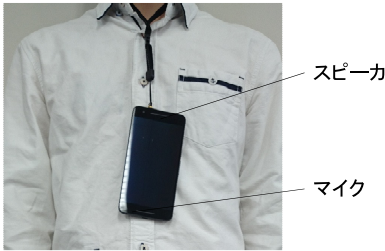


図 8 測定姿勢

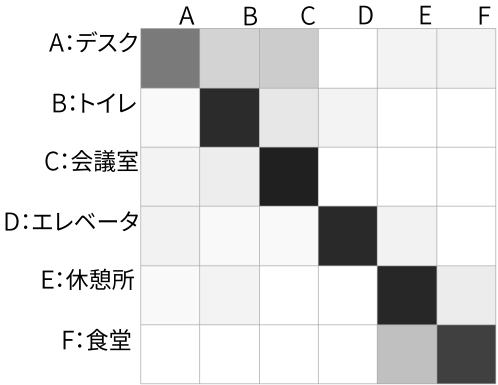


図 9 全ての環境下での推定精度

表 3 全ての環境下での推定精度

場所クラス	適合率	再現率	F 値
デスク [%]	77.8	52.5	62.7
トイレ [%]	72.9	83.3	77.8
会議室 [%]	73.5	87.8	80.0
エレベータ [%]	94.1	84.2	88.9
休憩所 [%]	70.8	85.0	77.3
食堂 [%]	85.7	75.0	80.0

4.1 データセット・評価手法

4 つの異なる実験環境の各環境において、デスク、エレベータ、食堂、休憩スペース (屋外喫煙所)、会議室、トイレの 6 つの場所クラスをランダムに訪れ停留するセッションを合わせて 10 回ずつ行った。各セッションでは図 8 のようにネックストラップにつけたスマートフォン (Nexus6P) で音声・気圧・磁場・加速度・Wi-Fi 電波のセンシングを継続的に行なった。このように取得されたデータを用いて、ある環境で得られたデータをトレーニングデータ、他の環境で得られたデータをテストデータとして評価を行う Leave-One-Environment-Out 交差検定を行った。

4.2 評価結果

提案手法の推定精度を図 9 と表 3 に示す。また、各環境ごとの推定精度を図 10 から図 13 と表 4 に示す。これらの結果から、どの環境においてもほぼ同程度の精度で位置のセマンティクスが推定されていることがわかる。また各場所クラスごとの精度を見ると、デスクの推定精度は他のクラスに比べて低いことが確認できる。この原因として、デスクへの推定に重要な特徴量が各環境ごとに異なっ

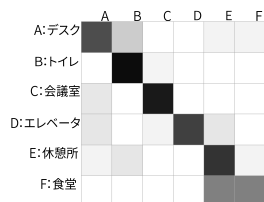


図 10 環境 1 の推定精度

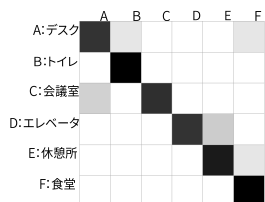


図 11 環境 2 の推定精度

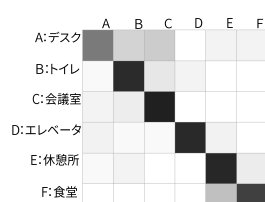


図 12 環境 3 の推定精度



図 13 環境 4 の推定精度

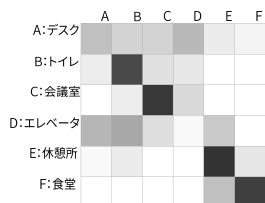


図 14 インパルス応答の MFCC のみを
利用した推定精度



図 15 磁気分散のみを
利用した推定精度

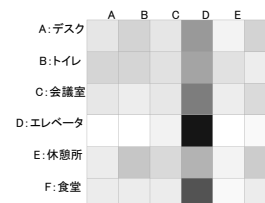


図 16 気圧分散のみを
利用した推定精度

表 4 各環境ごとの分類精度

環境	適合率	再現率	F 値
1	80.1	76.8	76.9
2	89.3	88.6	88.5
3	79.1	78.0	77.8
4	80.7	78.8	79.1

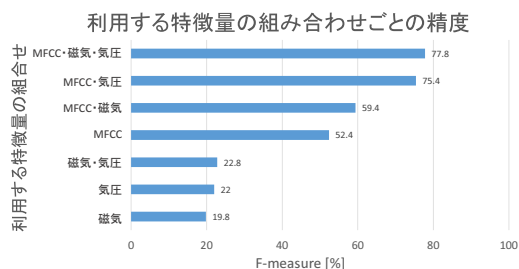


図 17 利用する特徴量の組み合わせごとの推定精度

おり、各環境のデスクに共通する特徴量を取得できていなかったことが挙げられる。しかし、トイレ・会議室・エレベータ・休憩所・食堂は非常に高い精度で推定することができており、全体としては高い精度での位置セマンティクスの推定が可能であると言える。

4.2.1 利用する特徴量による精度の変化

本研究では、分類に利用する特徴量として磁気の分散、気圧の分散、インパルス応答から抽出した MFCC を利用しているが、利用する特徴量の組合せによって精度がどのように変化するかを検証した結果を図 17 に示す。図中の IR はインパルス応答から抽出した MFCC を利用したものを示しており、磁気、気圧はそれぞれ磁気分散、気圧分散を利用したものを示している。この図 17 より、アクティブセンシングによって得られるインパルス応答の MFCC が他のパッシブセンシングで得られた磁気分散・気圧分散に比べ精度の向上に大きく寄与していることがわかる。また、1 つのセンサからの特徴のみを利用したときの分類結果の混合行列を図 14 から図 16 に示す。これらより、アク

同期加算回数による精度の変化

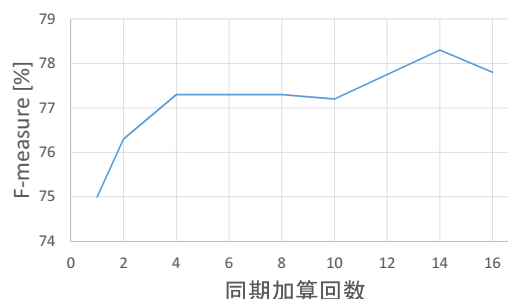


図 18 同期加算回数による精度の変化

ティブセンシングで測定したインパルス応答の MFCC 特徴量は様々な場所クラスの推定に大きく貢献しており、気圧分散は高度の変化するエレベータの推定のための重要な特徴となっていることが確認できた。

4.2.2 スイープ信号の同期加算回数による精度の変化

3.2.2 節のインパルス応答の測定の際、周期的でない雑音の除去のために複数の録音波形を同期加算する処理を行ったが、このとき同期させ加算する録音波形の数を変化させたときの精度の変化を図 18 に示す。これより、同期加算回数を多くすれば精度が高くなっていくことがわかるが、単調増加となっていないことも確認できる。原因として、複数のスイープ信号を再生し録音をしている期間に扉が閉まる音などの突発的な雑音が生じることにより加算回数を増やした方が大きな雑音が生じることがあることが挙げられる。同期加算の回数をさらに多くすると、突発的な大きな雑音の影響が小さくなると予想されるが、同期加算の回数分だけスイープ信号を発信する必要があるため、情報収集コストが大きくなることが考えられる。

4.2.3 クラスタリングの評価

環境 1 における Wi-Fi 電波を用いた階層的クラスタリングにより得られたデンドログラムと各停留地点の正解ラベルの例を図 19 に示す。この図のように同じ場所クラスは近いクラスタとして認識されており、閾値による分割を

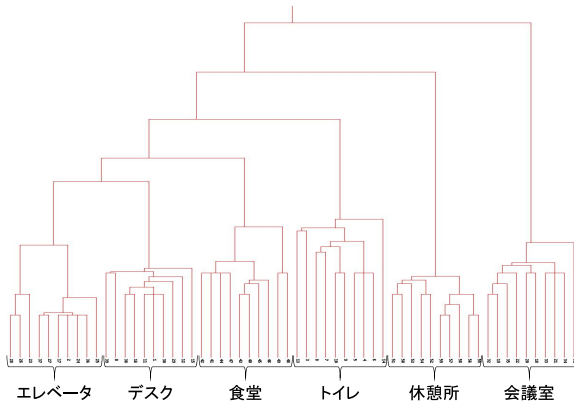


図 19 環境 1 の Wi-Fi データから得られた
デンドログラム

行くと正しくクラスタリングされた。また、環境 1 だけでなく全ての環境について正しく行われていることが確認された。

4.3 通常のランダムフォレスト分類器との精度の比較

3.4.3 節においてランダムフォレストを基に学習環境への過剰適合を考慮した分類器を提案したが、通常のランダムフォレストを利用したときの評価結果と提案手法の分類器を利用したときの評価結果を表 5 に示す。これら 2 つの分類器について、評価実験の 1 セッションごとに得られた F 値を用いて平均値の差で t 検定を用いたところ、両側 10% 水準で有意差が認められた。提案手法の分岐関数は、特定の環境に過剰適合することなく複数の環境に共通するクラス分類ルールを反映する決定木の分岐関数を作成するため、学習環境の総数が少ない場合や学習環境の中に他の環境と比べて得られる特徴量が大きく異なる環境が含まれるような場合に精度が大きく向上すると考えられる。

表 5 ランダムフォレストと提案手法の分類器による評価結果

	適合率	再現率	F 値
通常のランダムフォレスト [%]	76.8	73.0	75.4
提案手法の分類器 [%]	79.1	78.0	77.8

5. 結論

一般的な屋内位置セマンティクス推定手法では対象とする環境ごとに事前学習が必要であるという問題に対して、本研究ではアクティブセンシングとパッシブセンシングを組み合わせることで場所の普遍的な特徴を取得し、対象とする環境ごとに事前学習を必要としない屋内位置セマンティクスを推定する手法を提案した。提案手法では、加速度より推定された各停留地点ごとに、アクティブ・パッシブセンシングにより取得した磁気分散・気圧分散・インパルス応答の MFCC の特徴量を利用し位置のセマンティクスを推定した。さらに特定の学習環境への過剰適合をし

ないような分類器をランダムフォレストを基に構築し用いた。提案手法の有効性を示すために、評価実験を行った。4 つの異なる環境で 6 つの場所クラスをランダムに訪れ停留しながらセンシングを行い、提案手法を用いて位置セマンティクスの推定を行った提案手法の結果、高い精度で屋内位置のセマンティクスを推定することができ、提案手法の有効性を示すことができた。今後の課題としては、分類精度の向上のため、さらに多くの環境における学習データの収集が挙げられる。

謝辞

本研究の一部は、JST CREST の助成を受けて行われたものです。

参考文献

- [1] Borish, J. and Angell, J. B.: An efficient algorithm for measuring the impulse response using pseudorandom noise, *Journal of the Audio Engineering Society*, Vol. 31, No. 7/8, pp. 478–488 (1983).
- [2] Oppenheim, A., Willsky, A. S. and Nawab, S. H.: *Signals and systems*, University of Michigan. Prentice Hall (1997).
- [3] Paulo, J., Martins, C. R. and Bento Coelho, J.: Time Segmented Swept Sine Technique for Room Impulse Response Estimation, *Proceedings of the ICSV*, Vol. 12 (2005).
- [4] Rossi, M., Seiter, J., Amft, O., Buchmeier, S. and Tröster, G.: RoomSense: an indoor positioning system for smartphones using active sound probing, *Proceedings of the 4th Augmented Human International Conference*, pp. 89–95 (2013).
- [5] Tarzia, S. P., Dinda, P. A., Dick, R. P. and Memik, G.: Indoor Localization without Infrastructure using the Acoustic Background Spectrum, *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services (MobiSys '11)* (2011).
- [6] 伊達玄：数論の音響分野への応用，*日本音響学会誌*，Vol. 42, No. 11, pp. 884–893 (1986).