

Wi-Fi CSIを用いた行動認識のための強化学習を用いた 環境非依存なサンプリングレート選択

丹野 友華¹ 前川 卓也¹ 原 隆浩¹ 尾原 和也² 岸野 泰恵² 村上 友規³ アベセカラ ヒランタ³

概要: 近年, Wi-Fi 電波のチャネル状態情報 (Channel State Information: CSI) を用いる行動認識手法が注目されている. CSI を用いた従来の行動認識手法は, 一定の高いサンプリングレートで取得された CSI を利用しているため, 取得や解析に高い計算コストを必要とした. この問題を解決するために, 現在の行動や状態に応じて適切にサンプリングレートを調整しながら, できるだけ低いサンプリングレートで行動認識を行う手法が研究されている. しかし, この既存手法では, CSI が取得される環境ごとに行動認識モデルとサンプリングレートの選択を行うモジュールの学習を行う必要がある. 特に, サンプリングレート選択モジュールの学習手法には強化学習が用いられているため, 多くの時間と計算コストが必要となる. そこで本稿では, 環境に依存せずに動作するサンプリングレート選択モジュールを実現することで, 対象の環境における CSI を用いた行動認識システムの導入コストを低減する. サンプリングレートの選択には, 現在行われている行動の潜在表現である, 行動認識モデルの特徴抽出器の出力を利用することが有効であるが, 行動認識モデルはデータの特徴が異なる環境ごとに用意する必要があるため, 環境非依存なサンプリングレート選択には用いられない. そこで本研究では, 行動認識モデルに Domain Adversarial Neural Networks (DANN) を用いることで, 環境非依存な特徴抽出を行った. また環境によって, 異なるサンプリングレートで認識を行ったときの認識性能が異なるため, 特定の行動をどのサンプリングレートでするかを事前に学習することは不可能である. そこで, 環境に依存しない情報と考えられる行動認識モデルの現在の行動認識結果の信頼度を, サンプリングレート選択モジュールの入力に追加で用いる. これにより, 例えば信頼度が低い場合は高いサンプリングレートに切り替えるなどの環境非依存な方策を学習できると期待される. 3つの環境で, それぞれ7種類の行動を行った際の CSI を取得し, 評価実験を行った.

1. はじめに

近年のセンシング技術や情報処理技術の発達により, 実世界から得られたセンサデータを解析し, 人の行動を認識する研究が盛んに行われている. 行動認識は, データを基に対象の行動の種類を推定する認識技術であり, 高齢者の見守りや防犯の目的で利用されている. これまでに様々なセンサデバイスを用いた行動認識手法が研究されており, 近年注目されている手法として, Wi-Fi 電波のチャネル状態情報 (Channel State Information: CSI) を解析することによって行動認識を行う手法がある [1], [2]. チャネル状態情報とは, 無線 LAN 通信を行う送受信機の物理層から得られる電波の伝搬路に関する情報であり, 各送受信アンテナ間およびサブキャリア間の振幅変化と位相変化の情報として得られる. Wi-Fi 電波を用いる行動認識手法には, 従来の手法に対して, プライバシーの問題を軽減できる点や,

ウェアラブルデバイスの装着を必要としない点などの利点がある.

Wi-Fi の CSI を用いた従来の行動認識手法は, 一定の高いサンプリングレートで取得されたデータを利用していたため, データ取得と解析に高い計算コストを必要とした. この問題を解決するために, 現在の行動や状態に応じて適切にサンプリングレートを調整しながら, なるべく低いサンプリングレートで行動認識を行う手法が研究されている [3]. この研究において提案されたサンプリングレート選択モジュールは, 例えば環境内に人がいないときや, 単純な行動の認識を行う際にはサンプリングレートを低く抑え, 認識精度を下げることなく, 計算コストを抑えることができる. この研究ではサンプリングレート選択モジュールに強化学習エージェントが用いられており, サンプリングレートの調整は行動認識モデルの中間出力を利用して行われる. この手法の概略を図 1 に示す.

しかし, この手法では, 強化学習エージェントを対象とする環境ごとに学習することを前提としており, そのため

¹ 大阪大学大学院情報科学研究科

² NTT コミュニケーション科学基礎研究所

³ NTT アクセスサービスシステム研究所

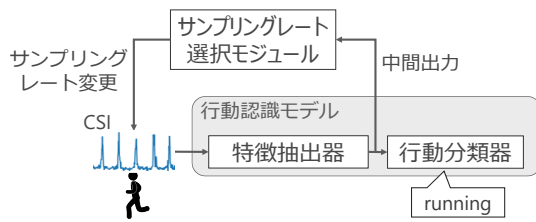


図 1 従来のサンプリングレート選択手法の概略 [3]

のデータ取得や学習に大きなコストがかかる。部屋（環境）のレイアウトや家具の位置が異なると、その環境で得られる CSI も大きく値が変化するため、ある環境で取得した CSI で行動認識モデルとサンプリングレート選択モジュールを学習しても、それを別の環境で取得した CSI に対して使用することはできない。特に強化学習によって学習されるサンプリングレート選択モジュールの学習には多くの時間と計算コストがかかるため、環境ごとに用意することは困難である。

そこで本研究では、環境非依存なサンプリングレート選択モジュールを実現することで、対象環境への行動認識システムの導入コストを低減することを目指す。環境非依存なサンプリングレート選択モジュールでは、サンプリングレートの選択基準を環境非依存となるよう設計する必要がある。従来手法では、環境ごとに用意した行動認識モデルの中間出力（環境の状態の潜在表現）をサンプリングレート選択モジュールの入力に用いていたため、その選択基準は環境依存となる。一方、本研究では、主に (1) 環境に依存しない行動認識モデルの潜在表現および (2) 行動認識の信頼度に注目したサンプリングレート選択を行う。すなわち、図 1 における行動認識モデルの特徴量抽出器の出力を環境非依存にすることで、サンプリングレート選択モジュールを環境非依存にする。これにより、環境内の状態に関する環境非依存な表現を用いたサンプリングレート選択が可能になる。また、特徴量抽出器とサンプリングレート選択モジュールが環境非依存となれば、CSI を取得する環境が変化した際には図 1 における行動分類器のみを再学習すればよい。

本稿の構成は以下の通りである。まず、2 節にて CSI を用いた行動認識に関する関連研究を紹介する。次に、3 節では提案手法について説明する。そして、4 節にて、提案手法の評価実験と、その結果について報告する。

2. 関連研究

2.1 CSI を利用した環境非依存な行動認識

CSI は、取得環境や被験者の違いにより、値が大きく変化してしまう。そのため、ある取得環境で、ある被験者によって得られた CSI で学習された行動認識モデルは、異なる取得環境で、異なる被験者によって得られた CSI ではうまく機能しないことが一般的である。Jiang ら [4] は、取得

されたデータに含まれる環境や被験者固有の情報を除去し、環境や被験者に依存しない特徴量を抽出できる、深層学習ベースのデバイスフリー行動認識フレームワーク「EI」を提案した。EI では、ある取得環境で、ある被験者によって得られたラベル付きデータと、異なる取得環境で、異なる被験者によって得られたラベルなしデータの両方を学習に使用する。また、全てのデータは、どの取得環境で、どの被験者によって取得されたかが既知であるとする。入力データから、畳み込みニューラルネットワーク (Convolutional neural network: CNN) を用いて特徴量抽出を行い、その特徴量を、全結合層で構成された行動分類器へと入力する。また、CNN によって得られた特徴量と、行動分類器の出力を結合し、全結合層で構成されたドメイン判別器へと入力する。ここで、ドメインとは、データの取得環境と被験者の組のことである。そして、行動認識精度は上げて、ドメイン分類精度は下げるように学習を行うことで、CNN が環境や被験者に依存しない特徴量を抽出できる。また、ラベルなしデータに関しては、より自信をもった推定ができるように損失関数が設計された。それに加えて、モデルの性能の向上のため、損失関数にはいくつかの制約項が加えられた。Wi-Fi や、超音波などを利用した、4 種類のデバイスフリーの行動認識手法に関して実験が行われ、EI の有効性と汎用性が実証された。

この手法は、CSI を用いて環境非依存な行動認識を行うことができるが、新しい未知のドメインに適応するために再学習をすることを前提としている。行動認識モデルの再学習を行うと、1 節で述べたように、サンプリングレート選択モジュールも再学習を行う必要があるため、本研究には用いることができない。

2.2 CSI を利用した行動認識のためのサンプリングレート選択

行動認識のための CSI データの取得を、ある程度の広さをもった空間（例えば複数の部屋から構成される家庭環境）で行うことを考える。このとき、Wi-Fi 電波の送受信機を複数設置する必要がある。その台数に比例したデータ量の CSI データが収集される。通常は、それらのデータを 1 つの計算機でまとめて解析することが多いが、2 節で紹介した手法を含めた CSI による従来の行動認識手法では、一定の高いサンプリングレートで取得された CSI が利用されており、それらのデータをまとめて解析するには高い計算コストが必要となる。CSI のデータサイズは CSI のサンプリングレートに依存するため、例えば環境内に人がいないときや、単純な行動の認識を行う際にはサンプリングレートを低くすることで、認識精度を下げることなく、データサイズを抑えることができると考えられる。そこで、現在の行動や状態に応じて適応的にサンプリングレートを調整しながら、なるべく低いサンプリングレートで行動認識を行

う手法が提案された [3]。これらの手法では、サンプリングレートの選択に強化学習が用いられた。強化学習における「行動」をサンプリングレートの選択とし、低いサンプリングレートで正しく行動認識できた際に高い「報酬」を与える。また、「状態」として、現在の行動に関する情報を反映すると考えられる、CSI を用いた行動認識モデルの中間出力を用いる。任意のサンプリングレートの入力を可能とする行動認識モデルを設計し、その中間層で獲得される、サンプリングレート非依存な中間表現を「状態」として用いている。2つの環境で取得された CSI に対して、単一のサンプリングレートで認識した場合と比較し、認識精度を維持しつつ、計算コストを低減できた。

これらの手法は、2.1 節で述べたように、未知のドメインで取得された CSI に対しては行動認識モデルがうまく機能しない。同様に、その中間出力を入力としているサンプリングレート選択モジュールもうまく機能しない。サンプリングレート選択モジュールは強化学習のエージェントであり、未知のドメインのデータに適応するための再学習には多くの時間と計算コストがかかってしまう。そこで本研究では、環境非依存なサンプリングレート選択モジュールを実現することで、対象環境への行動認識システムの導入コストを低減することを目指した。

3. 提案手法

3.1 提案手法の概要

本研究では、屋内に Wi-Fi 電波の送信機と受信機がそれぞれ 1 つずつ設置されている環境を想定する。送信機が Wi-Fi パケットを送信し、それを受信機が受信する際に CSI が得られる。本研究では送信機のパケットの送信頻度を、CSI のサンプリングレートとする。提案手法は、環境内の人の行動の種類を推定対象とし、同時に、次の時刻で使用すべきサンプリングレートを選択する。

図 2 に示すように、提案手法は行動認識モデルとサンプリングレート選択モジュールから構成される。また、提案手法は、任意の複数の環境で取得された CSI を入力とする。行動認識モデルとサンプリングレート選択モジュールの学習を事前に行う環境を学習環境、学習したモデルとモジュールにより行動認識を行う環境を対象環境とする。

行動認識モデルは、CSI を入力として、行動ラベルを推定するように学習されたニューラルネットワークであり、特徴抽出器と行動分類器の 2 つに分けられる。入力とする CSI は、指定された 3 種類のサンプリングレートのうちのいずれかで取得されたものとする。学習環境で取得された CSI を用いて学習を行う際は、Domain Adversarial Neural Networks (DANN)[5] を利用し、特徴抽出器が入力 CSI の取得環境に依らず機能するようにする。

サンプリングレート選択モジュールは、強化学習によって実現される。CSI から得られた現在の行動に関する情報

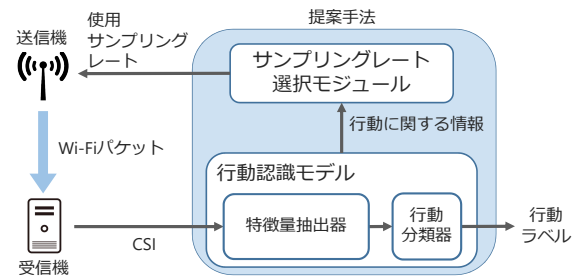


図 2 提案手法の概要

を強化学習における「状態」とし、それをもとに、CSI を観測するサンプリングレートを指定された 3 種類の中から決定する。学習の際の「報酬」は行動認識モデルの推定行動の正誤と、推定に使用したサンプリングレートをもとに値を決定する。また、環境に依存しないサンプリングレート選択を可能とするため、現在のサンプリングレートおよび現在のサンプリングレートより低いサンプリングレートを仮に用いた場合の認識の信頼度の推定値を「状態」に含める。

提案手法では、サンプリングレートを選択したあと、選択されたサンプリングレートで計測された CSI は行動認識モデルにより認識されると同時に、その中間出力などを用いて、次の時刻のサンプリングレートを決定する。この手順を繰り返すことにより、逐次的に行動認識を行う。また、対象環境で取得された CSI を入力とするときは、学習環境で取得された CSI を用いて学習を行ったモデルの、行動分類器のみを再学習したものを使用する。学習に大きなコストを要するサンプリングレート選択モジュールは、学習環境にて学習されたものを用いる。

3.2 行動認識モデル

先行研究 [3] のサンプリングレート選択モジュールの入力は行動認識モデルの特徴抽出器の出力である。そのため、環境非依存なサンプリングレート選択モジュールの実現のためには、特徴抽出器による環境非依存な特徴の抽出が必要である。そこで本研究では、行動認識モデルの設計に DANN を導入することで、特徴抽出器による環境非依存な特徴の抽出を行う。DANN は、異なるドメイン (例：環境や被験者) から得られたデータから、ドメイン非依存な特徴抽出を行うニューラルネットワークを実現する。ニューラルネットワークは特徴抽出器と 2 つの分類器 (ドメイン分類器とクラス分類器) から構成され、学習の際は複数の既知のドメインのデータを使用する。このとき、特徴抽出器の出力を両方の分類器に入力し、ドメイン分類精度のみを下げるように、すなわち、入力されたデータがどのドメインのものなのかを識別できないように学習する。これにより、特徴抽出器はドメイン非依存な特徴を抽出するように学習され、未知のドメインのデータでクラス分類を行う

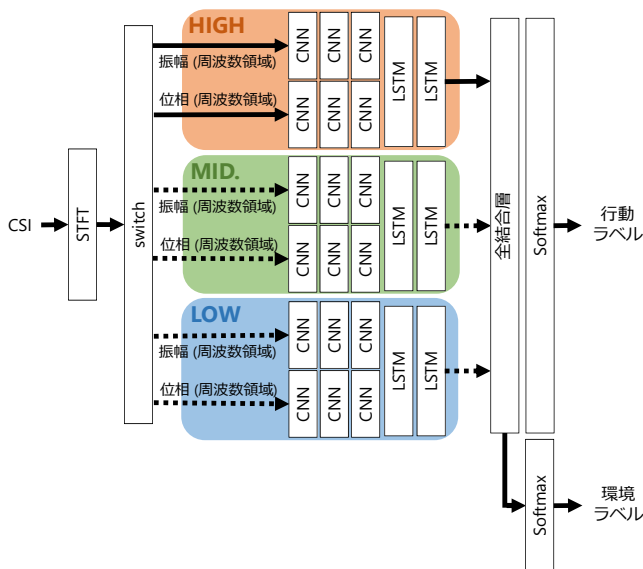


図3 提案手法の行動認識モデル

際の助けになると考えられる。提案手法では、ドメインをCSIの取得環境とした。また、行動の際の身体部位の動きの周期などの周波数にかかわる情報はCSIの取得環境に依存しないと考え、行動認識モデルの入力として、CSIに短時間フーリエ変換 (Short-term Fourier Transform: STFT) を適用した結果を用いた。

学習環境で取得されたCSIから学習される、提案手法の行動認識モデルの構造を図3に示す。先行研究[3]の行動認識モデルと同様に、取得されたCSIデータに、スライド幅が1秒で窓幅が5秒のスライディングウィンドウを適用し、ウィンドウごとに中央の時刻の行動を推定する。したがって、サンプリングレートを s [Hz]とすると、1つのウィンドウ中でのCSIの取得回数、すなわちサンプル数は、 $5s$ となる。よって、1ウィンドウのCSIの次元数は $5s \times 3 \times 1 \times 30$ となる。入力に用いるCSIのサンプリングレートは、HIGH, MIDDLE, LOWの3種類のうちの1つとし、HIGHは200Hz, MIDDLEは100Hz, LOWは25Hzとした。前処理として、1ウィンドウ分のCSIデータを $5s \times 90$ の行列に変形し、CSIの位相のノイズ除去[6]を行う。ここで、 $5s \times 90$ の1次元目は時間を表しているため、変形されたCSIデータは、90次元の時系列データの集合と捉えることができる。その次元ごとに、振幅と位相の標準化とSTFTを行い、その結果をニューラルネットワークへ入力する。200, 100, 25HzのCSIに対するSTFTのウィンドウサイズは、それぞれ256, 128, 32サンプル、そのオーバーラップはそれぞれ128, 64, 16サンプルとなる。STFTにより、1つの時系列データの入力に対し、時間×周波数の2次元のデータが得られるため、時間×周波数×90の3次元のデータがニューラルネットワークの入力となる。

CSIを処理するCNN層のカーネルサイズは連続する3層間で等しく、HIGH, MIDDLE, LOWでそれぞれ 3×20 , 3×10 , 2×5 とした。また、全てのCNN層において、スライド幅は、1次元目の方向に1, 2次元目の方向に1とし、カーネル数は32で、活性化関数はReLU関数とした。その後のLong short-term memory (LSTM)層のノード数は128, 活性化関数はsigmoid関数とした。1つ目のLSTM層は全時刻, 2つ目のLSTM層は最終時刻のみ出力を行う。最終のLSTM層の出力は、サンプリングレート間で共通の全結合層へ入力される。1つ目の全結合層のノード数は64, 活性化関数はsigmoid関数である。ここで、提案手法では敵対的学習を行うため、1つ目の全結合層の出力を、行動ラベルの出力層と環境ラベルの出力層の両方に入力し、行動分類と同時に環境分類も行う。行動ラベルと環境ラベルの出力層は全結合層であり、活性化関数はSoftmax関数である。学習環境で取得されたCSIを用いてこのモデルを学習する際、誤差逆伝搬法において、環境ラベルの出力層の勾配を反転させて逆伝搬することで、環境分類性能は下げるように学習を行う。これにより、1つ目の全結合層は、行動の種類には依存するが、取得環境には依存しないような特徴量の抽出ができるようになる。この1つ目の全結合層を行動認識モデルの中間出力層とし、その出力をサンプリングレート選択モジュールへの入力に用いる。

3.3 サンプリングレート選択モジュール

提案手法のサンプリングレート選択モジュールは強化学習によって実装される。強化学習では、ある環境下で、エージェントが現在の「状態」を観測し、「行動」をする。エージェントは行動をする度に、それに応じた「報酬」が得られ、一連の行動の中での報酬の総和を最大化する行動方策を学習する。本研究の場合は、「行動」を適切なサンプリングレートの選択とし、「状態」を、観測したCSIから得られる情報とした。また、なるべく低いサンプリングレートでの認識を実現するために、エージェントが低いサンプリングレートで正しく行動認識できた際は高い「報酬」を与えた。提案手法のサンプリングレート選択モジュールは、学習環境で取得されたCSIのみを用いて学習され、対象環境で使用する際には再学習が不要となるように設計されている。

3.3.1 行動と状態の設計

「行動」の値は0, 1, 2の3種類で、行動認識モデルの入力となるCSIのサンプリングレートを決定する。0, 1, 2はそれぞれ行動認識モデルにおけるLOW, MIDDLE, HIGHに対応する。

「状態」としては、以下の4つの値を用いた。

- 行動認識モデルの中間出力
- サンプリングレートの変更が連続で遅延している回数

- 前の時刻での選択サンプリングレート
- 現在のサンプリングレートおよびそれより低いサンプリングレートを仮に用いた場合の認識の信頼度の推定値

学習済みである行動認識モデルの中間出力とは、3.2節で説明した行動認識モデルの全結合層の出力であり、64次元ベクトルである。提案する行動認識モデルの中間出力は環境非依存な表現である。サンプリングレートの変更が連続で遅延している回数と前の時刻での選択サンプリングレートは、先行研究 [3] で用いられているものと同様であり、サンプリングレートの変更の遅延に対応した学習を可能とするために用いられる。サンプリングレートの変更の遅延とは、サンプリングレート選択モジュールが使用するサンプリングレートを選択したあと、選択されたサンプリングレートで取得された CSI による行動認識を直ちに次の時刻から行うことが不可能となる場合があり、その際に起こることである。サンプリングレートの変更が連続で遅延している回数は 0 から 4 の整数であり、前の時刻での選択サンプリングレートは、エージェントの「行動」の値に対応する 0, 1, 2 のいずれかの値である。ただし、サンプリングレート選択モジュールへの入力時は、サンプリングレートの変更が連続で遅延している回数は 4 で割った値、前の時刻での選択サンプリングレートは 2 で割った値を用い、正規化を行った。

3.2節で説明したような、CSI を取得した環境に依存しない特徴量を「状態」に用いるが、それだけでは環境非依存なサンプリングレート選択を行うことはできない。なぜなら、対象環境では、サンプリングレートの変更により精度が変化する行動クラスや精度変化の度合いも未知であるため、学習環境で学習された方策では適切なサンプリングレートを選択できないためである。一方、行動認識結果に自信がないときは、より高いサンプリングレートを選択するといった選択方策は CSI を取得する環境に依存しない。そこで提案手法では、現在のサンプリングレートおよび現在のサンプリングレートより低いサンプリングレートを仮に用いた場合の認識の信頼度の推定値を「状態」に用いる。これによりエージェントは、現在のサンプリングレートより低いサンプリングレートの推定認識信頼度が十分に高い場合は、そのサンプリングレートにまでサンプリングレートを下げることでコストの削減が可能である、と学習すると期待される。推定認識信頼度としては、現在の CSI を入力したときの行動認識モデルの出力ベクトルの最大の要素の値を用いた。この値は、行動認識モデルが最も分類確率が高いと推定した行動の分類確率を表しており、行動認識モデルによる認識の自信の度合を示すと考えられる。また、現在選択されているサンプリングレートより低いサンプリングレートの推定認識信頼度は、観測された CSI をダウンサンプリングすることによって得る。サンプリング

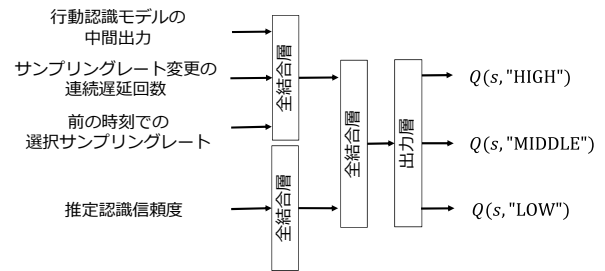


図 4 提案手法のサンプリングレート選択モジュール

レート選択モジュールへの入力時は、HIGH, MIDDLE, LOW それぞれの推定認識信頼度の、3次元ベクトルを用いた。ただし、現在のサンプリングレートより高いサンプリングレートの推定認識信頼度に対応する値は推定できないため、-1 と設定した。

3.3.2 報酬の設計

報酬の設計は、先行研究 [3] と同様である。正しく行動が認識できた場合に高い値の報酬を設定する。同時に、行動認識処理に必要なデータサイズを小さくするため、低いサンプリングレートで認識するほど、高い値の報酬を設定する。具体的な報酬 R の式を以下に示す。

$$R = \alpha R_{acc} + (1 - \alpha) R_{cost}$$

α は行動認識性能の向上と認識コストの削減のバランスを表すパラメータである。0 から 1 の実数であり、1 に近いほど行動認識性能の向上を優先し、0 に近いほど認識コストの削減を優先する。 R_{acc} は行動認識精度に関する報酬で、正しい行動認識が行われた場合は 1、そうでない場合は -1 とする。 R_{cost} は認識コストに関する報酬で、LOW, MIDDLE, HIGH が選択された場合、それぞれ 1, β , -1 とする。 β は、-1 から 1 の実数であり、LOW と HIGH に対する MIDDLE の価値を定めるパラメータである。報酬 R の値は、学習の安定のために -1 から 1 の値となるよう設計した [7]。

3.3.3 エージェントの構造

提案手法では、強化学習のモデルとして Deep Q Network [7] を使用した。エージェントとして、図 4 に示すような、全結合層からなるニューラルネットワークを用いた。入力は「状態」である、(1) 行動認識モデルの中間出力である 64次元ベクトルと、(2) サンプリングレートの変更が連続で遅延している回数、前の時刻での選択サンプリングレート、現在のサンプリングレートおよびそれより低いサンプリングレートを仮に用いた場合の推定認識信頼度を連結した 5次元ベクトルの 2つであり、まずはそれぞれに用意された全結合層によって処理される。その後、それらの処理結果を連結し、さらに全結合層によって処理する。全ての隠れ層のノード数は 32 であり、最終出力層のノード数は「行動」の種類の数と同じ 3 である。

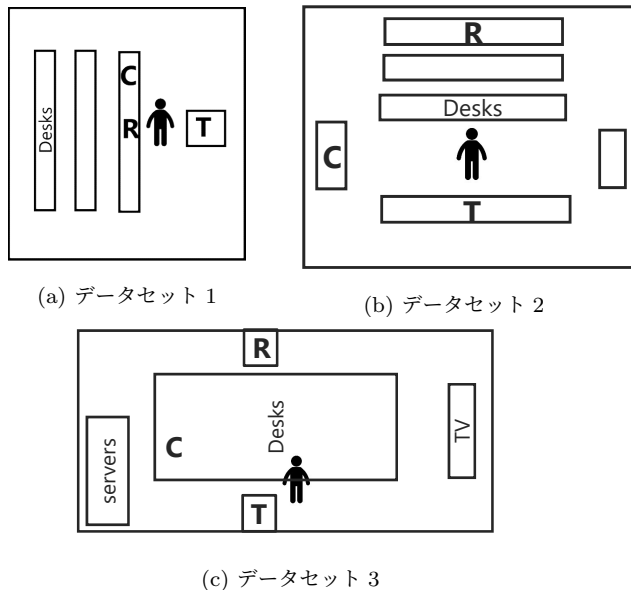


図 5 評価実験のデータ取得環境

3.3.4 エージェントの学習

提案手法では、任意の環境で取得された CSI に対して適切なサンプリングレート選択を行うために、学習環境で取得された CSI にランダムにノイズを付与したものをサンプリングレート選択モジュールの学習に用いる。これにより、エージェントは様々な環境下で観測された CSI に関する経験を疑似的に得ることができる。また、Experience Replay の手法として、Prioritized experience replay[8]を用いた。経験を取得する方法としては、 ϵ -greedy 法を用いた。通常、経験をj得る際、エージェントは現時点で最も最適な「行動」をとる。しかし、この方法では未知の「行動」に対する経験をj得ることができないため、 ϵ -greedy 法では、エージェントは ϵ の確率でランダムな「行動」をとり、経験をj取得する。

また、より効率的な学習のため、Rainbow[9] モデルをもとに学習方法に工夫を行った。

4. 評価実験

4.1 データセット

図 5 に示す、3 つの環境で取得したデータで評価実験を行った。図 5(a), (b), (c) の環境で取得されたデータセットを、それぞれデータセット 1、データセット 2、データセット 3 とする。図 5 中の T は送信機、R は受信機、C は正解ラベル取得用カメラである。送信機と受信機は Intel 5300 NIC Wi-Fi モジュールを搭載した PC である。それぞれに Halperin らによって公開されている CSI ツール [10] をインストールし、それを用いて CSI を取得した。送信機のアンテナは 3 つ、受信機のアンテナは 1 つであり、搬送波の数は 30 である。また、データ取得時の CSI のサンプリングレートは約 4000Hz に設定した。

データセットは、各環境 12 セッション分の CSI と、カメ

ラ映像をもとにした行動ラベルデータで構成されている。1 セッション中では、表 1 に示すような 7 種類の行動をランダムな順番で 20 秒間ずつ行った。

表 1 データ取得時に行った行動

行動	ラベル名
歩く	walk
歯を磨く	teeth
髪をとく	hair
掃除機をかける	vacuum
腹筋運動をする	ab
スクワットをする	squat
反復横跳びをする	side

4.2 環境非依存な行動認識に関する評価実験

既存手法の行動認識モデルの認識性能と、提案手法の行動認識モデルの認識性能を比較し、評価するための実験を行った。

4.2.1 評価方法

取得した CSI データを間引くことで、1, 5, 25, 100, 200Hz のデータを作成し、実験を行った。実験方法として、Own と Others の 2 つを用意した。Own は、各取得環境で取得した学習データを用いて学習、テストを行う方法である。Own では、対象環境への適応を行わないため、提案手法の行動認識モデルの、環境分類器は使用しない。Others は、他の環境で取得した学習データを用いて行動認識モデルの学習を行った後、対象環境で取得したラベルありデータを用いて行動分類器のみを再学習し、テストを行う方法である。例えば、データセット 1 の取得環境を対象環境とした場合、データセット 2 とデータセット 3 の学習データにより行動認識モデルとサンプリングレート選択モジュールの学習を行った後、データセット 1 の学習データを用いて行動分類器のみを再学習し、データセット 1 のテストデータを用いてテストを行う。各データセットのセッション 1 から 9 を学習データ、セッション 10 から 12 をテストデータとした。また、既存手法 [3] の行動認識モデルへの入力としては、1, 5, 25Hz の CSI の時間領域の情報を、提案手法の行動認識モデルへの入力としては、25, 100, 200Hz の CSI の周波数領域の情報を使用した。行動分類と環境分類の損失関数には交差エントロピーを使用し、最適化手法には Adam[11] を用いた。学習は、学習環境で取得された CSI を用いた場合には 200 エポック、対象環境で取得された CSI を用いた場合にも 200 エポック行った。また、評価指標として、各行動のマクロ平均 F 値を全てのサンプリングレートで平均したものをを用いた。

4.2.2 評価結果

実験の結果を図 6 に示す既存手法の行動認識モデルでは、Others のマクロ平均 F 値は、Own のマクロ平均 F 値

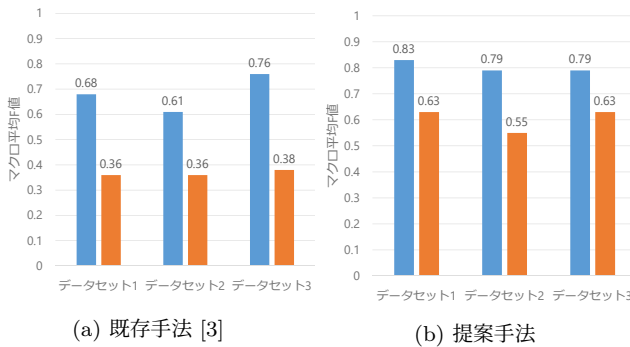


図 6 データセットごとのマクロ平均 F 値

の 50%から 60%程度となったのに対し、提案手法では 70%から 80%となった。すなわち、提案手法の行動認識モデルは、既存手法に比べて環境非依存な特徴を抽出することができたと言える。

4.3 環境非依存なサンプリングレート選択に関する評価実験

4.3.1 評価方法

データセット 2 とデータセット 3 の取得環境を学習環境、データセット 1 の取得環境を対象環境として、対象環境における行動認識時のサンプリングレート選択の評価を行った。提案手法の有効性を検証するために下記のような比較手法を用意した。

- HIGH：サンプリングレートとして 200Hz のみを利用した手法
- MIDDLE：サンプリングレートとして 100Hz のみを利用した手法
- LOW：サンプリングレートとして 25Hz のみを利用した手法
- 既存手法：既存手法 [3] のサンプリングレート選択モジュールを用いた手法

行動認識モデルは、どの比較手法でも前節の実験で学習したものを用いた。サンプリングレート選択モジュールの学習には、学習環境であるデータセット 2 とデータセット 3 のセッション 10 から 12 を用いた。また、対象環境であるデータセット 1 のセッション 10 から 12 をテストに用いた。経験の取得は、以下の 2 つの手順を繰り返すことにより行われる。

- (1) 学習に用いるセッション中からランダムに 1 セッションを選択
- (2) 選択セッション中でサンプリングレートの選択を逐次的に実行

この 2 つの手順を 1 度行うことを 1 エピソードとし、サンプリングレートの選択を 1 度行うことを 1 ステップとする。学習は 168000 ステップ行った。Replay buffer のサイズは 14000 とし、経験が 140 回取得されてから学習を行った。学習の際のバッチサイズは 64 とした。また、ターゲット

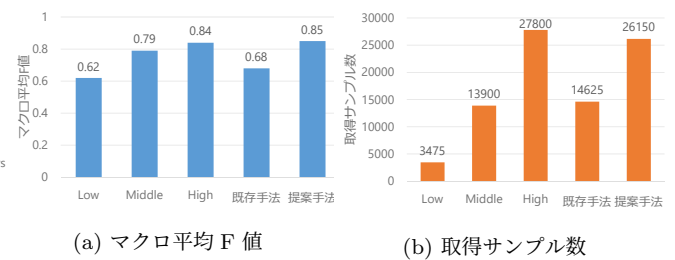


図 7 セッション 10 における実験結果

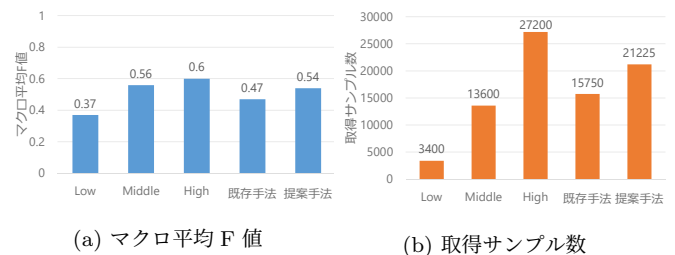


図 8 セッション 11 における実験結果

ネットワークは、28000 ステップごとに最新のものに更新した。ε-greedy 法の ε は、初期値を 1 とし、84000 ステップ目まで線形に 0.1 まで減少して、それ以降は 0.1 になるように設定した。3.3.2 節にて説明したパラメータについては、α を 0.9、β を以下の式で表される値とした。

$$\beta = \frac{\sqrt{HIGH} - 2\sqrt{MIDDLE} + \sqrt{LOW}}{\sqrt{HIGH} - \sqrt{LOW}}$$

ここで、HIGH, MIDDLE, LOW は、それぞれ HIGH, MIDDLE, LOW のサンプリングレートの値である。損失関数には最小二乗誤差を使用し、最適化手法には Adam[11] を用いた。3.3.4 節にて述べたように、学習に用いるセッションの CSI にはランダムにノイズを付与した。ノイズは、STFT 適用後の CSI に付与する。まず、STFT 後の CSI の各要素ごとに、0 以上 M 未満の一様分布の乱数を生成する。M はエピソードの開始時に 2 以上 5 以下の整数からランダムに設定される。その後、生成した乱数を、STFT 後の CSI の各要素に乗算する。上記の手法はウィンドウごとに行動の推定結果を出力するため、行動認識性能の評価指標として、それらを用いて計算したマクロ平均 F 値を用いる。また、計算コストを評価する指標として、テストデータを認識するために取得した総サンプル数を用いる。

4.3.2 評価結果

対象環境のセッション 10, 11, 12 における実験結果をそれぞれ図 7, 図 8, 図 9 に示す。提案手法は、セッション 10 において、HIGH より小さい取得サンプル数で、HIGH より大きいマクロ平均 F 値を達成した。この結果より、提案手法は、対象環境でも行動認識性能を保ちながらコストの低減を達成するような、適切なサンプリングレートを選択できたと言える。

しかし、セッション 11 においては、取得サンプル数は

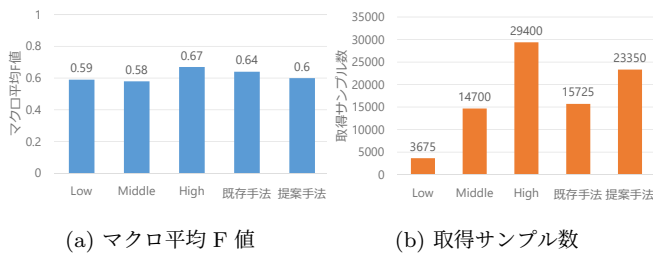


図 9 セッション 12 における実験結果

MIDDLE よりも大きく、マクロ平均 F 値は MIDDLE よりも小さくなった。セッション 11 においては、LOW の行動認識性能が低く、行動認識モデルの最終出力において、正解の行動以外の行動の分類確率が誤って高く推定されていた可能性が考えられる。そのため、行動認識性能の低さに反して、LOW の認識信頼度が高く推定されてしまった可能性がある。これにより、正しい認識ができない場合に LOW が選択されることが増え、行動認識性能が低くなってしまったと考えられる。また、LOW を選択したあとすぐに、それより高いサンプリングレートを選択したとしても、サンプリングレート選択の遅延により、CSI の取得は高いサンプリングレートで行われ、認識は LOW で行われる。これにより、取得サンプル数の削減がされにくかったと考えられる。

また、セッション 12 においては、取得サンプル数は MIDDLE よりも大きく、マクロ平均 F 値は MIDDLE や LOW と同程度となった。セッション 12 においては、LOW と MIDDLE の行動認識性能が同程度である。一方、学習に用いたセッションに付与したノイズは、低いサンプリングレートで取得された CSI を用いた行動認識ほど、より行動認識性能を低下させることが実験的に明らかになった。そのため、学習に用いられたセッションでは LOW と MIDDLE の行動認識性能が同程度である場合が少なく、強化学習のエージェントは、そのような場合の経験が得られず、適切なサンプリングレート選択方策が学習できなかったと考えられる。

また、既存手法のマクロ平均 F 値と取得コストは、提案手法のものと比較して低くなっていることが多い。これは、既存手法では学習に用いるセッションにノイズを付与しておらず、かつサンプリングレート選択モジュールの入力として、各サンプリングレートの認識信頼度が用いられていないためであると考えられる。4.2 節で示したように、対象環境での行動認識性能は、学習環境での行動認識性能よりも低くなる。そのため、学習環境で十分な精度で認識できるサンプリングレートが、対象環境では十分ではなく、より高いサンプリングレートが適切となる場合が多い。これにより、既存手法は提案手法よりも低いサンプリングレートを選択することが多くなると考えられる。

以上のように、今回の実験では、提案手法は対象環境に

において適切なサンプリングレートを選択できる場合とそうでない場合があった。全ての場合に対応するには、サンプリングレート選択モジュールへの入力や、学習に用いるセッションへ付与するノイズにより工夫が必要であると考えられる。

謝辞 本研究の一部は科研費 JP16H06539, JP17H04679, JP21H03428, 21H05299, JP21K19769 の支援によるものです。

参考文献

- [1] Wang, W., Liu, A. X., Shahzad, M., Ling, K. and Lu, S.: Device-Free Human Activity Recognition Using Commercial WiFi Devices, *IEEE Journal on Selected Areas in Communications*, Vol. 35, No. 5, pp. 1118–1131 (2017).
- [2] Wei, B., Hu, W., Yang, M. and Chou, C. T.: From Real to Complex: Enhancing Radio-based Activity Recognition Using Complex-Valued CSI, *TOSN*, Vol. 15, No. 3, pp. 35:1–35:32 (2019).
- [3] Tanno, Y., Maekawa, T., Hara, T. et al.: Wi-Fi CSI-based Activity Recognition with Adaptive Sampling Rate Selection, *MobiQuitous '21* (2021).
- [4] Jiang, W., Miao, C., Ma, F., Yao, S., Wang, Y., Yuan, Y., Xue, H., Song, C., Ma, X., Koutsonikolas, D., Xu, W. and Su, L.: Towards Environment Independent Free Human Activity Recognition, *the 24th Annual International Conference on Mobile Computing and Networking, MobiCom 2018*, pp. 289–304 (2018).
- [5] Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M. and Lempitsky, V.: Domain-adversarial training of neural networks, *The Journal of Machine Learning Research*, Vol. 17, No. 1, pp. 2096–2030 (2016).
- [6] Sen, S., Radunovic, B., Choudhury, R. R. and Minka, T.: You are facing the Mona Lisa: spot localization using PHY layer information, *The 10th International Conference on Mobile Systems, Applications, and Services, MobiSys'12*, pp. 183–196 (2012).
- [7] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M. A.: Playing Atari with Deep Reinforcement Learning, *CoRR*, Vol. abs/1312.5602 (2013).
- [8] Schaul, T., Quan, J., Antonoglou, I. and Silver, D.: Prioritized Experience Replay, *4th International Conference on Learning Representations, ICLR 2016* (2016).
- [9] Hessel, M., Modayil, J., van Hasselt, H., Schaul, T., Ostrovski, G., Dabney, W., Horgan, D., Piot, B., Azar, M. G. and Silver, D.: Rainbow: Combining Improvements in Deep Reinforcement Learning, *CoRR*, Vol. abs/1710.02298 (online), available from <http://arxiv.org/abs/1710.02298> (2017).
- [10] Halperin, D., Hu, W., Sheth, A. and Wetherall, D.: Tool release: gathering 802.11n traces with channel state information, *Computer Communication Review*, Vol. 41, No. 1, p. 53 (2011).
- [11] Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, *3rd International Conference on Learning Representations, ICLR* (2015).