

推薦論文

スマートフォン内蔵マイクを用いた 歩道路面状況推定における汎用性の向上に向けた改善

金澤 玲実^{1,a)} 石田 繁巳² 白石 陽²

受付日 2023年6月2日, 採録日 2024年1月10日

概要: 湿潤や凍結といった不安定な路面状況によって、歩行者の転倒・転落事故が発生する。このような事故を防止するためには、路面状況を事前に把握することが重要である。本研究では、路面状況によって歩行音が異なることに着目し、スマートフォンの内蔵マイクを用いた音響センシングによって路面状況を推定する手法を提案する。本稿では、複数人の歩行音を収集し、歩行動作の個人差に頑健な推定手法を検討および評価する。評価実験の結果、1歩ごとに抽出したMFCC (Mel Frequency Cepstral Coefficients) と複数歩における歩行安定度を組み合わせて用いることで、一般分類モデルの推定精度が F-measure 0.697 となり、先行研究で用いた手法から F-measure が 0.017 向上することを確認した。

キーワード: 路面状況, スマートフォン, 音響センシング

A Proposal of a Method to Increase the Versatility of Sidewalk Surface Condition Estimation Using a Smartphone's Microphone

REMI KANAZAWA^{1,a)} SHIGEMI ISHIDA² YOH SHIRAISHI²

Received: June 2, 2023, Accepted: January 10, 2024

Abstract: Unstable sidewalk surface conditions, such as wet or icy conditions, can cause pedestrian falls. To prevent such accidents, it is important to know the road surface condition in advance. In this study, focusing on the fact that walking sounds are highly affected by the road surface condition, we propose a method to estimate the sidewalk surface condition by acoustic sensing using a smartphone's built-in microphone. In this paper, we collect walking sounds of multiple people with the same footwear and road surface conditions, and investigate and evaluate an estimation method that is robust to individual differences in walking behavior. The results of the experimental evaluations showed that the common classification model that combines the Mel Frequency Cepstral Coefficients (MFCC) extracted at each step and the gait stability at multiple steps, resulted in an F-measure of 0.697, which is 0.017 higher than the conventional method.

Keywords: road surface condition, smartphone, acoustic sensing

1. はじめに

歩道路面の凹凸や湿潤による転倒・転落など、歩道路面状況に起因する歩行者の事故が問題となっている。東京都は、降雨後の路面に対し危機感を持った経験や怪我をした

経験を持つ人が4割に上ることを示している [1]。このような歩行者事故が起こる一因として、歩行者が路面状況を事前に把握できず、降雨後の水膜やぬかるみを予期しづらく不安定な路面状況の道を通りしてしまうことが考えられる。

安全かつ快適な歩行を実現するためには、歩道路面状況を把握することが重要となる。永田らの報告 [2] では、歩行者事故が起こりやすい地点で路面状況の危険性に関する

¹ 公立はこだて未来大学大学院システム情報科学研究科
Graduate School of Systems Information Science, Future University Hakodate, Hakodate, Hokkaido 041-8655, Japan

² 公立はこだて未来大学システム情報科学部
School of Systems Information Science, Future University Hakodate, Hakodate, Hokkaido 041-8655, Japan

a) g2122018@fun.ac.jp

本稿の内容は 2022 年 10 月の第 30 回マルチメディア通信と分散処理ワークショップ (DPSWS2022) で報告され、同プログラム委員長により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である。

情報共有を行うことや、路面状況に合わせた服装を勧めることが事故予防に効果的であると考えられている。著者は路面状況の把握にあたって、多数の歩行者による協働での情報共有が効果的であると考え、多数の歩行者が各地点における路面状況を把握し、その情報を周辺エリアにいる歩行者および在宅者に向けて共有することで、路面状況を考慮した経路案内や適切な服装の推薦といった事前の情報把握を行うことが可能となる。本研究では、安全かつ快適な歩行を実現するための歩行者支援に向け、多数の歩行者による協働を想定した路面状況の把握手法を確立することを目指す。

路面状況の把握に向けて、スマートフォンやスマートシューズといったモバイルデバイスのセンサデータを用いて路面状況を推定する研究が行われている [3], [4], [5], [6], [7], [8], [9]。これまでに慣性データや圧力データ、音響データを用いた手法が提案されてきたが、これらの手法では推定可能な路面状況に限られる、デバイスの装着による手間や導入コストがかかるといった問題がある。

本研究では、スマートフォンの内蔵マイクを用いた音響センシングによって歩道路面状況を推定する。アスファルト舗装路と砂や土で覆われた未舗装路では歩行した際に生じる音が異なるように、歩行音は路面状況に応じて変化する。各路面状況における歩行音の特徴をとらえることで、路面状況を推定することが可能であると考え、音響センシングは慣性や圧力のセンシングと比較してデバイスの向きや設置箇所における制限が少ないため、デバイスの固定や装着といったデータ収集者への負担を減らすことが可能である。また、スマートフォンのような既存デバイスに内蔵されているマイクを用いることで、歩道路面状況推定を低コストで実現できる。

著者は、先行研究で各路面状況における音響特徴として MFCC (Mel Frequency Cepstral Coefficients) と歩行テンポを抽出し、それらを組み合わせた学習データを用いることで、F-measure 0.957 の精度で路面状況の推定が可能であることを確認した [10]。しかし、推定精度の評価に際して使用したデータは被験者 1 人分の低雑音環境下における音響データであり、歩行動作の個人差が推定結果に及ぼす影響については考慮できていない。また、先行研究で使用したデータ分割手法や特徴量は収集者の歩幅や歩行速度による影響を受けやすく、複数の歩行者が収集したデータに対して汎用的に用いることができない可能性がある。そこで、本稿では歩行動作の違いに対する頑健性を高める改善手法を提案する。

本稿では、歩行動作に対する汎化性能向上を主目的として、複数の被験者から収集した歩行音データをもとに個人ごとの学習データからなる固有分類モデルおよび被験者当人以上の学習データからなる一般分類モデルの構築を行

い、路面状況を推定する。路面状況推定の各モデル構築にあたって、MFCC と、歩行間隔の規則性を表す歩行安定度を組み合わせた学習データを用いる。

また、モデルの構築および推定に際して使用するデータ内に雑音が混入する状況を想定し、雑音除去手法の初期検討を行う。

本稿の構成は以下のとおりである。2 章では、歩道路面状況の推定や音響センシングに関する研究について述べる。3 章で歩行動作の個人差に対応した改善手法の詳細について述べ、4 章で改善手法の評価実験について述べる。また、5 章で雑音に対する頑健性能向上に向けた検討について述べる。最後に 6 章でまとめとする。

2. 関連研究

関連研究として、まず 2.1 節では、路面状況の推定に関する研究について述べる。次に 2.2 節では、音響センシングに関する研究について述べる。

2.1 歩道の路面状況推定に関する研究

路面状況を推定する手法としては、慣性センサを用いる手法 [3], [4], [5] や圧力センサを用いる手法 [6], [7], [8]、音響センサを用いる手法 [9] などが報告されている。

慣性データを用いた歩道路面状況推定では、段差や傾斜といった形状に関する路面状況を推定することができる。藤井ら [3] は、スマートフォンをポケットに入れた場合とカバンに入れた場合の 2 つの状況下において、それぞれ加速度データを用いて路面状況を推定している。路面の形状に応じて歩行動作が変わることに着目し、3 歩分の加速度データから平坦な道、段差、傾斜の緩やかな坂道を推定できると示している。佐藤ら [4] は、靴の甲に装着した慣性センサのデータを用いて路面状況を推定している。センサの傾きから路面の傾斜が推定できることや、複数の歩行者から得た路面傾斜の分散から凹凸の有無が推定できることを示している。桂田ら [5] は、ポケットに入れたスマートフォンの慣性センサを用いて路面状況を推定している。慣性データから抽出した特徴量を用いて個人ごとにモデルを構築することで、冬季の積雪や圧雪といった路面状況を推定できることを示している。しかし、慣性データのみを用いた路面状況推定は、歩行動作の個人差による影響を受けやすいと考えられる。また、タイル舗装路の乾燥・湿潤状態といった歩行動作への影響が少ない路面状況の推定が困難であるという課題がある。

圧力データを用いた歩道路面状況推定では、傾斜や路面の材質、状態に関する路面状況を推定することができる。大橋ら [6] は、足圧布センサを用いて路面状況を推定している。路面の凹凸に応じて足底にかかる圧力が変わることに着目し、平面と視覚障害者誘導用ブロックを推定できることを示している。Matthies ら [7] は、静電容量式の圧力セ

ンサを装着したインソール型デバイスを用いて路面状況を推定している。センサ値から足底にかかる圧力を推定し、砂路や敷石、タートンなど6種類の路面状況を推定できることを示している。また、センサ値の変動から路面の湿潤を検知することが可能であると示している。若林ら [8] は、圧力センサを装着したインソール型デバイスを用いて路面状況を推定している。路面状況に応じて足の接地動作が変わることに着目し、冬季の積雪や圧雪、凍結といった路面状況を推定できることを示している。しかし、圧力データのみを用いた路面状況推定では、歩行動作の個人差や、推定対象と近い硬度・形状の路面状況の影響により誤推定を起す可能性がある。また、デバイスを履物に装着する手間やデバイスの導入コストがかかる。

音響データを用いた歩道路面状況推定では、路面の材質や状態に関する路面状況を推定することができる。三嶽ら [9] は、靴に装着した慣性センサとマイクを用いて路面状況を推定している。1歩ごとに抽出した慣性データの統計量と歩行音のMFCCを用いて、砂利道や草むら、水たまりなど6種類の路面状況を推定できることを示している。また、データ収集時の背景雑音の大きさに応じて慣性データの寄与度を高めた推定を行うことで、雑音環境下における推定精度の低下を抑えることができると示している。しかし、慣性データの寄与度が高まることで路面の材質や状態についての推定が困難になると考える。また、マイクを搭載したスマートシューズの使用を想定しており、デバイスの特殊性から地域住民によるデータ収集への協力を想定した場合に協力者の確保が困難であると考えられる。

2.2 音響センシングに関する研究

本研究で目的音とする歩行音は、踵やつま先が地面と衝突する際の衝撃音と、足底と地面が擦れる際の摩擦音からなる [11]。こうした環境音における特性を分析し、推定に用いる研究が報告されている [12], [13]。堀ら [12] は、床に固定したマイクを用いて1歩分の歩行音から人物を推定している。歩行音から抽出したメル周波数スペクトログラムを用いたSVM (Support Vector Machine) や、水増しした歩行音を用いたCNN (Convolutional Neural Network) など、複数の推定手法について調査し、人物推定が可能であることを示している。荒川ら [13] は、家庭内における生活音を用いて日常生活動作を推定している。荒川らの手法では、音響特徴量としてMFCCを用いている。推定の対象となる生活音に合わせてMFCCの次元数、周波数範囲を拡大し、主成分分析で重要な特徴のみを抽出する手法を用いることで、生活音からの行動推定が可能であることを示している。このように、環境音を対象とした研究では特徴量としてメル尺度を適用した周波数情報が有効であると報告されている。本研究においても、歩行音の周波数情報をとらえるにあたってMFCCを利用する。

3. 歩行者への汎化性能向上に向けた改善

本章では、著者らの先行研究 [10] での提案手法（以降、先行手法）を複数の歩行者に適用することを考慮した改善手法について述べる。まず3.1節で改善手法の概要について述べ、以降の3.2～3.5節で改善手法における具体的な処理について述べる。

3.1 手法概要

本稿で用いる改善手法の全体像を図1に示す。

改善手法はデータの準備フェーズと推定フェーズから構成される。準備フェーズでは、歩行者から収集した音響データ内から歩行音データを切り出し、雑音除去処理を施す。推定フェーズでは、入力された歩行音データから特徴量抽出や機械学習モデルの構築処理を行い、路面状況推定の結果を出力する。以降では、フェーズ内の各処理について述べる。

3.2 データ分割

データ分割処理部では、入力された音響データから1歩ずつ歩行音データを切り出す。先行手法では、1データの長さを約3歩分の歩行音が含まれる2秒とし、これを1件の歩行音データとして扱っていた。しかし、歩行者の癖や歩行速度などにより1データ内の歩行音が減少し、歩行間隔の変化を正確に抽出できない可能性がある。また、音響データ内に含まれる歩行音以外の音への考慮を行っておらず、足の接地から次の接地までの遊脚期に雑音のみの区間が生じていた。それにより、データ収集地点における雑音が各路面歩行時の特徴として誤認識され、推定結果に影響していたと考える。改善手法では、歩行動作の変化をより正確にとらえたうえで遊脚期の雑音を可能な限り除去するため、1データの長さを踵が接地した際のピークを中心とした0.3秒分に変更する。なお、改善手法では、この1データを1歩分の音響データとして扱う。

3.3 雑音除去

本稿では、ステレオ形式で収録した歩行音データをモノ

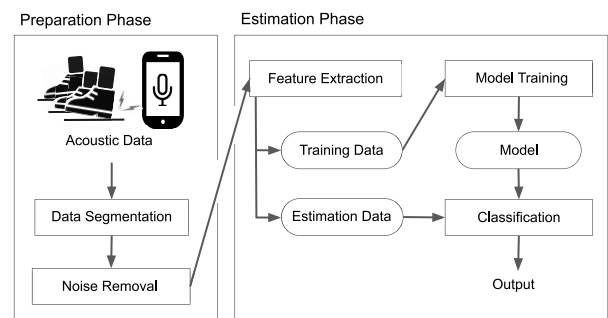


図1 推定のプロセス

Fig. 1 Estimation process.

ラル形式に変換することで雑音による影響を軽減する。なお、本稿の雑音処理部では目的音である歩行音以外の音を雑音と定義している。通常、スマートフォンは端末上部と下部に内蔵マイクを有している。これらのマイクではデバイスの携行状況によってそれぞれ異なる音データが収集できる。たとえば、端末の上下部いずれかが布地と接触している場合では片側のマイクチャンネルで定常的に大きな振幅値を持つ雑音が収録される。2つのマイクによるステレオ形式で収録された音響データのうち、振幅値の大きなチャンネルを除去することで雑音による処理への影響を軽減できると考える。本稿では、ステレオ形式データのうち、平均振幅値が小さいチャンネルのデータを抽出することで、より雑音が少ないデータに変換する。

3.4 特徴量抽出

特徴量の抽出処理部では、改善手法における特徴量としてMFCCと歩行安定度を抽出する。以降では、先行手法で利用したMFCCおよび歩行テンポと、改善手法で導入する歩行安定度の概要について述べる。

(1) MFCC

MFCCは、メルフィルタバンクを適用したケプストラム係数で、スペクトルの概形を表す20次元の特徴量である。算出にあたっては、まず音響データに対してフーリエ変換を行い、スペクトルを求める。次に、得られたスペクトルに対してメルフィルタバンクを適用し、得られたメルスペクトルを対数へと変換する。変換した対数スペクトルに対し離散コサイン変換を行うことでメル周波数ケプストラムが得られる。これは低周波数成分を重視したスペクトルを時間尺度に戻したケフレンシ領域の数値列であり、低次成分がスペクトル包絡、高次成分がスペクトル微細構造を表す。本稿では、フィルタバンクのチャンネル数を128、特徴量として使用する係数を低次から20としてMFCCを抽出する。なお、MFCCの0次の係数は対数スペクトルの直流成分を表すため除外する。

(2) 歩行テンポ

歩行テンポは、1分あたりの歩数を表す1次元の特徴量である。算出にあたっては、音響データ内にほぼ一定間隔で出現する歩行音を拍と見なし、Ellis [14]の拍検出手法を適用する。具体的には、音響データ開始時の振幅値を測定し、その振幅値から相対的に近い振幅値を持つピークを動的計画法によって検出していく。その後、先行手法で分割した2秒間の音響データ内におけるピークの出現数を1分あたりの出現数に直すことで、歩行テンポを求める。

(3) 歩行安定度

歩行安定度は、歩行間隔の規則性を表す1次元の特徴量である。齊田ら [15]は、路面状況の悪化にともない

歩行の規則性が失われる点に着目し、歩行時の加速度から加速度スペクトル密度を算出し路面の転倒危険度を定量的に評価する手法を提案している。また、路面状況の悪化にともなう歩行危険度の変動量は、年齢や性別といった歩行者の属性によって異なることや、路面状況ごとの転倒危険度の出現頻度の傾向から悪路の検出が可能であることを報告している。本稿では、歩行間隔 i を音響データ内の $k-1$ 番目と k 番目の歩行開始時刻の差分とし、歩行間隔 $i-2$ から歩行間隔 i における標準偏差を歩行安定度とする。算出にあたっては藤井ら [3]の報告に基づき、3歩分のデータを使用することとした。

著者らの先行研究では、MFCCと歩行テンポを組み合わせた学習データを用いることで、路面状況推定における推定精度がMFCC単体の場合よりも向上することを確認した。しかし、歩行テンポは2秒間の音響データに含まれる歩行音の頻度を1分あたりへと変換しているため、わずかな歩行の乱れが大きく値に影響する。また、歩行テンポは歩幅や歩行速度などにおける個人差の影響を受けるため、複数の歩行者に適用した場合には歩行動作の個人差から汎用的に用いることができない可能性がある。それに対して歩行安定度は、ある1歩とその前2歩分の歩行間隔との相対的なずれを算出しているため誤差が生じにくく、歩行者の個人差による影響も受けにくいと考える。そのため、本稿では改善手法としてMFCCと歩行安定度を組み合わせる学習データを作成する。

3.5 モデル構築

モデル構築部では、作成した学習データを標準化した後に分類モデルの学習を行う。モデル構築にあたっての機械学習アルゴリズムにはSVMを用いる。SVMは著者らの先行研究で6種類の路面状況をF-measure 0.957の精度で推定することが可能であったため、改善手法においても採用した。SVMのパラメータはグリッドサーチで決定する。

4. 歩行動作の個人差に関する汎化性能評価

本章では、改善手法を複数人の歩行音データに適用した際の汎化性能を評価した実験について述べる。路面状況推定システムの実用に際しては、歩行者ごとでの処理を必要とせず、複数の歩行者に汎用的に適用できることが望ましい。そこで、予備実験として先行手法における歩行テンポと改善手法における歩行安定度を被験者別に比較分析した。また、改善手法を用いて固有分類モデルと一般分類モデルをそれぞれ構築し、被験者の違いによる影響を調査した。

以降では、4.1節で実験環境について述べ、4.2節で歩行テンポおよび歩行安定度の汎用性に関する比較分析結果について述べる。



(a) 舗装路/乾燥 (b) 舗装路/湿潤 (c) 草地/乾燥

図 2 対象路面

Fig. 2 Target road surface.

表 1 被験者データ一覧

Table 1 Subject data list.

被験者	実験グループ	履物	性別	年代	雑音レベル
A	1	共通	女	20	.0045
B	2	共通	男	20	.0009
C	3	共通	女	20	.0015
D	3	共通	男	20	.0010
E (1 回目)	3	共通	男	20	.0012
F	4	共通	男	20	.0008
G	5	個人	男	20	.0025
E (2 回目)	5	個人	男	20	.0039
H	6	共通	女	20	.0006
I	6	共通	男	20	.0005
J	6	個人	男	20	.0007
K	7	個人	女	20	.0097
L	7	個人	男	20	.0086
M	8	共通	男	20	.0018
N	8	個人	男	20	.0011
O	8	個人	男	20	.0018

4.1 実験環境

本稿での実験に用いる音響データの収集は、図 2 に示す 3 種類の路面状況で行った。状態差による差異をみるため同地点の舗装路における乾湿の路面状況の一例として a, b を選択した。また、材質差による差異をみるため舗装路と路面の材質が異なる路面状況の一例として c を選択した。先行研究 [10] では、歩きづらく舗装路と材質の異なる路面状況として未舗装路（土道）を用いたが、路面表面の凹凸による不安定さが類似している点や、伸びた草により歩きづらさが増す点などから草地に変更した。

データ収集者となる被験者は男性 11 名、女性 4 名の計 15 名である。データ収集は 6 日間をかけて行い、データ収集地点はすべての被験者で同じである。各被験者のデータ収集時における情報を表 1 に示す。表中の実験グループが同一の被験者らは同時刻にデータ収集を行ったことを表す。履物は実験者が用意したものを使用した場合に「共通」、被験者本人のものを使用した場合に「個人」としている。雑音レベルは関連研究 [9] の処理を参考に、被験者ごとの歩行音における振幅値の平均分散値を求めたものである。この雑音レベルは、雑音の混入によって振幅値が急激に変化した際に高い分散値を示しやすい。被験者は、ゴムベルトを用いて衣服の上からスマートフォンを腰の位置に固定した状態（図 3）で約 5m の区間を往復し、歩行デー

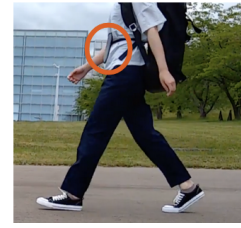


図 3 データ収集の様子

Fig. 3 During data collection.

表 2 検証する推定手法

Table 2 Estimation methods for comparison.

手法名	データ分割方法	特徴量
先行手法	2 秒ごと	MFCC・歩行テンポ
改善手法 1	1 歩ごと	MFCC
改善手法 2	1 歩ごと	MFCC・歩行安定度

タを収集した。実験に際して、被験者にはデータ収集中に立ち止まらないことや発声しないことを指示した。実験時のマイクの録音形式はステレオで、サンプリング周波数は 44.1 kHz である。

収集したデータは、表 2 に示す 3 つの手法でそれぞれ処理し、学習データを作成した。使用したモデルのパラメータは、カーネルが RBF, C (コストパラメータ) が 10, gamma (カーネル係数) が 0.01 である。なお、著者らの知る限りでは歩行者の所持するスマートフォンのマイクによるデータ収集を想定した路面状況推定手法は報告されていないため、改善手法の比較対象として先行手法を取りあげる。

4.2 歩行テンポ・歩行安定度の比較分析

予備実験として、女性 2 名（被験者 A, C）と男性 3 名（被験者 B, D, E）の計 5 名分のデータを用いて歩行テンポと歩行安定度の比較分析を行った。なお、使用する被験者のデータ選出にあたっては、同日、各日でのデータを最低 3 人分ずつ含むこと、履物による音響への影響を省くため同一の履物を使用していることなどを考慮した。分析にあたって歩行音の特徴を正確にとらえるため、データ内の非歩行区間および歩行音がかき消された高雑音区間については収集後に提案手法と別途に除去した。除去にあたっては非歩行区間および高雑音とラベル付けした区間のデータを空白値とし、以降の処理で用いないものとした。

図 4 は、ローレンツプロット法を適用して歩行間隔 i を x 軸、歩行間隔 $i+1$ を y 軸とし、歩行テンポおよび歩行安定度について時系列上で隣接するデータ間の変動をプロットしたグラフである。また、表 3 は、歩行テンポおよび歩行安定度のばらつきを表す、ローレンツプロットの面積 S をまとめたものである。グラフ上で $y = x$ の場合または S が小さい場合には、歩行間隔の乱れがない安定した歩行時のデータであることを意味する。

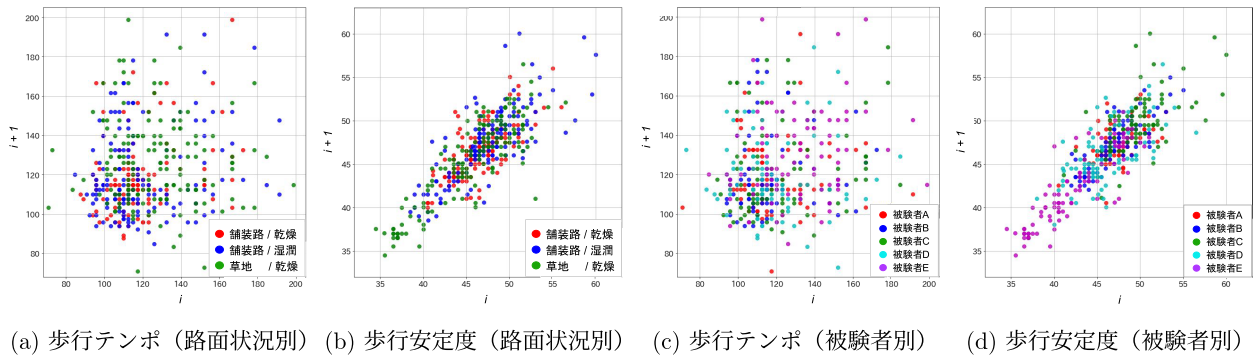


図 4 歩行テンポと歩行安定度のローレンツプロット

Fig. 4 Lorenz plot of gait tempo and stability.

表 3 特徴量別のローレンツプロット面積 S

Table 3 Lorenz plot area S by feature.

	被験者別					路面状況別		
	被験者 A	被験者 B	被験者 C	被験者 D	被験者 E	舗装路/乾燥	舗装路/湿潤	草地
歩行テンポ	271.5	395.6	235.6	318.3	438.3	266.2	370.0	378.3
歩行安定度	3.6	7.1	2.9	7.0	6.9	4.5	7.9	9.4

表 4 路面状況分類モデルの推定精度

Table 4 F-measure of road surface estimation.

	被験者 A		被験者 B		被験者 C		被験者 D		被験者 E		平均		p 値 (<0.05)
	固有	一般	固有	一般	固有	一般	固有	一般	固有	一般	固有	一般	
先行手法	.920	.591	.981	.654	.765	.722	.732	.737	.704	.695	.820	.680	-
改善手法 1	.827	.607	.947	.633	.749	.725	.832	.766	.763	.636	.824	.673	0.05
改善手法 2	.882	.636	.941	.658	.760	.757	.862	.812	.813	.625	.852	.697	0.04

路面状況別の結果より、両特徴量ともに S が草地/乾燥、舗装路/湿潤、舗装路/乾燥の順で大きいという結果が得られた。これは、歩行間隔が前述の順で不安定であることを意味し、歩きやすさの異なる路面状況を推定するにあたっての特徴量として有用であると考えられる。また、この結果は歩行間隔に関する特徴量が舗装路や積雪路の状態差における誤推定を減少させ、推定精度を向上させるという先行研究 [10] での結果と類似しており、歩行テンポや歩行安定度に関する有用性を裏付けられた。

しかし、被験者別の結果より、被験者ごとにプロットの分布や S の大きさが異なることから、歩行間隔の規則性における個人差が現れていることが分かる。特に、被験者別の歩行安定度に注目すると、女性（被験者 A, C）は歩行安定度の S が 3.0 前後、男性（被験者 B, D, E）は歩行安定度の S が 7.0 前後になる傾向が見受けられる。被験者数の少なから歩行者の属性や歩行動作の個人差によるものであるか明らかでないが、汎用性の向上にあたって歩行者の属性や歩行動作特性における類似度を考慮したデータの校正手法が必要である可能性が示唆された。

4.3 固有分類モデル・一般分類モデルの評価

先行手法、改善手法 1、改善手法 2 をそれぞれ用いて固

有分類モデルおよび一般分類モデルの推定精度を評価した結果を表 4 に示す。また、改善手法の比較対象として、筆者らの先行研究 [10] を取り上げる。なお、その他の関連研究については筆者らの調べた限りでは同様のデバイス、収集状況、収集対象を想定した研究が報告されていないため今回の比較対象として取り上げない。また、一般分類モデルで各手法を評価した際の混同行列を図 5 に示す。

推定精度は、固有分類モデルに 5 分割交差検証、一般分類モデルに LOPOCV (Leave One Person Out Cross Validation) を用いてそれぞれ評価した。なお、両分類モデルとも使用したデータは 4.2 節と同一であり、路面状況別に学習データが 120 件、テストデータが 30 件となるようデータを分割した。また、一般分類モデルでは上記の学習データと別途に各被験者で 10 件ずつランダムにデータセットを作成し、得られた推定精度をマン・ホイットニーの U 検定にかけ、片側での有意差を確認した。なお、表 4 中の p 値は先行研究に対して有意に大きい推定精度を得られたかを表す。推定精度の評価指標として F-measure をマクロ平均で算出した。

4.3.1 歩行安定度の有用性に関する分析

実験結果より、複数の被験者を対象とした推定においても MFCC と歩行間隔に関する特徴量を組み合わせて用い

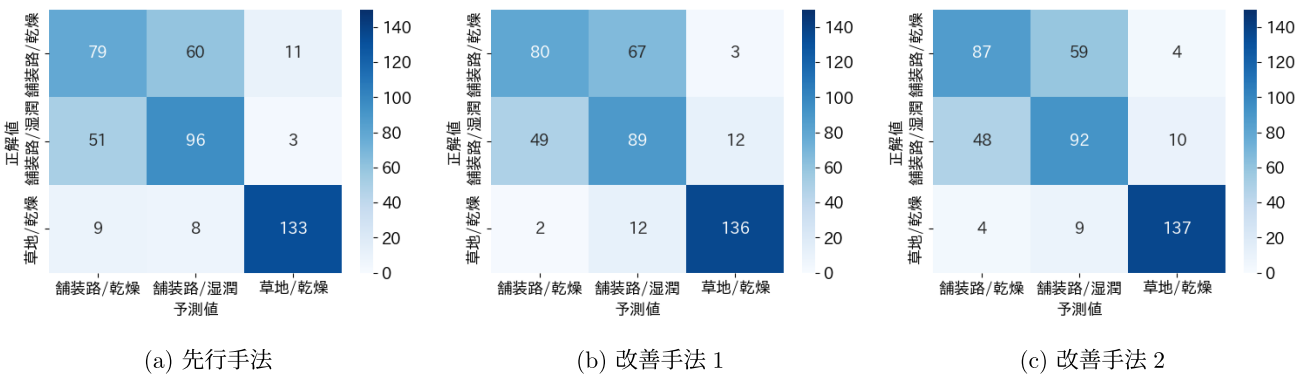


図 5 一般分類モデルの混同行列

Fig. 5 Confusion matrix for common classification model.

表 5 収集日時別の推定精度

Table 5 F-measure by collection date and time.

収集日時	被験者					平均	p 値 (<0.05)
	被験者 A	被験者 B	被験者 C	被験者 D	被験者 E		
同日時	-	-	.589	.597	.530	.572	-
別日時	.723	.590	.660	-	-	.658	.002

る手法が有用であることを確認できた。改善手法 1 と改善手法 2 の間で被験者によって推定精度に差があったものの、平均では MFCC と歩行安定度を組み合わせた改善手法 2 の方が高い推定精度を得られた。有意差検定においても、改善手法 2 は先行手法、改善手法 1 のそれぞれよりも有意に高い推定精度が得られることが確認できた。これは先行研究と同様の結果であり、MFCC に歩行間隔に関する特徴量を組み合わせると推定精度が向上する点は被験者数が増えても変わらなかった。

固有分類モデル、一般分類モデルともに改善手法 2 の平均推定精度で最も高い結果が得られた。これは路面状況によって変動する歩行の規則性を特徴としてとらえられたためであると考えられる。図 5 の混同行列より、改善手法 1 と改善手法 2 を比較して舗装路の乾燥、湿潤状態間における誤推定が減少したことからも周波数情報だけではとらえられない歩行動作の差があったと考えられる。また、先行手法より改善手法 2 を用いた方が高い推定精度を得られたことから、歩行テンポよりも歩行安定度の方が複数人を対象とした特徴量として適していると考えられる。

しかしながら、すべての推定手法において、一般分類モデルの平均推定精度が固有分類モデルよりも 2 割程度低下した。本稿で考案した改善手法 2 における一般分類モデルの平均推定精度は F-measure 0.697 であった。これは、テストデータとなる被験者と異なる歩行者属性や歩行動作特性を持つ被験者の学習データを用いて分類モデルの学習を行ったことで、推定に対する歩行テンポや歩行安定度の影響度が低下したためと考える。今後は歩行者属性や歩行動作特性の類似度を考慮した改善が必要である。

4.3.2 周囲雑音の影響に関する分析

実験結果より、データ収集日時の違いによる推定結果への影響が考えられる。表 4 より、被験者ごとの推定精度に着目すると、被験者 A, B は固有分類モデルの推定精度が各推定手法で 8 割以上と高い結果が得られた。特に先行手法を用いた際の推定精度が 9 割以上と高い結果になった。この要因として、遊脚期に含まれる雑音を各路面歩行時の特徴として誤認識した可能性が考えられる。各路面歩行時の周囲雑音が顕著に異なったことで、固有分類モデルで高い推定精度が得られたと考える。このことから、データ収集日時の違いによって周囲雑音の状況が変化し、推定結果に影響を及ぼすことが考えられる。

データ収集日時の違いによる推定精度への影響を調査するため、日時別の被験者データを改善手法 2 に適用し、一般分類モデルを構築することで追加評価を行った。その評価結果を表 5 に示す。表 5 は各被験者で 10 件ずつランダムにデータセットを作成し、収集日時別のグループによる LOPOCV で推定精度を評価したものである。平均値は被験者 3 名の計 30 試行における推定精度から求めている。

異なる日時にデータを収集した被験者 A, B, C の音響データを用いた場合の推定精度は F-measure 0.658 であった。また、同日時にデータを収集した被験者 C, D, E の音響データを用いた場合の推定精度は F-measure 0.572 であった。同日時に収集したグループについては別日時のグループと比較して平均推定精度が 1 割ほど減少している。有意差検定においても別日時のグループによる推定精度の方が有意に高いという結果が得られた。これは、別日時でのデータ収集によって背景雑音に大きな差が生じたためと考えられる。このことから、歩行音と重ならない遊脚期の

雑音だけでなく、歩行音と重なるように生じる雑音を除去する必要がある。

5. 雑音に対する頑健性能向上に向けた検討

本章では4章の結果を受け、提案手法の雑音に対する頑健性を向上させるために行った検討について述べる。実環境において歩行者がスマートフォンを所持してデータ収集を行う際には、衣類やカバン類などによる衣擦れ音が歩行音と連動して生じると考えられる。また、その際の衣擦れ音は足よりもマイクに近い位置で生じるため歩行音よりも振幅値が大きくなる傾向にあると考える。除去対象となる音の振幅値が歩行音よりも大きい場合には、歩行音の音素のみを取り出すことが難しい。そこで雑音が影響しないデータの併用および、歩行音と重なる雑音の除去手法について検討する。

5.1節ではデータ分割および特徴量抽出に加速度データを用いる改善手法について、5.2節では雑音除去処理として音源分類を行う改善手法について述べる。最後に、5.3節で上記の改善手法を用いた場合の路面状況推定精度を評価する。

5.1 歩行検知方法の改善

雑音への頑健性を高めるため、データ分割処理および歩行動作に関する特徴量抽出に加速度データを用いる。改善手法1, 2では歩行音の分割や特徴量の抽出にあたって、ピーク音に基づく歩行音の検知を行っていた。しかし、雑音が多量に含まれる音響データの場合、歩行音と雑音のピーク音を混同し、歩行音が含まれない区間を切り出してしまう可能性がある。同様に、歩行テンポや歩行安定度といった歩行動作に関する特徴量の抽出においても、雑音の影響により実際の歩行動作との乖離が発生する可能性がある。そこで、加速度データを用いる改善手法3を新たに設ける。改善手法3では改善手法2におけるデータ分割および特徴量抽出処理で用いていたピーク音検知を加速度の3軸合成値に基づくピーク検知に置き換える。加速度の3軸合成値を用いることで、センシングデバイスの位置や向きの制限を受けずに接地時の身体の揺れを検知できると考える。他手法との比較評価については、5.3節で後述する。

5.2 雑音除去処理の改善

同じく、雑音への頑健性を高めるため、雑音除去に音源分類モデルを用いる。改善手法1, 2では、3.2節のデータ分割処理や3.3節のモノラル化処理によって、雑音による影響を抑えていた。しかし、上述の雑音除去処理では振幅値の大きな雑音が歩行音と同時に発生している場合やステレオマイクの両チャンネルに多量の雑音が含まれる場合などにおける除去性能が不十分である。そこで、3.3節の雑音除去部におけるモノラル化処理後に音源分類による雑音除

去処理を追加する。

音源分類による雑音除去処理では、ピーク検知後に切り出された音響データを入力値とし、その音響データが歩行音であるか否かを判定する。モノラル音源に対する雑音除去手法として、一般にスペクトルサブトラクション法や非負値行列因子分解などを用いた周波数成分の分離手法が知られている。しかし、歩行音を含め実環境下で収集した環境音の多くは低周波で構成されているため、これらの手法では歩行音の特徴を誤って除去する可能性がある。そこで、本稿では雑音が多量に含まれるデータ区間を除去することで雑音除去を行う。音源分類モデルを用いて歩行音と雑音を分類し、歩行音であると分類されたデータ区間のみを路面状況推定に用いることで、路面状況推定における雑音の影響を軽減する。以降では、音源分類モデルの構築方法について述べる。

音源分類モデルの構築にあたっては、路面状況の分類モデルと同様にSVMを用いる。分類モデルの学習にあたっては、オープンデータセットFSD50K [16]に収録された歩行音、風切り音、車両走行音に関する音響データ計6,278件と、著者らが収集した歩行音、衣擦れ音に関する音響データ計6,021件を使用する。分類に用いる特徴量としてMFCCを3章と同様の手順で抽出し、学習データとする。

5.3 改善手法の評価

5.3.1 データ多様化と雑音混入による影響とその対応

5.1, 5.2節の各手法を用いた場合の路面状況推定精度の違いを表6に示す。学習および検証にあたっては4.1節で収集した15名分のデータを使用した。また、雑音除去処理の有効性を確認するため、雑音によって歩行音がかき消されている高雑音区間についても各路面状況のラベル付けを行ったうえで混入させた。また、データの多様性による影響をはかるため、被験者データが5名分の場合と15名分の場合でそれぞれ検証を行った。なお、5名分のデータを用いた検証については4.2, 4.3節と同様の被験者データを用いている。推定精度はLOPOCVを用いて評価し、評価指標としてF-measureをマクロ平均で算出した。表中のp値は音源分類ありにおける15名分の被験者によるLOPOCV結果をマン・ホイットニーのU検定にかけたものである。各値は先行手法に対して各手法の推定精度が有意に高かつ

表6 改善手法別の路面状況推定精度

Table 6 F-measure of road surface estimation by improvement methods.

	音源分類なし		音源分類あり		p 値 (<0.05)
	5名	15名	5名	15名	
先行手法	.484	.454	.378	.373	-
改善手法1	.334	.326	.326	.323	1.000
改善手法2	.449	.337	.463	.350	0.999
改善手法3	.396	.391	.372	.376	0.992

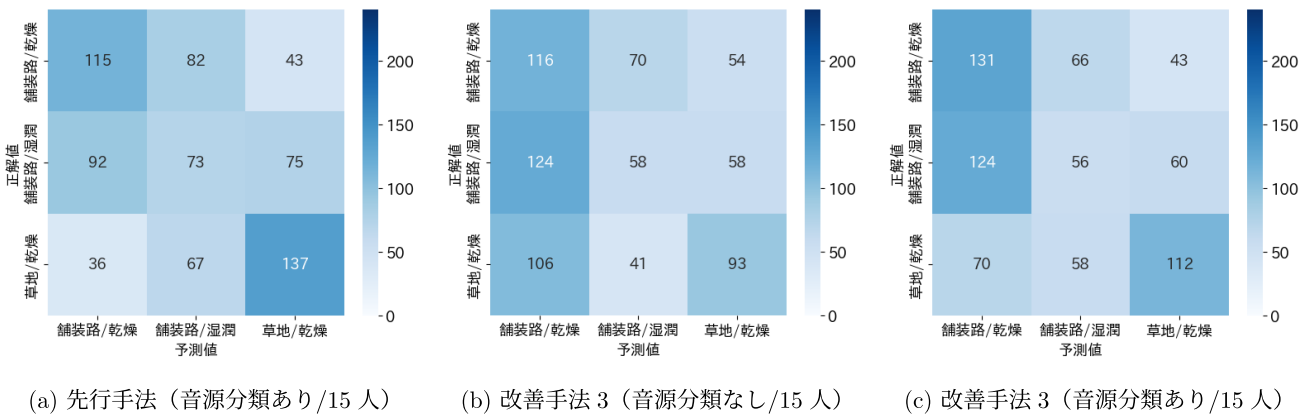


図 6 音源分類有無での混同行列
 Fig. 6 Confusion matrix with and without noise removal.

たかを表す。また、先行手法における「音源分類あり/15人」と改善手法3における「音源分類なし/15人」「音源分類あり/15人」との混同行列を図6に示す。図6では雑音除去による学習データ量の差を考慮せず、分類クラス間の傾向を確認するためにそれぞれの混同行列を求めるとあって各路面状況のデータ数を240に統一した。なお、雑音除去後に利用可能なデータ数において被験者による偏りが生じたため混同行列に使用するデータは乱数で選出した。

まず、使用する被験者データを5名分から15名分に拡張したことによる差に着目する。表6より、先行手法および改善手法1, 2において雑音除去処理の有無によらず、データの多様化によって推定精度が低下していることが分かる。一方で、改善手法3では音源分類ありの場合に被験者データを増やしたことでF-measure 0.372~0.376 (+0.004)に推定精度が向上していることが分かる。また、音源分類なしの場合もF-measure 0.396~0.391 (-0.005)と、他手法と比較して被験者データの多様化による精度低下を最も抑えられていることが分かる。この一因として、被験者データ数の増加によって収集されるデータに含まれる音源の種類が多様化したことや、改善手法3が雑音による外れ値を生みにくいことなどが考えられる。改善手法3は他手法とは異なり、データ分割や特徴量抽出にピーク音検知を使用していない。そのため、音響データに多種多様な雑音が入った場合にも雑音の影響を抑えることができたと考える。以上の結果から、改善手法3がデータの多様化、特にデータに含まれる雑音の多様化に対して頑健性を有していることが示唆された。

次に、音源分類処理の有無による推定精度への差に着目する。表6より、先行手法および改善手法1, 改善手法3において音源分類による雑音除去を行った場合の推定精度が、行わなかった場合よりも低いことが分かる。また、図6(b), (c)より、路面状況ごとの推定精度の偏りがそれぞれの手法で類似していることから本稿の雑音除去処理では路面状況推定における雑音の影響を軽減する効果が見ら

れないと考えられる。このことから、雑音除去処理を行った明瞭な歩行音のみを路面状況推定に用いるよりも、不明瞭なデータをそのまま入力値として用いた方がより高い推定精度を得られると考える。このような結果となった要因として、音源分類モデルの学習データが適していなかったことがあげられる。本稿では、音源分類モデルの学習データセット内にオープンデータからの歩行音および雑音を含んでいた。これは限定的な環境でのみ有効なモデルになることを避ける意図があったが、結果として本研究の実験で収集されたデータに適さないデータセットとなっていたと考える。実際に、音源分類によって除去されたデータの内訳を確認すると、実験グループ5や7といった雑音レベルの高いデータにおいて雑音除去が機能していたが、同時に草地のデータが雑音であると誤分類されていた。これは、オープンデータ内の歩行音について街中の舗装路を対象としたものが多かったためと考える。今回の実験では、雑音除去の改善手法における有意差を確認できなかったが、路面状況推定および音源分類モデルにおけるデータセットを変更することで推定精度が向上する可能性がある。

5.3.2 雑音以外による影響

本稿では、4章では歩行動作以外による影響を考慮しないため、被験者の履物を統一した。しかし、実用に際しては靴底の異なる履物によるデータが多数収集される。その際の影響をはかるため、15名分のデータを履物によってグループ分けして推定結果への影響を確認する。表7は被験者15名を実験者が用意した履物を使用した「共通」と被験者個人のものを使用した「個人」に分けて改善手法3における推定精度を比較したものである。「共通」および「個人」における履物はすべてスニーカータイプであった。平均はそれぞれのグループにおける試行の平均推定精度で評価指標はF-measureのマクロ平均である。p値はマン・ホイットニーのU検定における片側検定の値を表す。平均およびp値の算出にあたっては各グループの人数比による影響を低減するため、nC3の組合せによる各試行の推定精度

表 7 履物別の推定精度

Table 7 F-measure of road surface estimation by shoes.

履物	平均	p 値 (<0.05)
共通	.280	-
個人	.309	.009

を用いた。なお、表中の p 値は「個人」が「共通」に対し有意に大きいかを表す。

表 7 の結果より、p 値 0.009 で別の履物を用いた場合の方が推定精度が有意に大きくなる事が確認できた。これは、履物の材質などによって生じる歩行音が変わるであろうという予想に反する結果であった。この一因として、マイクと歩行音の距離が離れていることで詳細な音の差異が取得しづらかったことや、収集当日の周囲環境による影響が考えられる。しかし、本稿の実験ではヒールのように靴底の形状や材質がスニーカーと異なり、歩行音に大きな変化が生じると考えられる履物データについては未検証である。また、本稿の実験では被験者属性に偏りがあるうえに、年齢差や身長差といった被験者属性を考慮していない。そのため、今後は被験者の属性やデータ収集状況をさらに拡張したうえで路面状況推定への影響を調査する必要がある。

6. おわりに

本研究の目的は、安全かつ快適な歩行に向けた支援のため、スマートフォンを用いた音響センシングによって歩道路面状況を推定することである。著者らの先行研究 [10] では、推定に有用な特徴量を検討したものの、歩行動作の個人差による影響について十分に調査できていなかった。本稿では、歩行動作の個人差に頑健な推定手法を検討すべく、1 歩ごとに抽出した MFCC と複数歩における歩行安定度を特徴量として用いる改善手法を提案した。評価実験より、上述の改善手法を用いて構築した一般分類モデルで F-measure 0.697 の推定精度が得られることを確認した。また、評価実験の結果から改善手法において雑音による影響が見られたため、加速度の 3 軸合成値を用いた歩行動作検知や、音源分類による雑音区間の除去といった改善手法の検証を追加で行った。しかし、こちらの検証では改善手法が被験者数の増加に対する頑健性を有していることが示唆されたものの、先行研究の手法から雑音に対する頑健性能が向上したかについては確認できなかった。

今後は、高雑音環境下においても低雑音時と同等の推定精度が得られるよう雑音除去用のモデルにおけるアルゴリズムや学習データセットの見直しを行う。屋外環境下において無数に存在する雑音のデータを収集し、モデルを構築することは難しいため、半教師あり学習による歩行音のみの学習について検討する。また、本稿では 3 つの路面状況

に焦点を当てたが、今後は実験対象となる路面状況を実環境に合わせて拡充したうえでの手法評価が必要となる。分析対象とする歩行者属性を増やしたうえで、歩行者属性や歩行動作特性の類似度を考慮した手法改善に取り組む。

謝辞 本稿で示した研究の一部は、JSPS 科研費 JP20K11793 の助成を受けたものである。

参考文献

- [1] 東京都生活文化局消費生活部：平成 25 年度降雨時のヒヤリ・ハット調査、東京都（オンライン）、入手先 https://www.shouhiseikatu.metro.tokyo.jp/anzen/-hiyarihat/documents/report_rainy.pdf（参照 2021-07-18）。
- [2] 永田泰浩、金田安弘：令和元年度冬期の札幌市における転倒による救急搬送者の状況、寒地技術論文・報告集寒地技術シンポジウム/北海道開発技術センター編，Vol.36, pp.169–172 (2020)。
- [3] 藤井海斗、羽田野真由美、西田京介、戸田浩之、澤田 宏、鹿島久嗣：歩行者クラウドセンシングによる路面状態の推定、第 8 回データ工学と情報マネジメントに関するフォーラム，pp.1–6 (2016)。
- [4] 佐藤 匠、廣森聡仁、山口弘純、東野輝夫：靴型ウェアラブルセンサによる通行路状況推定手法の提案、情報処理学会研究報告，Vol.2015-CDS-12, No.9, pp.1–8 (2015)。
- [5] 桂田 連、長谷川達人：歩行者のための積雪マップ実現に向けたスマートフォンを用いた路面積雪量推定、マルチメディア、分散、協調とモバイルシンポジウム 2018 論文集，Vol.2018, pp.88–93 (2018)。
- [6] 大橋勇介、榎堀 優、間瀬健二：歩行時の最大圧選択足圧データを用いた路面コンテクストの識別、情報処理学会研究報告，Vol.2014-HCI-44, No.17, pp.1–8 (2014)。
- [7] Matthies, D.J.C., Roumen, T., Kuijper, A. and Urban, B.: CapSoles: Who is Walking on What Kind of Floor?, *Proc. 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '17*, pp.9:1–9:14, Association for Computing Machinery (2017)。
- [8] 若林勇汰、白石 陽：インソール型デバイスによる路面状況推定の精度向上に向けた特徴量の検討、情報処理学会研究報告，Vol.2021-DPS-189, No.15, pp.1–8 (2021)。
- [9] 三嶽寛人、渡邊拓貴、杉本雅則：足音と慣性データに基づく路面状況認識手法、情報処理学会論文誌，Vol.61, No.10, pp.1578–1590 (2020)。
- [10] 金澤玲実、石田繁巳、白石 陽：スマートフォンを用いた音響センシングによる歩道路面状況推定手法の提案、情報処理学会第 84 回全国大会，pp.1–2 (2022)。
- [11] 塩田陽也、板井陽俊：時間周波数解析と畳み込みニューラルネットワークを用いた足音の個人識別と特徴解析、情報処理学会研究報告，Vol.2019-HCI-182, No.6, pp.1–6 (2019)。
- [12] 堀 佑貴、安藤崇央、福田 晃：一步分足音を用いた個人識別手法、情報処理学会研究報告，Vol.2019-AVM-107, No.10, pp.1–6 (2019)。
- [13] 荒川正規、大石優彦、撫中達司：生活音を用いた行動認識のための特徴量抽出についての一考察、情報処理学会研究報告，Vol.2019-CDS-24, No.30, pp.1–8 (2019)。
- [14] Ellis, D.P.W.: Beat Tracking by Dynamic Programming, *Journal of New Music Research*, Vol.36, No.1, pp.51–60 (2007)。
- [15] 齊田 光、徳永ロベルト、高橋尚人、渡部武朗、高野伸栄：スマートフォンを用いた冬期転倒危険度の定量評価手法に関する研究、土木学会論文集 D3 (土木計画学)，Vol.75,

No.5, pp.L999–L1008 (2019).

- [16] Fonseca, E., Favory, X., Pons, J., Font, F. and Serra, X.: FSD50K: An open dataset of human-labeled sound events, *IEEE/ACM Trans. Audio, Speech, and Language Processing*, Vol.30, pp.829–852 (2022).

推薦文

本稿は、スマートフォンの内蔵マイクを用いた音響センシングによって路面状況を推定する手法を提案している。履物と路面状況を揃えた複数人の歩行音を収集し、歩行動作の個人差に頑健な推定手法を評価しており、斬新性・新規性ともに高い。評価結果から有用性も認められ、今後の発展・展開に期待できる。以上の理由により、本稿を推薦する。

(第30回マルチメディア通信と分散処理ワークショップ
プログラム委員長 荒川 豊)



白石 陽 (正会員)

2002年慶應義塾大学大学院理工学研究科博士課程単位取得退学。東京大学空間情報科学研究センター研究機関研究員、助手、助教を経て、2009年公立はこだて未来大学助教。2016年より同大学教授。博士(工学)。センサネットワーク、データベース、高度交通システムに関する研究に従事。ACM、電子情報通信学会、地理情報システム学会等各会員。



金澤 玲実 (学生会員)

2022年公立はこだて未来大学システム情報学部情報アーキテクチャ学科卒業。現在、同大学大学院システム情報科学研究科システム情報科学専攻博士前期課程在学中。モバイルセンシング、高度道路交通システムに関する研

究に従事。



石田 繁巳 (正会員)

2006年芝浦工業大学工学部卒業。2008年東京大学大学院新領域創成科学研究科修士課程修了。2012年同大学院工学系研究科博士課程修了。博士(工学)。2008年(株)アクティス入社。2013年米国ミネソタ大学客員研究員。2013年九州大学システム情報科学研究院助教。2021年公立はこだて未来大学准教授。無線通信、センサネットワークに関する研究に従事。2016年度山下記念研究賞、2023年IPSJ/IEEE Computer Society Young Computer Researcher Award。IEEE、電子情報通信学会各会員。