

# スマートフォンを用いた音響センシングによる 歩道路面状況推定に向けた特徴量の検討

金澤 玲実<sup>1,a)</sup> 石田 繁巳<sup>1,b)</sup> 白石 陽<sup>1,c)</sup>

**概要:** 路面の形状や材質、状態などの路面状況に起因する歩行者の転倒・転落などの事故が問題となっている。このような歩行者事故を防止するためには、歩道路面状況を考慮した経路案内や適切な靴の推薦といった歩行者支援を行うことが有効である。本研究では、路面状況ごとで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歩道路面状況を推定する手法を提案する。歩行者が持つスマートフォンで収集した歩行音から特徴量を抽出し、教師あり学習により歩道路面状況を推定する。本稿では、実際に4種類の路面状況を歩行しながら収集した音響データを用いて歩道路面状況推定に使用する音響特徴量を検討した結果を報告する。複数の音響特徴量を評価した結果、周波数スペクトルを縮約した音響特徴量のうち、低周波数成分の分解能が高いものが有効であることを確認した。

**キーワード:** 路面状況推定, スマートフォン, 音響センシング, 歩行者支援

## 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]。歩道路面状況に起因する歩行者の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、降雨後のぬかるみや冬季の部分的な積雪・凍結などの予期しづらく不安定な場所を通行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。歩道路面状況を把握することで、路面状況を考慮した経路案内や適切な靴の推薦といった歩行者支援を行うことが可能となる。

しかし、歩道路面状況については十分な情報提供が行われていない。北海道の冬期における歩道路面の滑りやすさを予報する「つるつる予報 (R)」のように特定の地域や路面状況を対象とした推定および情報提供は行われているものの、広範囲かつ多様な歩道路面状況を対象としたサービスは提供されていない。

歩道路面状況の推定に関しては、これまでに慣性センサや圧力センサを用いた研究が行われている[2][3][4][5]。しかし、これらの研究では推定対象となる路面状況に限られる、デバイスの装着による手間やコストがかかるという問題がある。

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手法を提案す

る。アスファルト舗装路と砂利道で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捉えることで、路面状況を推定する。音響センシングはデバイスの向きや位置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。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。

本研究でのデータ収集方法として、参加型センシング[6]を想定する。参加型センシングとは、多数の協力者が所持するモバイルデバイスを利用して計測したセンサデータを共有する方法である。本研究の推定対象には降雨や降雪といった気象による変化の激しい路面状況を含むため、参加型センシングによって広範囲のセンサデータを網羅的かつ継続的に収集する。

本稿では、実際にアスファルト舗装路(乾燥・湿潤状態)と砂路(乾燥・湿潤状態)を歩行し、収集した音響データを用いて歩道路面状況推定に使用する特徴量を検討した。その結果、MFCC (Mel-Frequency Cepstrum Coefficient)を用いた際の F-measure が 90.0%、CQT (Constant-Q transform) を用いた際の F-measure が 88.4%となり、周波数スペクトルを縮約した音響特徴量のうち、低周波数成分の分解能が高いものが有効であることを確認した。

本稿の構成は以下の通りである。2章では歩道路面状況推定の関連研究を示す。3章では歩行音を分析した予備実験の結果について示し、4章で分析結果を踏まえた提案手法について示す。5章では、推定に用いる特徴量の評価を行い、最後に6章でまとめとする。

1 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.

a) b1018092@fun.ac.jp  
b) ish@fun.ac.jp  
c) siraisi@fun.ac.jp

## 2. 関連研究

### 2.1 歩道路面状況推定に関する研究

歩道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[2][3]や圧力センサを用いる手法[4][5]、音響センサを用いる手法[7]などが報告されている。

慣性センサ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[2]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抽出した特徴量を用いることで平坦な道、段差、傾斜の緩やかな坂道をそれぞれ約70%、60%、30%の精度で推定できることを示している。佐藤ら[3]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。しかし、慣性データのみを用いた路面状況推定では、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。

圧力センサ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。若林ら[4]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平均97%の精度で推定できることを示している。Matthiesら[5]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、タータンなど6種類の路面状況を平均82%の精度で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。しかし、圧力データのみを用いた路面状況推定では、推定対象とする路面状況と硬度の近い路面状況を区別できずに誤推定する可能性がある。また、デバイスを靴に装着する手間やデバイスの導入コストがかかる。

音響センサ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[7]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種類の路面状況を平均83%の精度で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参加型センシングにおける収集者の確保が困難であると考えられる。

### 2.2 本研究の位置付け

本研究では、歩行者支援に向けた歩道路面状況の把握を目的として、スマートフォンを用いて収集した音響データから歩道路面状況の推定を行う。普及率の高いスマートフォンをセンシングデバイスとして用いるため、センシングへの参加難度や導入コストの面で優れている。また、歩道路面状況の推定に使用する音響特徴量については文献[7]でMFCCを取り上げているが、本研究ではMFCC以外の音響特徴量も対象として比較検討を行い、有効な特徴量を明らかにする。

## 3. 予備実験

本章では、予備実験として行った歩行音の周波数分析について述べる。文献[8]では場所や履物によって特定の周波数帯域にピークが現れること、文献[9]では歩行音が低周波数成分を多く含むことが報告されている。しかし、これらの関連研究と本研究では想定する路面状況やデータ収集状況が異なるため、同様の結果が得られるか定かでない。そこで、周波数スペクトル上で路面状況やデータ収集状況の違いによる差異が見られるかを調査した。

### 3.1 音響データの収集

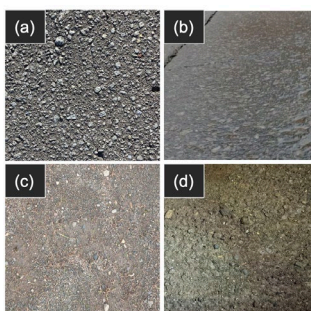
#### 3.1.1 対象とする路面状況

路面状況の要素として、1) 路面形状、2) 路面材質、3) 路面状態がある。以下に各要素の定義と例を示す。

- (1) 路面を横から見た際の形状 (例: 傾斜, 段差)
- (2) 路面を覆う材質 (例: アスファルト, タイル)
- (3) 路面上の状態変化 (例: 湿潤, 凍結)

本研究の推定対象としては、これらの3要素の組み合わせのうち、歩行者事故を誘発するリスクの観点から滑りやすいものや足を取られやすいものを選出する。具体的には、傾斜のあるタイル舗装路や砂路の湿潤状態などである。

本稿では予備実験として、アスファルト舗装路の乾燥状態 (asphalt\_dry)、アスファルト舗装路の湿潤状態 (asphalt\_wet)、砂路の乾燥状態 (sand\_dry)、砂路の湿潤状態 (sand\_wet) の4種類の路面状況においてデータ収集を行った (図1)。なお、いずれも路面形状は平坦である。



(a) アスファルト舗装路の乾燥状態 (b) アスファルト舗装路の湿潤状態 (c) 砂路の乾燥状態 (d) 砂路の湿潤状態

図1 データ収集を行った路面状況

### 3.1.2 データ収集時の状況

データ収集は、1人のデータ収集者が人通りや車通りが少ない静かな環境下で行った。また、デバイスを足首(ankle)と腰(waist)の位置にそれぞれ固定し、マイクを遮蔽せずに収録した(図2)。データ収集者の靴はADIHOOPS 2.0 (Adidas)を使用した。デバイスは足首側でZenFone5Z、腰側でPixel5を使用した。デバイスの録音形式はモノラルで、サンプリング周波数は44.1kHzである。

収集した音響データは、分析前に車両走行音や発声音といった明らかな雑音を除去した。また、路面状況ごとのデータ長を1,000秒に揃えた。

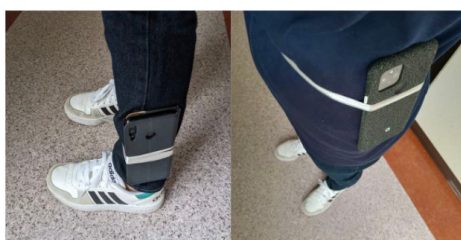


図2 デバイスの位置

### 3.2 収集したデータの分析

分析では各路面状況における音響データ1,000秒をフーリエ変換し、横軸を周波数、縦軸を正規化した振幅の大きさとする周波数スペクトルを求めた。

図3は、図1に示した4種類の路面状況から算出した各周波数スペクトルをデバイスの位置ごとに平均した結果を示す。デバイス位置が異なる2つの周波数スペクトルに共通して1,000Hz以下の低周波数成分における振幅が明らかに大きいという特徴が見られる。このことから本研究で想定するデータ収集状況においても歩行音の特徴が得られると考える。しかし、腰位置では500Hz以降で振幅が急激に減衰していることから、路面と靴の摩擦音のような歩行音の中でも比較的小さい音による情報が欠落していると考えられる。その原因として、デバイスと路面間の距離が離れていることや、デバイスのマイク性能差による影響が考えられる。

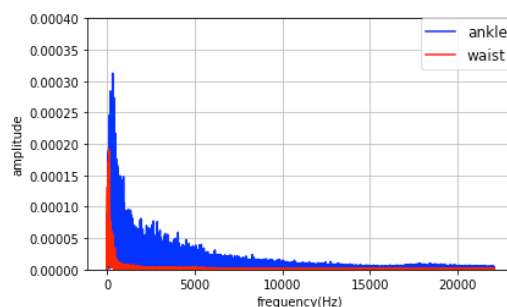
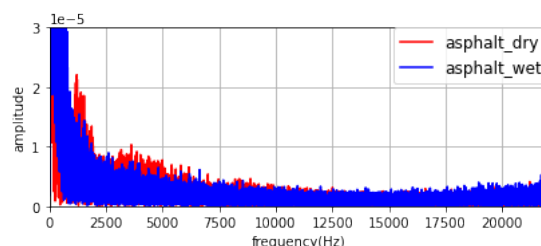
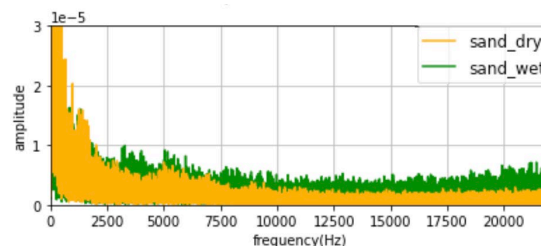


図3 歩行音の周波数スペクトル

図4、図5は、図1に示した4種類の路面状況において、腰位置のデバイスで収集した音響データに対する周波数スペクトルを示す。図4(a)はアスファルト舗装路、図4(b)は砂路の周波数スペクトルを示し、図5(a)は乾燥状態の

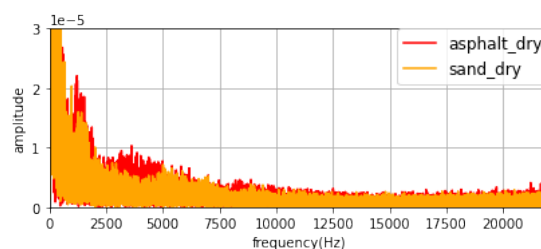


(a) アスファルト舗装路

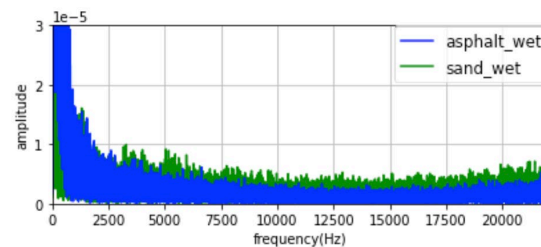


(b) 砂路

図4 路面材質別の周波数スペクトル



(a) 乾燥状態の路面



(b) 湿潤状態の路面

図5 路面状態別の周波数スペクトル

路面、図 5 (b)は湿潤状態の路面での周波数スペクトルを示す。なお、振幅の差をわかりやすくするためにグラフを拡大している。

各図より、路面材質および状態の違いによって周波数スペクトルに差異が現れていることが確認できる。特に 2,500~10,000Hz の周波数帯域において概形の違いが見られる。このことから、特徴量を抽出する周波数帯域を絞ることで、本稿で対象とする路面状況についても歩行音の周波数情報による推定が可能であると考えられる。

## 4. 提案手法

### 4.1 路面状況推定の流れ

本研究で行う路面状況推定の手順を図 6 に示す。まず、複数の歩行者がスマートフォンをポケットまたはカバンに入れた状態で歩行し、各路面状況における連続した音響データを収集する。次に、前処理で音響データから固定長の音響データを切り出す。そして、切り出した音響データから特徴量を抽出し、路面状況の正解ラベルとともに学習データとする。最後に推定モデルを構築し、路面状況推定に用いる。

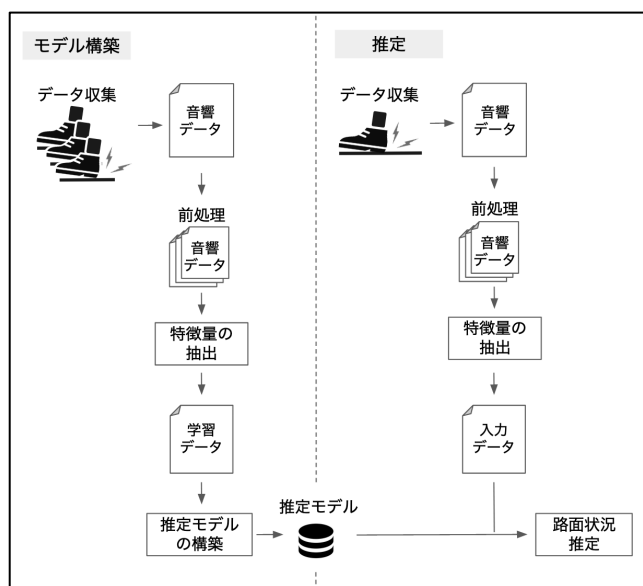


図 6 推定の流れ

### 4.2 前処理

音響データは、あらかじめ定めた一定時間ごとに切り出す。1データの長さは、1分間に 100 歩進むと仮定した際に 3 歩分の歩行音が含まれるよう 2 秒とする。音響データの切り出しはオーバーラップ率 0%で行う。

### 4.3 特徴量抽出

本稿では特徴量の基礎検討のため、以下の音響特徴量を抽出する。

- ・ MFCC: 音響データに対してフーリエ変換を行い、メル

フィルタバンクを適用した後に離散コサイン変換した音響特徴量である。周波数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 とする。

- ・ Cepstrum: 音響データに対してフーリエ変換を行い、周波数軸上で等間隔なフィルタバンクを適用した後に再度フーリエ変換した音響特徴量である。こちらもケフレンシ領域の数値列である。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 とする。
- ・ Chroma: 同音階の周波数における振幅を積算し、12 音に縮約した音響特徴量である。
- ・ CQT: 周波数と窓幅を反比例的に変動させてフーリエ変換した音響特徴量である。本稿では 3.2 節の周波数スペクトル上で差異が見られた 2,500~10,000Hz の周波数帯域にあたる 24 次を用いる。
- ・ ZCR (Zero-crossing rate): 信号振幅の正負が切り替わる頻度を表す音響特徴量である。信号に含まれる周波数が高いほど ZCR の値も高くなる。

各音響特徴量を本稿における検討の対象とした理由は以下の通りである。MFCC や Chroma は環境音推定の関連研究[7][9][10]において使用例があるため、本稿でも採用した。Cepstrum は MFCC と比較して高周波数成分の分解能を下げない場合の推定結果を調査するために採用した。CQT は、低周波数成分の分解能が高く、歩行音の特徴量として適していると考えたために採用した。ZCR は歩行音のピッチによる差異を捉えられることを期待して採用した。

### 4.4 推定モデル構築

推定モデルは教師あり学習で構築する。本稿ではアルゴリズムとして SVM (Support Vector Machine) を用いる。SVM は少ないデータ数で安定した推定結果が得られる手法であるため採用した。標準化した学習データを用いて推定モデルを学習し、構築した推定モデルを用いて歩道路面状況を推定する。

## 5. 特徴量の評価と考察

### 5.1 評価実験

評価実験では、各音響特徴量のみを用いた際の推定精度を 5 分割交差検証にて評価した。交差検証では、学習データをランダムに並び替えた後に 5 分割し、4/5 をトレーニングデータ、それ以外の 1/5 をテストデータとした。学習データの作成にあたって使用した音響データは、予備実験で収集した腰位置のデータを前処理で切り取ったもの 2,000 件(各路面状況 500 件)である。交差検証では、評価指

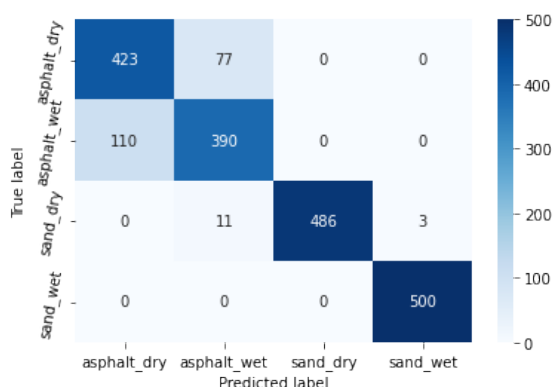
標として, Accuracy, Precision, Recall, F-measure をマクロ平均で算出した.

音響特徴量ごとの推定精度を表 1 に示す. スペクトル包絡を表す特徴量である MFCC と Cepstrum でそれぞれ F-measure 90.0% と 89.6% が得られた. CQT でも F-measure 88.4% と, 9 割近い推定精度が得られた. しかし, Chroma と ZCR の推定精度は他 3 つを大きく下回る結果となった.

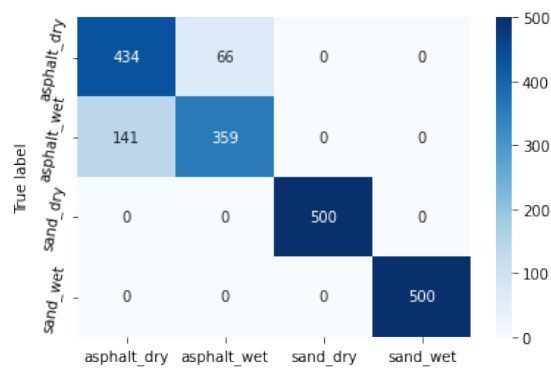
表 1 路面状況推定の精度評価結果 (%)

Feature	Accuracy	Precision	Recall	F-measure
MFCC	90.0	90.1	90.0	90.0
Cepstrum	89.7	90.1	89.7	89.6
Chroma	45.4	46.4	45.0	45.3
CQT	88.6	88.5	88.6	88.4
ZCR	65.0	65.8	65.0	62.0

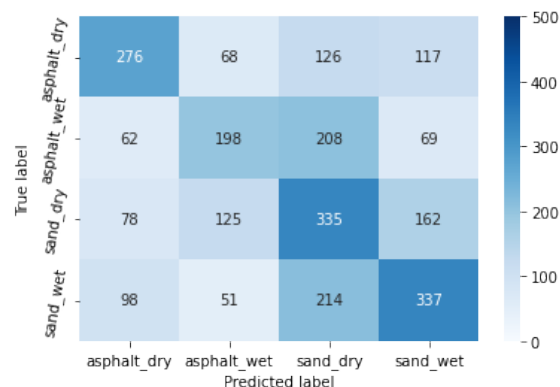
図 7 は, 表 1 に示した推定結果における混同行列を示す. まず, 表 1 で F-measure の高かった MFCC, Cepstrum, CQT の混同行列である図 7 (a), (b), (d) について述べる. Cepstrum は乾燥状態での材質差 (asphalt\_dry, sand\_dry) と砂路での状態差 (sand\_dry, sand\_wet) における推定精度が最も高いことが確認できる. また, CQT は湿潤路面での材質差 (asphalt\_wet, sand\_wet) とアスファルト舗装路での状態差 (asphalt\_dry, asphalt\_wet) における推定精度が最も高いことが確認できる. MFCC は各路面状況において最高または次点の推定精度を持ち, 汎用性が比較的高いため F-measure が高くなったと見られる. 一方, 表 1 で F-measure の低かった Chroma と ZCR では, 図 7 (c), (e) より, Chroma は全ての路面状況で, ZCR は湿潤状態の材質差以外で他 3 つの音響特徴量と比較して誤推定が多いことが確認できる.



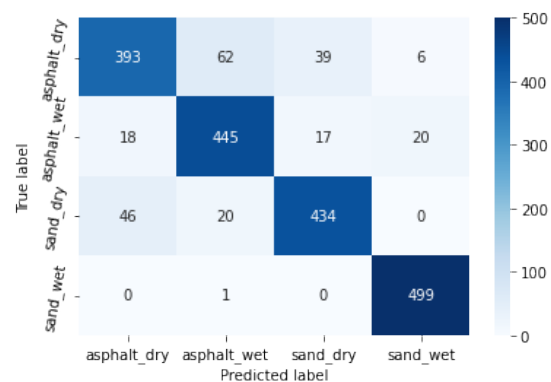
(a) MFCC



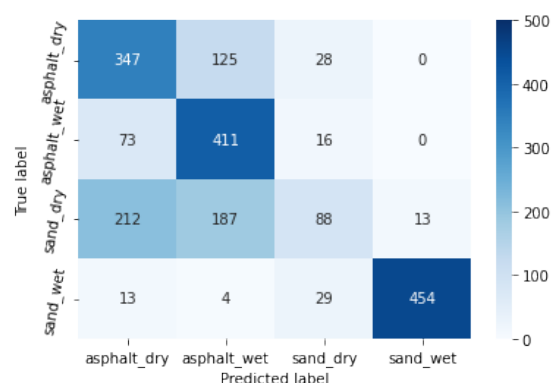
(b) Cepstrum



(c) Chroma



(d) CQT



(e) ZCR

図 7 音響特徴量ごとの混同行列

## 5.2 考察

MFCC と Cepstrum の結果から, 歩道路面状況推定に用い

る特徴量としてスペクトル包絡が有効であることが確認できた。これは2,500~22,050Hzで間断なく差異のあった砂路の状態差における推定精度が高かったためであると考えられる。しかし、全周波数帯域を音名に従って縮約した Chroma の F-measure が 50%を下回ったことから、各特徴量要素にどの周波数帯域を含めるかが結果に影響するものと考えられる。

また CQT の結果から、本稿で対象とした路面状況においては 10,000Hz 以下の周波数成分に含まれる路面状況ごとの特徴を捉えることが重要であると考えられる。これは MFCC と Cepstrum のうち、Cepstrum よりも低周波数成分の分解能が高い MFCC の方が F-measure が高くなったことから裏付けられる。

しかし、推定対象とする路面状況によって注視すべき周波数帯域が異なる可能性がある。Cepstrum と CQT の混同行列から、低周波数成分と高周波数成分でそれぞれ異なる路面状況の特徴が現れると考えられる。今後の研究においては対象とする路面状況を増やし、多様な路面状況で汎用的に推定可能な特徴量を調査する必要がある。

## 6. まとめ

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。本稿では、歩道路面状況の推定に使用する特徴量について評価し、周波数スペクトルを縮約した音響特徴量のうち、低周波数成分の分解能が高いものが有効であると示した。しかし、本稿の実験で用いた音響データは収集状況を限定しており、収集状況の違いによる推定への影響について調査できていない。また、複数の特徴量要素を組み合わせた場合については未検討であり、本稿で示した特徴量要素が最も有効であるかは定かでない。

今後は、推定対象とする路面状況や背景雑音の大きさが異なる場合や、複数の特徴量要素を組み合わせた場合での実験を行い、それぞれの要素が路面状況推定に与える影響について調査する。また、異なる条件下のデータを収集・分析することで、データ収集状況の違いに対応した推定手法を検討する。

## 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”, [https://www.shouhiseikatu.metro.tokyo.jp/anken/hiyarihat/documents/report\\_rainy.pdf](https://www.shouhiseikatu.metro.tokyo.jp/anken/hiyarihat/documents/report_rainy.pdf) (最終アクセス: 2021/07/18).
- [2] 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, D6-5 (2016).
- [3] 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(GN), 93(9), pp.1-8 (2015).

- [4] 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手法の冬季環境への適用”, 情報処理学会研究報告(MBL), 98(30), pp.1-7 (2021).
- [5] D. J. C. Matthies, T. Roumen, A. Kuijper, B. Urban., “CapSoles: Who is Walking on What Kind of Floor?”, Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI, pp1-14 (2017).
- [6] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, M. B. Srivastava, “Participatory Sensing”, World Sensor Web at SenSys, pp1-6 (2006).
- [7] 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).
- [8] 安田浩大, “周波数スペクトルの時間変化に着目した足音の解析”, 計測自動制御学会東北支部第 237 回研究集会, 10, pp.1-7 (2007).
- [9] 三嶽寛人, 渡邊拓貴, 杉本雅則, “足音を用いた路面状況推定手法”, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム 2018 論文集, pp.999-1007 (2018).
- [10] 立花巧樹, 中村優吾, 松田裕貴, 諏訪博彦, 安本慶一, “スマートウォッチの音響センサを用いたポイ捨てごみの種別認識手法の提案と評価”, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム 2021 論文集, pp.325-332 (2021).