

ユーザタイプ推定モデルの構築

— 決定木分析を用いたユーザタイプ分類モデルの効率化 —

小幡 浩大, 杉本 匡史, 長田 典子

関西学院大学感性価値創造インスティテュート

Constructing a User Type Estimation Model

— An Efficient User Type Classification Model Using Decision Tree Analysis —

Kodai OBATA, Masashi SUGIMOTO and Noriko NAGATA

Kwansei Gakuin Institute of Kansei Value Creation, Kwansei Gakuin University, Hyogo 669-1337, Japan

Abstract : In product design, it has become important to classify users into several types and to design products for each type. In the field of Kansei engineering, users have been classified based on questionnaires. However, questionnaires that have many items require a lot of effort from users. We solved this problem by extracting and modeling key questionnaire items for classification using decision tree analysis. We applied this method to the classification of motorcycle users based on emotion evaluation, and we conducted type estimation. As a result, we developed a method to perform type estimation with a small number of items. The accuracy of the type estimation model was 0.85. In addition, types classified with this method retained the Kansei information of the original rider types, indicating the method's high validity.

Keywords : *User type estimation model, Efficient, Decision tree analysis*

1. はじめに

近年、製品に対するユーザのニーズの多様化にともない、製品開発の分野では、機能性などの従来の価値に加え、新たな付加価値付与の必要性が高まっている。特に、新たな付加価値として、モノを通じての経験や感動が重視されるようになり、自己のありたい姿を実現する製品 [1] やイメージに合った製品 [2] が求められている。そのため、マーケティングにおけるセグメンテーションの変数の1つとしてユーザの嗜好や感情といった感性情報を把握することが重要であり、ユーザの感性モデルを構築するための主観評価実験が行われている [3, 4]。しかし、ユーザの感性モデルを構築する対象ドメインによっては、評価対象サンプル数や評価語数が非常に多くなるため、参加者の負担が大きいといった問題がある [5]。そのため実験で用いたアンケートをマーケティングの現場で利用することは実用的でない。実際著者らの研究 [3] でも、バイクユーザをタイプ分類するために90枚のバイク画像に対して、1枚につき4つの感情評定、つまり計360回の評定が必要であり、アンケートとしてはユーザへの負担が大きすぎる。

一方で、匂坂ら [6] は、決定木分析とIRS分析（項目関連構造分析）を用いることで回答に70～90分要する試験問題を、最長でも16.5分で回答可能な試験問題に効率化した。この研究は、試験問題の得点結果に基づき1次元のクラスタ分析を行うことでプログラミング学習者を4つの成績グループに分類し、目的変数に成績グループを、説明変数に試験問題

全てを用いた決定木分析を行った。その結果、4つの成績グループを分ける主要な問題を決定木で抽出し、試験問題を効率化した。また、IRS分析を用い、問題間の包含関係や難易度といった関連構造を考慮することでWeb-based診断システムを作成した。

このようなクラスタ分析の結果であるクラスタを目的変数、クラスタ分析に用いたデータを説明変数とした決定木分析による効率化は、主に以下の2点によってその妥当性を評価できる。1点目はコストであり、クラスタ分けの重要な変数だけを抽出することによって問題など項目をいかに減らせるかという点である。2点目は精度であり、コストとのトレードオフを考慮しつつ分類精度をいかに保てるかという点である。しかしながら精度に関しては、クラスタ分析は教師なし学習の一種でありデータは正解を持たないため、精度の評価が難しい。匂坂ら [6] は精度の評価のために診断システムの試験問題とは別に紙のテスト問題を用意し、その得点に基づくクラスタ分析を行い、結果のクラスタを正解と見なすことによって分類精度87%（31人中27人）を算出した。このケースではデータが得点という1次元データであったため、診断システムに基づくクラスタと紙のテスト問題に基づくクラスタが同じ性質（学習成績）を保持すると見なすことが可能であった。しかし、本研究で扱う感情などの感性情報は多次元ではばらつきがあり、しかも評価データが実利用を前提とした多人数であるため、正解を付与することが困難である。

そこで本研究では、著者らの先行研究で確立した感情評定に基づいたバイクユーザのユーザタイプ分類モデル [3] を効

率化する目的で、決定木分析を用いて実用上十分に少ない評
定項目でユーザタイプ推定を可能とするモデルを構築する。
さらに構築したユーザタイプ推定モデルの妥当性を検証する
ため、約2400名の評価データを用いてタイプ推定を行い、
結果のクラスタを元のユーザタイプ分類モデルのクラスタと
比較し、クラスタの性質が保持されているかの検討を行う
ことで精度評価とする。決定木分析は分類や予測において
ルールを明確化し、また項目の重要度を示す利点がある。
しかし感情評定からユーザタイプを推定するような多次元で
ばらつきのある感性情報に決定木分析を適用して有用性を
示した例はない。以下、2章ではユーザタイプ推定モデルの
構築について、3章ではユーザタイプ推定モデルの妥当性の
検討について示す。

2. 決定木分析を用いたユーザタイプ推定モデルの構築

本研究の概要図を図1に示す。まず2.1節において著者ら
の先行研究であるユーザタイプ分類モデル [3] の概要を述
べる。次に2.2節において分類モデルを決定木分析によって
効率化したユーザタイプ推定モデルについて説明する。

2.1 ユーザタイプ分類モデル

著者らの先行研究 [3] で示したユーザタイプ分類モデル
は、バイク乗車時に喚起される感情について、バイクユーザ
240名に評定実験を行い、7つのタイプが存在することを
明らかにしたものである。以下に概要を述べる。

初めに二輪乗車時に喚起される様々な感情を、コアアフェ
クトモデルに基づき4つの感情（「楽しい・爽快な」「気持ちの
良い・心地よい」「退屈な・物足りない」「不安な・怖い」）とし
て指標化した。また実験刺激として走行中のバイクを撮影し
た画像90枚（バイクの種類や、シーンが異なるもの）を用意
した。次にバイクユーザ240名に対して、バイク画像90枚を
呈示し上記の4つの感情について5段階（1:まったく感じない
～5:とても感じる）で評定する実験を行った。240名の内訳は、
インターネット調査会社経由参加者90名、二輪情報サイト
経由参加者100名、二輪商品の開発技術者50名であり、この
データは網羅性のあるバイクユーザのデータと考えられる。

実験参加者全員の360回（バイク画像90枚×4つの感情）の
回答に対してクラスタ分析を行ったところ、7つのタイプに

分類された。各ライダータイプは、感情傾向に基づいて、クラス
タ1がスタンダードライダー、クラスタ2がポジティブライダー、
クラスタ3がクールライダー、クラスタ4がスーパーポジティ
ブライダー、クラスタ5がマイペースライダー、クラスタ6が
アクティブライダー、クラスタ7がアグレッシブライダーと命名
された（各タイプの感情評定値は後に示す）。

2.2 ユーザタイプ推定モデルの構築と評価

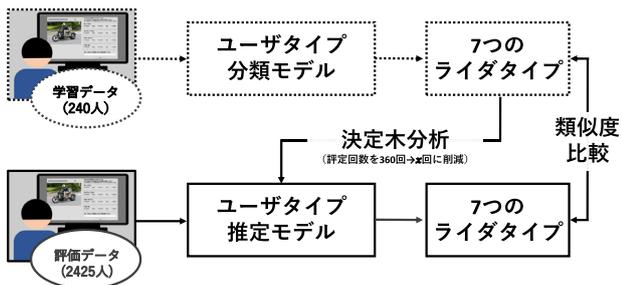
ユーザタイプ分類モデルでは7つのユーザタイプに分類す
るために360回の評定が必要であったので、これを効率化し、
少数の感情評定でタイプ推定ができるモデルを構築する。
7つのライダータイプを目的変数、バイク画像に対する4つの
感情の評定値（360回）を説明変数とした決定木分析を行っ
た。分析には統計ソフト「R」のパッケージrpartを使用した。
また、決定木の推定精度の評価指標にはF値を用い、F値の
算出には240名分の全学習データを検証用データとして用い
る代替推定法を用いた [7]。

決定木分析では、クラスターリングにおいて重要な評定項目
を木構造の上位ノード（ルール）として抽出することで、
木構造の効率化を図る。このとき、各ノードに必要とされる
最小データ数を下げるとF値は向上するが評定項目数も増加
し過学習の問題も生じる。そこでこの最小データ数を変えて
検討した結果を表1に示す。表1において、どの最小データ
数においても平均評定回数（評定項目数の平均）は全評定回
数（360回）と比較して十分に削減できていた。さらにF値
はいずれも0.75以上であり感性情報のような多次元でばら
つきを含む量のモデル化においては十分である。そこでF値
が0.85と最も高い決定木（ノードの最小データ数が3）を
本研究のユーザタイプ推定モデルとして採用した。

図2に採用したモデルを示す。左上のノードから条件分岐式
の質問項目に従って評定を行い、到達した右端の最終ノード
の最多タイプを実験参加者のライダータイプとして推定する。
具体的には例えば左上のノードは『背景が山で走行状態が
旋回でバイクカテゴリーがスポーツであるバイク画像』を見て
「不安な・怖い」に3未満の感情評定を付与する』というルールで
あり、これがユーザタイプ推定において最も重要な評定項目で
あることを示している。このモデルでは平均評定回数はわずか
4.5回（最小2回、最大7回）であり、平均評定回数が360回
であった元のユーザタイプ分類モデルと比較すると約98.8%の
大幅な削減となった。コストの点からは、実用上十分に少ない
評定項目数に削減でき、効率化が達成できたとと言える。

表1 最終ノードのデータ数の下限値ごとのF値と評定回数

葉における 最小データ数	F値	最大 評定回数	最小 評定回数	平均 評定回数
7	0.75	5	2	3.7
6	0.76	5	2	3.7
5	0.79	5	2	4.1
4	0.81	5	2	4.3
3	0.85	7	2	4.5



点線は著者らの先行研究 [3]、実線は本研究の内容である。

図1 研究概要図

ユーザタイプ推定モデルの構築

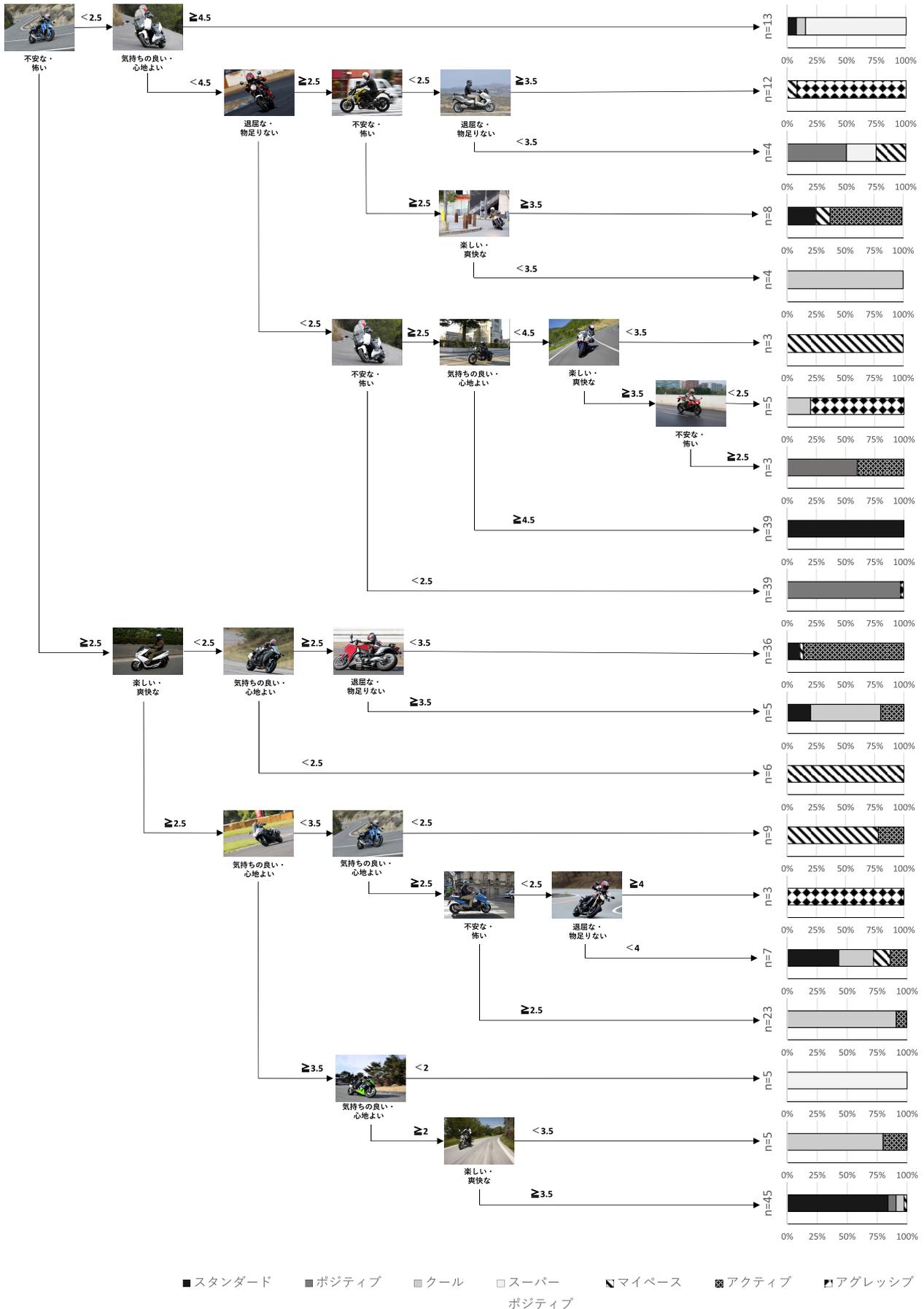


図2 バイクユーザのユーザタイプ推定モデル

3. ユーザタイプ推定モデルの妥当性の検討

本章では、ユーザタイプ推定モデルの妥当性を確認するため、新たな評価データを用いて推定精度の点から検証する。

3.1 実験方法

合計2425名(男性2374名, 女性51名, 平均年齢=45.2±10.5)のバイクユーザが参加した。実験参加者全員が普通自動二輪免許もしくは大型自動二輪免許を保有しており、また少なくとも1台のバイク(排気量不問)を所有していた。

実験参加者は、webアンケートに回答することによって調査に参加した。調査内容は、ユーザ情報(性別, 年齢, 所有しているバイクなど)の回答と画像評定実験であった。実験参加者は、画像評定実験において、90枚のバイク画像を1枚ずつ呈示され、「楽しい・爽快な」、「気持ち良い・心地よい」、「不安な・怖い」、「退屈な・物足りない」の4つの感情をどの程度感じるかをそれぞれ5段階(1:まったく感じない~5:とても感じる)で評定した。バイク画像は著者らの先行研究[3]と同様の90枚のバイク画像を使用した。

3.2 4次元ユークリッド距離を用いた類似度による検討

新たに取得した2425名分の評価データに関して、ユーザタイプ推定モデルに基づきタイプ推定を行い、推定結果の評価を行う。しかし前述したようにクラスタ分析は教師なし学習であり正解に基づく評価ができない。

そこで正解に基づく推定精度の算出の代わりに、推定結果である7つのライダータイプを、元のユーザタイプ分類モデルによる7つのライダータイプと比較して、対応するライダータイプ同士が同様の感情傾向を保っているかどうかを評価する。ライダータイプ毎に実験参加者個人の4つの感情毎の平均評定値を求め、さらにライダータイプ全体の感情毎の平均評定値を算出した。表2にライダータイプ毎に求めた各感情の平均評定値を示す。なおここでの推定結果のライダータイプは、スタンダードライダー(1')のように「タイプ名(クラスタ番号)」と表す。加えて表3にライダータイプ(1'~7')と元のライダータイプ(1~7)との類似度として、4つの感情の平均評定値の4次元ユークリッド距離を算出した。4次元ユークリッド距離が小さいほどライダータイプ間の類似度が高いと考えられる。

表2 ライダータイプ毎の4つの各感情の平均評定値

	楽しい・爽快な	気持ちのよい・心地よい	退屈な・物足りない	不安な・怖い
1'	3.95 (3.98)	3.92 (3.95)	2.39 (2.27)	2.95 (2.95)
2'	3.70 (3.88)	3.67 (3.86)	2.25 (2.13)	2.03 (1.94)
3'	3.42 (3.35)	3.37 (3.33)	2.61 (3.09)	2.87 (3.17)
4'	4.22 (4.63)	4.22 (4.33)	2.00 (1.49)	1.71 (1.73)
5'	2.89 (2.74)	2.88 (2.69)	2.88 (2.37)	2.94 (3.07)
6'	3.41 (3.36)	3.38 (3.28)	3.38 (2.73)	2.72 (2.91)
7'	3.39 (3.23)	3.37 (3.24)	2.74 (3.05)	1.96 (1.77)

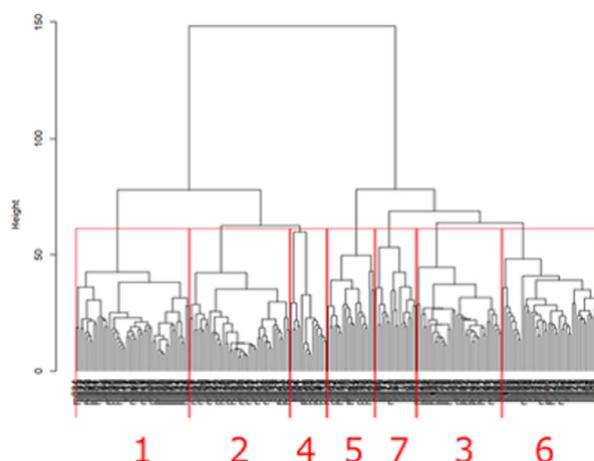
表3 本研究のライダータイプ(1'~7')と先行研究[3]のライダータイプ(1~7)との間の4次元ユークリッド距離

		先行研究[3]におけるクラスター						
		1	2	3	4	5	6	7
本研究におけるクラスター	1'	0.13	1.05	1.11	1.71	1.72	0.94	1.67
	2'	1.00	0.29	1.50	1.39	1.73	1.14	1.06
	3'	0.88	1.25	0.57	2.22	1.01	0.17	1.21
	4'	1.32	0.56	2.20	0.66	2.55	1.90	1.74
	5'	1.64	1.87	0.72	2.92	0.57	0.64	1.28
	6'	1.39	1.62	0.55	2.63	1.44	0.69	1.03
	7'	1.38	0.93	1.27	2.02	1.51	0.96	0.41

濃い灰色は類似度が最も高い値、薄い灰色は類似度が2番目に高い値である。

表3において、スタンダードライダー(1, 1'), ポジティブライダー(2, 2'), マイペースライダー(5, 5'), アグレッシブライダー(7, 7')の4つのライダータイプでは、同名のライダータイプとの類似度が最も高かった。また、クールライダー(3, 3'), スーパーポジティブライダー(4, 4'), アクティブライダー(6, 6')の3つのライダータイプでは、同名のライダータイプとの類似度が2番目に高かった。全てのライダータイプの類似度が上位2番目までにあり、本研究のライダータイプと著者らの先行研究[3]のライダータイプの高い類似度を確認した。

全てのライダータイプが同名のライダータイプと類似度が最も高いことが望ましいが、クールライダー(3'), スーパーポジティブライダー(4'), アクティブライダー(6')の3つのライダータイプは、同名のライダータイプとの類似度が2番目に高かった。その原因について考察するため、著者らの先行研究[3]で行われたクラスタ分析のクラスタの派生を確認した(図3)。図3から、ポジティブライダー(2)とスーパーポジティブライダー(4)、クールライダー(3)とアクティブライダー(6)が、クラスタ決定の直前で分離したタイプであり、それぞれの性質が類似していることを確認した。



赤文字は、クラスタ番号を示している。(1: スタンダードライダー, 2: ポジティブライダー, 3: クールライダー, 4: スーパーポジティブライダー, 5: マイペースライダー, 6: アクティブライダー, 7: アグレッシブライダー)

図3 クラスタの派生を示したデンドログラム

ユーザタイプ推定モデルの構築

また、ポジティブライダ (2) とスーパーポジティブライダ (4)、クールライダ (3) とアクティブライダ (6) の意味的な特徴を考慮すると、ポジティブライダ (2) とスーパーポジティブライダ (4) の意味的な特徴は類似している。一方で、クールライダ (3) とアクティブライダ (6) の意味的な特徴は若干の違いがある。以上の考察から、7タイプ中5タイプ、すなわち約71%は意味的にも正しく推定できていると考えられる。この精度は、ばらつきの多い感性情報のタイプ推定としては十分であると考えられる。これらの結果から、ユーザタイプ推定モデルの高い妥当性を確認した。

4. おわりに

本研究は、著者らの先行研究 [3] のユーザタイプ分類モデルに対し決定木分析を行い、タイプ分類において重要な評定項目を抽出することで、ユーザタイプ推定モデルを構築した。構築したユーザタイプ推定モデルは、先行研究と比較して、評定回数を大幅に削減できている。実用上十分な効率化が実現できた。また先行研究のライダタイプと本研究のライダタイプが、同様の感情傾向を示したことにより、ユーザタイプ推定モデルの高い妥当性を確認した。これらのことから、決定木分析を用いた効率化の手法は感性情報を扱うタイプ分類モデルにおいても有効であったと考えられる。

参考文献

- [1] 加藤哲夫：インダストリー4.0に想起される新たなイノベーション分類の提言，国際P2M学会誌，12(2)，pp.129-144，2018。
- [2] Schuitema, G., Anable, J., Skippon, S., and Kinnear, N.: The role of instrumental, hedonic and symbolic attributes in the intention to adopt electric vehicles, *Transportation Research Part A: Policy and Practice*, 48, pp.39-49, 2013.
- [3] 杉本匡史，今井将太，片平建史，山崎陽一，長田典子，益田綾子，岩田小笛，内山一：コアアフェクトモデルに基づいた二輪乗車時に喚起される感情の指標化：画像に対する感情評定を通じたライダのタイプ分類，電子情報通信学会技術研究報告，117(29)，pp.123-126，2017。
- [4] 向江亜紀，加藤俊一：3次元物体の形状と質感に対する感性のモデル化，情報処理学会論文誌データベース，47，pp.134-146，2006。
- [5] 秋山修一，鈴木謙介，Andreas, G.，永井義満，石津昌平：MTS法を用いた名詞による感性画像検索システムの構築，日本感性工学会論文誌，8(4)，pp.1171-1178，2009。
- [6] 匂坂智子，渡辺成良：プログラミング初学者のためのWeb-based学習診断システムの開発と評価，教育システム情報学会誌，27(1)，pp.29-38，2010。
- [7] 澤島秀成，杉山陽二，山岡俊樹：ユーザ属性と製品評価の関係分析における決定木の有用性，デザイン学研究，52(4)，pp.15-24，2005。



小幡 浩大 (非会員)

2018年 関西学院大学理工学部人間システム工学科卒業。2020年 同大学大学院理工学部研究科修士課程修了。専門は感性情報学。主にユーザのタイプ推定，ユーザの感情モデル構築に関する研究に従事。2019年度日本人間工学会関西支部大会優秀発表賞受賞。



杉本 匡史 (非会員)

2015年 京都大学大学院教育学研究科修了。同年，筑波大学人間系研究員。2016年 関西学院大学大学院理工学研究科博士研究員。2018年 同研究特別任期制助教。現在に至る。専門は認知心理学，感性工学。日本心理学会，日本認知心理学会，日本認知科学会会員。博士(教育学)。



長田 典子 (正会員)

1983年 京都大学理学部数学系卒業。同年 三菱電機(株)入社。1996年 大阪大学大学院基礎工学研究科博士後期課程修了。2003年より 関西学院大学理工学部情報科学科助教授。2007年 教授。2009年 米国パデュー大学客員研究員。2013年 感性価値創造研究センター長。2020年 感性価値創造インスティテュート所長。博士(工学)。専門は感性工学，メディア工学等。