

# 動作推定を用いた Web パーソナライズの開発

小野寺 斗弥<sup>†</sup> 猪股 一步希<sup>†</sup> 堀川 三好<sup>†</sup>

岩手県立大学 ソフトウェア情報学部<sup>†</sup>

## 1. はじめに

近年、モバイル端末の普及により、Web 上の利用者の情報を収集して一人ひとりに合わせてコンテンツの最適化を図る Web パーソナライズの導入が身近なものとなった。しかしながら、多くの Web パーソナライズ技術は、時・場所と検索・閲覧・購買履歴を組み合わせたものが多い。さらなる Web パーソナライズの進化による新たなサービス提供のためには、利用者の時・場所に加えて動作・状態に応じた情報配信を行う Web プラットフォームが必要となる。

本研究では、モバイル端末のブラウザにおいて収集可能なセンサ情報から利用者の動作・状態推定を行い、利用者の時・場所・状態に適した情報提供を行うことを目的としている。その一例として、本稿では、顧客が店舗内で行うモバイル端末操作を動作推定することでタップ等の操作なしで動的に情報提供を行う、動作推定を用いた店内情報提供システムを提案する。すなわち、顧客の端末保持姿勢のみで店内のフロアマップ、商品撮影からの口コミ情報提供、現在位置表示等を動的に切り替える Web システムを開発する。本稿では、利用者の動作推定モデルの精度、撮影画像の商品名認識精度、および小規模店舗での検証実験を通して得られた知見について報告をする。

## 2. 関連研究

専用アプリを用いてセンサ情報を利用した保持姿勢・動作推定を行う研究は多い<sup>[1]</sup>。しかしながら、Web アプリで取得できるセンサ情報を用いて動作推定する研究は少なく、データの特徴量が異なるため、動作推定モデルを構築する必要がある。また、行動状態を加味したアプリ開発が行われており、利用者の状態に合わせて適切な推薦を行うことの有用性が確認されている<sup>[2]</sup>。しかしながら、現実世界の利用者の姿勢・状態推定を考慮した研究は少ない。

Web personalization using motion estimation

Toya Onodera <sup>†</sup>, Ibuki Inomata <sup>†</sup>, Mitsuyoshi Horikawa <sup>†</sup>

<sup>†</sup> Software and Information Science, Iwate Prefectural University

## 3. 動作推定を用いた店内情報提供システム

### 3.1 提案システムの概要

本稿では、店舗内で Web 閲覧をしている利用者のモバイル端末の保持姿勢から動作推定を行い、欲しいと思われる情報提供や端末操作を動的に提供するプロトタイプを開発する。プロトタイプでは、利用者に3つの情報提供を行う。

#### (1) フロアマップ表示

利用者が歩行時に閲覧する場合にはフロアマップを表示する。

#### (2) 商品の口コミ情報提供

利用者がカメラ撮影を行う動作を行った場合、画面が切り替わり、カメラが起動する。その後、数秒後に自動で撮影した画像をクラウド上に送信し、画像解析から商品名等を取得する。さらに、取得された商品名をハッシュタグとして SNS から口コミ情報を収集し、利用者に提示する。

#### (3) 現在位置表示

利用者が天井看板などの高い位置のカメラ撮影を行う動作を行った場合、クラウド上に撮影写真を送付し、OCR から看板文字を読み取り現在位置を表示する。

### 3.2 動作推定モデル

上述の動作推定を行うにあたり、ブラウザから取得できるモーション情報（加速度・ジャイロ）を用いて動作推定モデルを構築する。複数の機械学習モデルによる分類器を比較し、CNN-LSTM を用いた動作推定モデル（図 1）を用いる。汎用的モデルとするため、動作分類は、「歩行・階段・直立（着席）・横たわる」の基本動

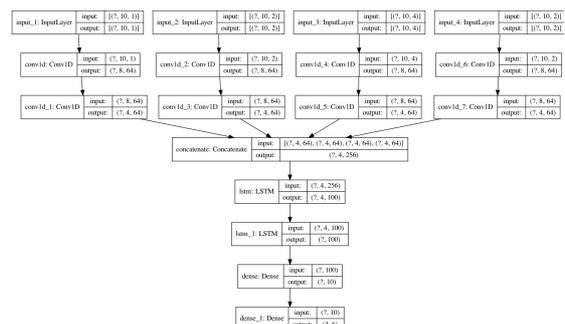


図 1 動作推定モデル (CNN-LSTM)

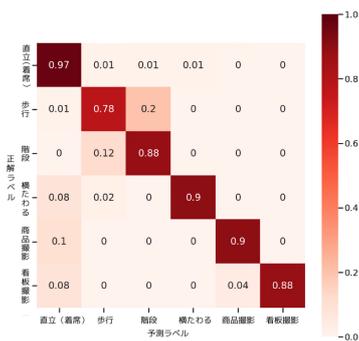


図2 動作推定モデルの混合行列

作に加えて、「商品撮影・看板撮影」を加えた6種類とする。

### 3.3 商品名の取得

商品のロコミ情報提供を行うにあたり、撮影した画像から商品名を取得する必要がある。本稿では、その方法として下記2つのクラウドサービスを利用する。

#### (1) 画像類似検索

撮影画像に類似した画像を Web 上から検索し、併せて画像の内容を表すラベルを取得する。商品ラベルが明確にない場合や、特徴ある外観等の場合に利用できる。

#### (2) OCR (光学的文字認識)

撮影画像から OCR 機能にて商品ラベルを検出する。取得した文字から商品ラベルのみを抽出するため、3つの手法を比較検討する。また、現在位置表示における看板文字認識も同様に OCR を利用する。

- 手法1: クラスタリング (k-means) による抽出
- 手法2: 中心座標にある文字ブロック抽出
- 手法3: 一番大きい文字ブロック抽出

## 4. 動作推定を用いた店内情報提供システム

### 4.1 実験目的および方法

動作推定モデルの精度および商品名の取得実験の精度比較を行った後、小規模店舗においてプロトタイプシステムの検証実験を行う。

### 4.2 動作推定モデルの精度検証

モバイル端末から加速度・ジャイロ値を 200 ミリ秒で取得し、特徴量変換を行う。サンプリング長を 2 秒とした場合の、混合行列を図 2 に示す。全体の精度は、F 値で 0.90 であるが、歩行・商品撮影・看板撮影の判別精度は高い。

### 4.3 商品名の取得の精度比較

画像類似検索および OCR による商品名抽出の精度は、対象商品や撮影方法に依存する。そのため、ペットボトル飲料 4 種を「棚にある商品を撮影」「手に取って撮影」で各手法 10 回撮影し認識精度を比較した。実装には、Google Cloud

表1 プロトタイプの実験結果

	動作	推定結果	機能提供	ロコミ
入/お菓子	歩行	歩行	○	-
お菓子	商品撮影	商品撮影	○	×
お菓子/カップ麺	歩行	歩行	○	-
カップ麺 (1)	看板撮影	看板撮影	○	-
カップ麺 (2)	商品撮影	商品撮影	○	○
カップ麺/飲み物	歩行	歩行	○	-
飲み物	商品撮影	商品撮影	○	○
飲み物/本	歩行	歩行	○	-
本 (1)	看板撮影	看板撮影	○	-
本 (2)	商品撮影	商品撮影	○	×

Vision API の Web Detection および Text Detection を用いた。その結果、中心座標にある文字ブロック抽出する手法の精度が高く 0.775 であったが、商品によっては類似画像検索の認識が高い場合も見られた。そのため、プロトタイプでは、両機能を並行して利用し、取得した商品名をハッシュタグとして SNS から取得した件数が多いものを利用することとした。

### 4.4 プロトタイプシステムの検証実験

動作推定モデルを TensorFlow.js を用いて実装することで、即時性を保ちつつ、データ取得から動作推定までをブラウザ上で完結可能となった。また、SNS として Instagram と Twitter API と連携してロコミ情報を提供する。検証実験は、大学構内の売店を用いて動作検証を行った。検証実験では、「お菓子/カップ麺/ペットボトル/本」について商品の情報を調べ、買い物を行うシナリオを想定した。実験結果の要約を表 1 に示す。動作に合わせた動的に情報提供は、概ね正しい結果となったが、ロコミ機能は正しく動作出来なかった。そのため、今後は商品名の取得アルゴリズムの再検討が必要だと考えられる。

## 5. おわりに

本研究では、モバイル端末のブラウザにおいて収集可能なセンサ情報から動作推定を行い、動的な情報提示アプリケーションの提案を行った。アプリケーションにより、利用者の時・場所・状態に応じた情報配信を行う Web サービスの提案への有用性を確認した。

今後の課題としては、提案手法の精度向上と状態推定への発展、妥当性の確認のために実環境での実験を進める必要がある。

### 参考文献

[1] 小川剛史, 朴燦鎬: ユーザの姿勢を考慮したモバイル端末の把持姿勢認識, 情報処理学会論文誌デジタルコンテンツ, Vol.5, No.1, pp.31-37, 2017.  
 [2] X.Wang, D.Rosenblum and Y.Wang, "Context-Aware Mobile Music Recommendation for Daily Activities, " in the ACM International Conference on Multimedia, 2012.