

# カメラを用いた顔位置計測による運動識別システム

加藤 花歩<sup>1,a)</sup> 夏成碩<sup>1,b)</sup> 杉浦 裕太<sup>1,c)</sup>

受付日 2021年1月26日, 採録日 2021年9月9日

**概要:** 健康を維持するために日常的な運動が重要であるが, 運動の継続が困難な人は多い. 本稿では手軽な運動管理や運動ゲーム作成のための, モバイル端末の内蔵カメラを利用した顔画像情報による運動識別システムを提案する. 本システムはカメラで運動中のユーザの顔画像情報を取得し, 運動の種類を識別できる. ユーザがカメラの画角に自身の顔を収めるように運動すると, システムは顔画像情報から顔上の追跡点を抽出する. 追跡点より算出した特徴量をサポートベクタマシンで学習することで9種類の運動を識別した. ユーザごとの精度検証の結果, 平均識別精度は97.2%となった. 加えて, システムの高速化や複数人の同時計測に向けて, 最適なウィンドウサイズや次元削減の検討, ユーザとカメラの位置関係による識別への影響の調査を行った.

**キーワード:** 運動識別, モバイル端末, 機械学習

## Exercise Recognition System Using Facial Image Information

KAHO KATO<sup>1,a)</sup> CHENGSHUO XIA<sup>1,b)</sup> YUTA SUGIURA<sup>1,c)</sup>

Received: January 26, 2021, Accepted: September 9, 2021

**Abstract:** Daily exercise has played a significant role for people in staying healthy; however, some people cannot do moderate exercise continuously. In this paper, we propose an exercise recognition system using facial image information for making exercise management convenient. The proposed system gets facial image information from a built-in camera on a mobile device and can recognize and classify multiple kinds of exercises. When a user exercises and their face is within the viewing angle, the system extracts features from the facial images. Via an experiment with user's own data, the average classification accuracy reached up to 97.2%. To improve the operation of the designed system, we also evaluated the suitable window size, dimension reduction, and influence of the user's standing position.

**Keywords:** exercise measurement, mobile devices, machine learning

### 1. はじめに

医療の発展や生活水準の向上にともない人々の平均寿命は延伸しており, 今後もさらに増加していくことが予想される [21], [25]. 一方で病気や高齢化に伴う運動機能の低下によって, 自由に動くことができず要介護期間が続いてしまうこともある. 近年では健康に生活できる期間「健康寿命」の延伸が重要だと唱えられており, これは平均寿命とともに伸びている傾向にあるが, 将来的にはそうではなく

なる可能性も提唱されている [14], [29].

健康を保つためには意識的に食生活と運動をバランスよく保つことが重要である. 特に運動には, 生活習慣病の予防や運動機能・認知機能の維持といった身体的な効果だけでなく, ストレスや不安の解消といった精神的な効果もある [3], [7], [18]. 一方で, 適度な運動を自力で継続することが困難な人は多い. トレーニング施設のトレーナーをはじめとする第三者による監督は運動の継続に有効である [8]. しかし, 毎日施設に通うのは金銭面でのコストが高い. このような問題の解決策として, 第三者の代わりに自動的に運動を計測し, ユーザへのフィードバックをする情報システムが提案されている [1].

このような情報システムを構築するうえで重要な点は,

<sup>1</sup> 慶應義塾大学  
Keio University, Yokohama, Kanagawa 223-8522, Japan  
<sup>a)</sup> kaho\_0128@keio.jp  
<sup>b)</sup> csxia@keio.jp  
<sup>c)</sup> sugiura@keio.jp



図 1 運動計測アプリケーション使用中の様子  
Fig. 1 Exercise with our application.

ユーザの身体情報を計測することである。身体の計測方法はこれまでも多く研究されてきており、代表的なものとしては、ウェアラブルセンサによる動作計測がある [16], [17]。ほかにも、カメラを用いて全身画像から抽出した身体の特徴点を用いる運動識別がある [2], [13]。これらの手法では共通して専用の機器や扱うための知識が必要であり、多くの一般的な家庭や施設にとって導入障壁となっている。

近年ではモバイル端末に多数のセンサやカメラが組み込まれており処理性能も向上しているため、これを用いた運動計測アプリケーションなども登場している [23]。モバイル端末は我々にとって身近な存在であり、専用の機器に比べ日常生活に導入しやすいという利点がある。我々はこのような市場に出回っているモバイル端末を用いた運動計測を考える。モバイル端末を用いた運動計測としては大きく2つに分類できる。1つは、端末自体を身体に装着する方法であり、この方法は身体動作や運動回数を正確に計測できる点で優れている。しかし、身体に装着することによる使用中の違和感や装着する手間が発生するため、運動のモチベーションが低下する恐れがある。もう1つは、内蔵カメラを用いた非接触計測であり、これは前者で課題となる違和感を解消できるが、計測時の環境や服装に制約を受けてしまう。本研究では、モバイル端末の内蔵カメラを用いるため、ユーザは端末を身体に装着する必要がない。加えて、顔画像情報のみを使用するため、ユーザは計測時の環境や服装を気にしなくてよい。

具体的には、モバイル端末の内蔵カメラを用いて顔画像を取得し、その特徴量を利用した運動の種類識別システムを構築する (図 1)。取得された特徴量に対しトレンド除去やハニング窓関数を適用した後、高速フーリエ変換 (以下、FFT) を行い、サポートベクタマシン (以下、SVM) による識別器を生成する。本稿では、顔画像情報のみを用いて計測できる運動の種類を調査することで提案手法の有効性を検証した。さらに、システムのリアルタイム処理に向けたパラメータの検証や複数人の同時計測実現のための検証を行った。

この手法によって、計測時の環境や服装の制約を低減した運動識別システムを構築できる。また複数人が重なり合うような空間においても同時に利用することができ、ユーザのモチベーション向上が期待できる。

## 2. 関連研究

### 2.1 身体装着型センサによる運動計測

身体を追跡し運動を計測する手段として身体装着型センサを用いる研究は多くなされている。Pernek らは上半身に装着した5つの加速度センサによる上半身の運動識別システムを提案した [16]。このシステムは2層のSVMで構築されており、運動の種類だけでなく運動中の負荷 (以下、運動強度) を推定できる。Prabhu らは腕に装着した慣性センサを利用して局所筋持久力を鍛える運動を識別し、運動回数を数える運動分析フレームワークを提案した [17]。周波数領域と時間領域の特徴量を抽出し、主成分分析 (以下、PCA) を用いて次元削減をすることで計算コストを小さくしている。Depari らは腕に装着した慣性センサを利用して運動を識別した [4], [5]。この研究では抽出した24個の特徴量に対しPCAを適用しており、計算リソースの要求が少ない手法を識別器生成に利用している。

これらの手法は身体にセンサを直接装着することで高精度に運動を識別できるが、ユーザが特別なセンサを購入する必要性や、センサを自身の身体に装着する手間を要する。本手法では、ユーザはモバイル端末を所持していれば利用できるため、導入コストは低く、かつ身体にセンサを装着する手間も不要という利点がある。

### 2.2 カメラによる非接触運動計測

RGB-DカメラやKinectなどの深度カメラを利用して、ユーザの全身情報を抽出して、人間の動作や姿勢を計測し、運動計測を行う研究が多く存在する。Ar らは運動を識別するための特徴量として運動の種類やユーザの姿勢、使用する道具の情報が利用できることを提案した [2]。3つの特徴量は市販のRGB-Dカメラから取得した全身画像より抽出することができ、それらを生成ベジアンネットワークで組み合わせることで、動作パターンを作成し運動回数を推定できる。Khurana らはジム内に設置したカメラで同時に複数人の運動を識別し計測するGymCamを提案した [13]。システムはキャプチャした画像からDense Optical Trajectoryで抽出したキーポイントより27種類の特徴量を生成し、ジムで行われる17種類の運動を識別したが、ユーザ間の距離が近すぎると正確に運動を検出できないことがある。

多様な運動を効果的に識別するためにSVMを使う研究も多い。Vox らは関節角度を利用し20種類の体操運動の識別システムを提案した [24]。関節角度はKinectでキャプチャした全身の3次元関節位置から計算され、14個の関節

角度をSVMで学習することで運動を識別している。Zhangらは深度カメラから取得した骨格形状の特徴量を利用した歩行パターン検出システムを開発した[26]。K-meansによってクラスタリングされた骨格形状の特徴量からマルコフ連鎖定常分布モデルを生成し、SVMで識別し歩行を検出している。このようにカメラ画像からの特徴量抽出とSVMによる人間の動作計測システムは有効であるため、本稿でもSVMで識別器を生成した。

一方で、上述の研究においてユーザ側は全身を撮影できる広い空間を用意する必要がある。このような問題に対し、身体の一部の情報を用いて運動を計測する研究も提案されている。Zulkifleyらは正面から撮影したカメラ画像から頭と上半身の領域のみを抽出して追跡することで膝角度を推定するシステムを開発した[27]。ただし、推定した膝角度をスクワットの運動強度推定に利用するのみで運動の種類識別を目指しているものではない。稲葉らはモバイル端末の内蔵カメラを用いて、顔画像から抽出した顔上の特徴点（以下、追跡点）の追跡による踵上げ運動計測システムを提案した[28]。追跡点の変化から運動回数が自動で計測され、さぼり検出のために、踵上げ運動と似たような動作からなる3種類の運動も識別できる。本研究ではこの稲葉らの研究を参考にカメラ画像から取得した顔画像情報のみを用いた運動識別を行うことで、運動計測システムの導入障壁を低減し、ユーザが手軽に運動を計測できるシステムを実現する。加えて、どのような種類の運動を正確に計測できるのかを調査することを目的としている。

### 2.3 モバイル端末を利用した運動支援

タブレット端末やスマートフォンなどのモバイル端末の普及にともない、これらを利用して人間の運動を計測・支援するアプリケーションが研究されている。特に医療やヘルスケア分野での利用を想定した研究が増えている。Gandomkarらは腕に装着したスマートフォンに内蔵されている加速度センサ、磁力計とジャイロセンサを利用した10種類の運動識別システムを提案した[9]。アームバンドでの装着というユーザにとって使い易い仕組みである上、高精度で運動の種類識別や回数計測が可能である。Depontiらはスマートフォンを所持した状態で行う、手首のリハビリテーションのためのゲームアプリケーションを開発した[6]。スマートフォン内蔵の加速度センサ値の急激な変化を検知すると、ゲームを停止し運動を止めるようユーザに警告する。これにより、ユーザは効果的なりハビリテーションを行えているかどうかを把握できる。Ockendonらはスマートフォン内蔵の加速度センサを利用した膝屈曲角度の測定システムを開発した[15]。ユーザが脛骨に沿うようにスマートフォンを装着すると、内蔵の加速センサ値から膝と地面間の角度を算出し膝の屈曲角度を推定する。また商用利用可能なアプリケーションも存在する[12], [22]。

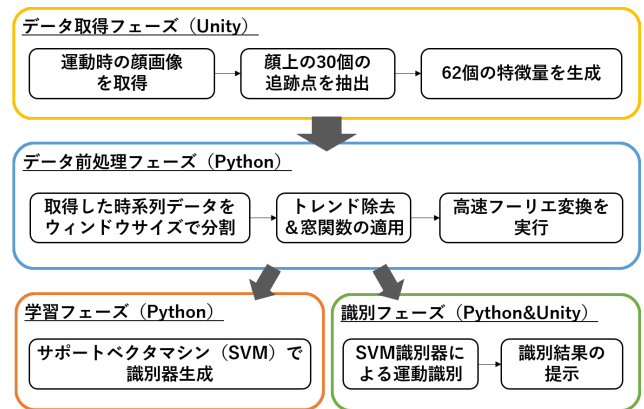


図 2 本システムの流れ

Fig. 2 Flow of our method.

これらのアプリケーションは全身画像から骨格情報を推定して運動を計測し、ユーザの姿勢の正確さを判定できるが、広い空間の用意や服装への配慮が必要である。

本研究では一般的な家庭や企業など様々な施設で簡単に導入でき、将来的に日々の運動を支援できるシステムを目指す。そのため、モバイル端末の内蔵カメラから取得できる顔画像情報のみを利用して運動を識別することで、導入障壁を低下し、かつ計測時の環境や服装を制限しないシステムを提案する。

## 3. 提案手法

### 3.1 システムの概要

我々のシステムではモバイル端末の内蔵カメラ画像より取得した顔上の追跡点を利用し、様々な運動を識別する(図2)。ユーザはカメラ画像に自分の顔全体が映る状態で運動する。次にSingle Face Tracker for Unity Plugin[20]を利用して、カメラ画像から顔上の30カ所の追跡点を抽出する。この際、画像解像度はカメラ自体の解像度によらず、640×480ピクセルに固定した。各追跡点は2つの座標値(x, y)を保持するので、30追跡点×2(x, y) = 60個の特徴量が取得でき、ユーザの運動中、60個の特徴量からなる時系列データが生成される。利用するモバイル端末は顔追跡と機械学習が実行できればよい。

### 3.2 運動の種類

健康に寄与する運動の中から自宅で簡単に実施できる9種類の運動を選んだ[10], [11](表1)。それらは立位状態で行う運動(以下、立位運動)と床上で行う運動(以下、床上運動)の大きく2つに分けられる。前者は「スクワット」、「かかとあげ」、「ジョギング」、「ももあげ」、「ウォーキング」の5種類であり、効果的に下半身を鍛えることができる。また立位状態で行うため、場所を問わず実施できる。後者は「腹筋」、「腕立て伏せ」、「背筋」の3種類であり、効果的に体幹を鍛えることができる。前者に対し、後

表 1 9種類の運動

Table 1 9 Kinds of exercise.

運動時の状態	運動
立位	スクワット
	かかとあげ
	ジョギング
	ももあげ
	ウォーキング
	無動作 (直立状態)
床上	腹筋
	腕立て伏せ
	背筋

者は床を利用するため室内のような限られた場所でしか実施できない。また運動していない状態を識別するために、立位運動には「無動作 (直立状態)」を加えた。ただし「スクワット」と「腕立て伏せ」は実施するのに一定の筋力や正しいフォームの習得が求められるため、ユーザ間の個人差が大きくなる可能性がある。そこで本稿では「スクワット」は椅子の立ち座りを繰り返す運動、「腕立て伏せ」は膝を床につけ足先を浮かせた状態で上体を上げ下げする運動とした。本システムは上記のような運動時の状態の異なる2タイプの運動を識別する。

### 3.3 運動の種類識別

本システムは Python を用いてデータ前処理や SVM 識別器生成を行い運動を識別する。60 個の特徴量のみでは z 軸方向の情報が欠如するため、z 軸方向の特徴量として2つの距離値を作成する。1つ目は x 軸方向の差分値である 0 番と 8 番の追跡点間の距離値、2つ目は y 軸方向の差分値である 5 番と 19 番の追跡点間の距離値である。これらの値が大きくなるとカメラに近づき (+z 軸方向)、小さくなると遠ざかる (-z 軸方向) と判断する。よって、合計 62 個の特徴量を識別器の学習に利用する (図 3)。フレームレートは 30 fps とした。

運動識別には運動中の顔上の 62 個の特徴量からなる時系列データを取得し、各特徴量ごとに処理を行う。まず基数 2 の Cooley-Tukey 型アルゴリズムの FFT を適用し、周波数成分を取得する。周波数成分を取得するにあたり、計測したデータは任意フレームごとに分割され、定数近似によるトレンド除去やハニング窓関数が適用される。このとき、サンプリングレートは 30 Hz、サンプル数は 128 であり、各特徴量につき 0 Hz から 14.8 Hz までの 64 次元の周波数成分を標準化した後、SVM での識別器生成に利用した。今回は SVM を用いて 9 種類の運動を識別しているが、SVM の他クラス識別方法は One-versus-Rest とした。これは任意のクラスを正例とし、その他のクラスを負例として 2 クラス分類を繰り返し行う方法であり、クラス数だけ識別器を学習すればよいという利点がある。

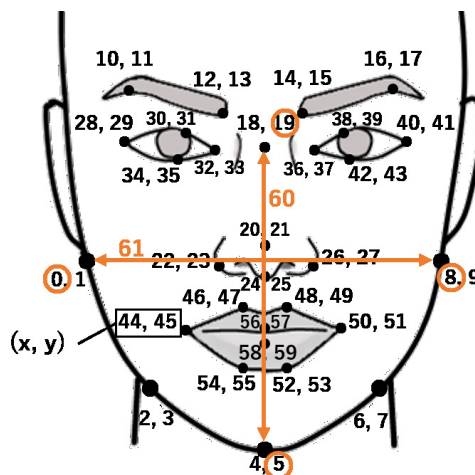


図 3 顔上の 62 個の特徴量 (偶数は x 軸, 奇数は y 軸の値を意味する)

Fig. 3 62 features on the face (The even means the value of x-axis, and the odd means the value of y-axis).

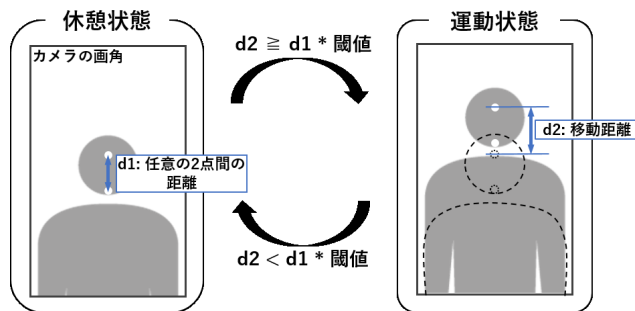


図 4 顔追跡による運動回数の計測

Fig. 4 Counting of exercise repetitions by face tracking.

### 3.4 運動回数の計測

稲葉ら [28] の研究を参考に運動回数を計測する。この手法は 2 点の距離と顔の移動距離を利用するため、運動開始時のユーザの位置や顔の大きさの個人差に対して頑健である。具体的には、顔の移動距離が任意の 2 個の追跡点間の距離に基づく任意の閾値を超えると、運動状態と休憩状態が入れ替わったと判定する (図 4)。判定後、2 つの追跡点における最大値または最小値が次の運動の始点となる。

本研究ではこの計測方法をもとに、識別した運動の種類に応じて自動で閾値を変化させ、様々な運動の回数計測を実現する。たとえば「スクワット」や「ジョギング」のような上下に動く運動をしているときは任意の 2 つの追跡点として図 3 における 5 番と 19 番 (y 軸値) を選択する。「ウォーキング」や「ももあげ」のような上下の動作が少ない運動の場合は 0 番と 8 番 (x 軸値) を選択する。このとき、運動の種類に紐づいて変化させる閾値は事前に決定する。

## 4. 検証実験

### 4.1 識別精度検証実験

#### 4.1.1 実験の概要

実験には 21 歳から 25 歳までの 8 人の実験協力者（男性：3 人，女性：5 人）が参加した。我々は実験開始前に表 1 の 9 種類の運動をそれぞれどのように実施するか実験協力者に説明した。実験協力者は各運動をそれぞれ 60 秒間実施し，各運動の間には次の運動の説明を含めそれぞれ 60 秒の休憩が与えられた。今回はモバイル端末としてノートパソコンを使用し，内蔵カメラの画角は約 60 度であった。モバイル端末は実験協力者の顔が画角内に収まるように配置した。具体的には，立位運動実施時には，モバイル端末は机の上に置かれ，実験協力者はモバイル端末からの距離 60 cm かつ正面の位置に立って運動を実施した。床上運動実施時には，運動の種類によってモバイル端末の置き場所や画角内に顔が収まる位置が異なった。「腕立て伏せ」と「背筋」ではモバイル端末は床上に置かれたが，「腹筋」では上体を起こした際の目線の高さと同様になるように椅子上に置かれた。モバイル端末からの距離は「腕立て伏せ」では 65 cm，「腹筋」と「背筋」では 40 cm であり，各距離において正面の位置に立って運動を実施した。連続での運動による疲労を考慮し，9 種類の運動の実施順序は各実験協力者でランダムに決定した。運動は合計で 8 人 × 9 種類の運動 = 72 回実施された。

#### 4.1.2 識別精度の評価

顔画像情報のみを用いて計測できる運動の種類を調査するために 2 つの手法で識別精度を評価した。

- 1 つ目は各実験協力者のデータにおける Leave-one-out 交差検証 (LOO) である。ユーザの運動をより正確に計測するために，学習データにユーザ自身のデータを使用することは重要である。LOO では各実験協力者のデータを学習データとテストデータに分けて識別精度を検証する。
- 2 つ目は全実験協力者のデータにおける Leave-one-subject-out 交差検証 (LOSO) である。他者のデータを学習データとして使用しても識別できるならば，ユーザのキャリブレーションの手間などを減らすことができる。LOSO では選択した実験協力者以外のデータを学習データ，選択した実験協力者のデータをテストデータとして識別精度を検証する。

全実験協力者の運動の最大周期はおよそ 100 フレームであったため，取得した各人各運動の時系列データをそれぞれ 100 フレームごとに分割した。運動周期より長いフレームサイズで分割したことで分割データ間で運動周期の開始位置のずれが発生するが，本手法では時系列データから抽出した周波数特性を使用するため位置ずれの影響は受けない。1 周期以上のデータが含まれることで発生するスペク

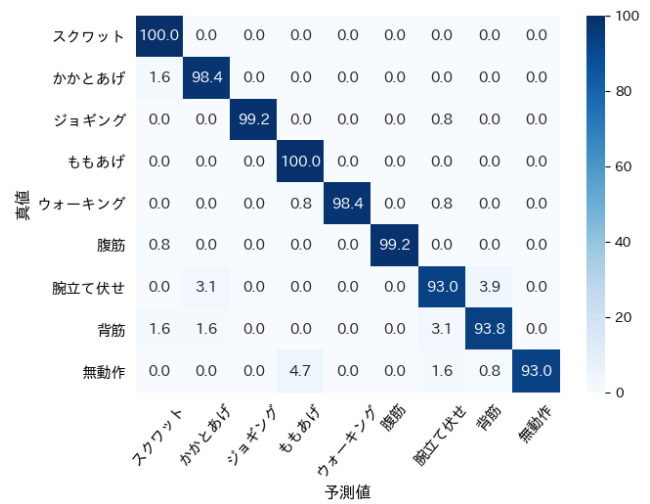


図 5 LOO での精度検証結果/%

Fig. 5 Result of classification by LOO(%).

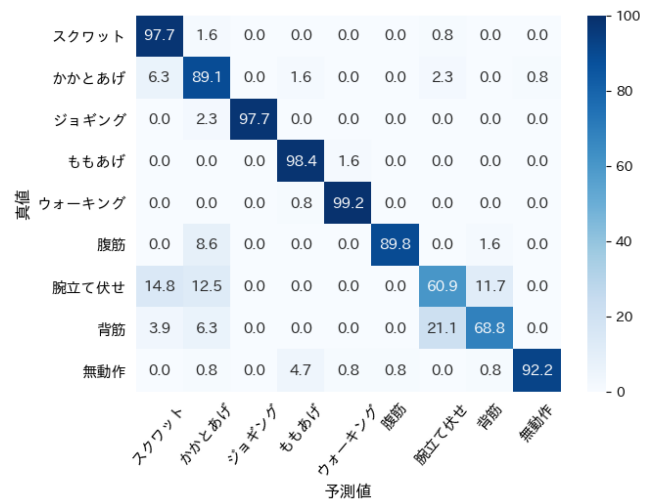


図 6 LOSO での精度検証結果/%

Fig. 6 Result of classification by LOSO(%).

トル値への影響も，学習前の標準化処理によって軽減される。この際，運動中に一時的なフレームレートの低下が発生し，取得した時系列データのフレームサイズが想定よりも小さくなってしまった。そこで最小フレームサイズのデータから分割可能な最大個数のデータを使用することとした。今回は，各人各運動で時系列データの開始位置から 16 個のデータを生成し，余剰部分は使用しなかった。FFT におけるサンプル数を 128 としたため，分割後のデータに対し不足部分は両端からゼロパディングした。62 個の特徴量に対しそれぞれ 64 次元の周波数成分を抽出し，合計で 3,968 次元のデータとなった。検証結果を図 5，図 6 に示す。

LOO での全実験協力者における 9 種類の運動の平均識別精度は 97.2%，立位運動のみの平均識別精度は 98.1%，床上運動のみの平均識別精度は 95.3% であった。LOO での識別精度は高く，運動時の状態が異なる運動も効果的に識別できた。図 5 より，「スクワット」と「ももあげ」の識別に対し高い性能を示していることが分かり，それぞれ

100.0%の精度で識別できた。一方で「腕立て伏せ」と「背筋」、「無動作」は他の運動よりも識別精度が低く、「腕立て伏せ」と「背筋」は互いに誤識別されることがあった。原因の1つとして、これらの運動の実施中、一部の実験協力者はカメラのある前方ではなく下方を向いてしまい顔追跡が外れてしまっていたことが考えられる。加えて、床上運動ではカメラを床または床に近い高さの位置に置いたため、十分な光を得られず高いフレームレートを維持できなかった可能性がある。また「無動作」は「ももあげ」と誤識別されているが、「ももあげ」中はユーザの身体が他の運動に比べてゆっくりと動くため、運動中に顔の位置があまり変化しなかったことが起因していると考えられる。

また LOSO での平均識別精度は 88.2%、立位運動のみの平均識別精度は 95.7%、床上運動のみの平均識別精度は 73.2%であった。図 6 より、特に「腕立て伏せ」と「背筋」の精度が LOO での識別結果より低かったことが分かる。これは顔追跡の欠如や運動強度に個人差があったことが原因と考える。また「かかとあげ」と「腹筋」の精度も少し低かった。7人の実験協力者は平均 98.2%の精度で「腹筋」を識別できていたが、ある実験協力者の結果では「腹筋」の識別精度は 31.3%であり、「腹筋」の多くを「かかとあげ」と誤識別していた。これは他のデータに比べると精度が低く、この実験協力者の「腹筋」は他の実験協力者とは異なる動作をしてしまっていたため、学習データとして適していなかった可能性がある。加えて「かかとあげ」自体が単純な上下運動であり移動量も大きくないため、様々な運動と誤識別されやすいと考える。LOO と比較すると LOSO の精度は 9%低下するが 80%を超える精度が得られており、識別の正確さとトレードオフではあるが他者の学習データを使用しても運動を識別できることが分かった。適度な運動ペースや運動強度で運動することは大切だが実施するのは困難であり、たとえばスクワットで腰を痛めるように、誤った運動を続けると怪我をするリスクがある。他者のデータを学習データに使用することで、トレーニング施設のトレーナーなどが実施したデータを識別器生成に使用できるようになり、ユーザの運動効果の向上が期待できる。

また実験中、話したり笑ったりしながら運動する実験協力者が存在しており、識別精度に何らかの影響を与えている可能性がある。現実にもそのような状態で運動する人々は存在しており、口の追跡点を除くことでさらに精度を向上できる可能性がある。

#### 4.1.3 ウィンドウサイズと識別精度の関係の検証

4.1.2 項では、ウィンドウサイズを実験協力者の運動の最大周期である 100 フレームとしてデータを分割した。本システムのフレームレートは 30 fps であるため、運動識別には約 3 秒分の時系列データを要したが、リアルタイムで運動を計測して結果をフィードバックするには計算時間の短縮が期待される。一方で、ウィンドウサイズが短すぎると精

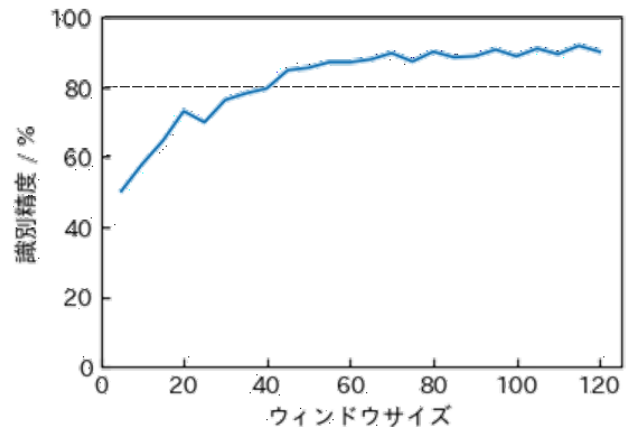


図 7 ウィンドウサイズと識別精度の関係

Fig. 7 Relation between window size and accuracy.

度の高い識別ができなくなると考えられ、これはトレードオフの関係である。そこで提案手法の高速化に向けて最適なウィンドウサイズの調査が必要であり、本節で実施した。

本検証では 5 フレームから 120 フレームまで 5 の倍数ごとにデータを分割し、LOSO で識別精度を検証した。このとき、FFT のサンプル数はすべて 128 とした。分割するウィンドウサイズによって取得できるデータ数に偏りが生じるため、120 フレームごとに分割した場合に合わせて学習に使用するデータ数を決定した。また 4.1.2 項と同様にフレームレートの低下を考慮したところ、学習に使用するデータ数は各人各運動で 13 個となった。結果を図 7 に示す。

ウィンドウサイズが 70 フレームを超えると、識別精度は 90%付近に安定した。全実験協力者の運動周期は平均 81 フレームであったため、最適なウィンドウサイズは運動の平均周期に近い値になることが予想できる。一方で、45 フレームを超えた時点で識別精度は 80%を超えているため、精度の低下と引き換えに、より小さいウィンドウサイズを使用できる可能性も示唆される。

#### 4.1.4 特徴量削減による識別精度の検証

これまでの実験では顔上の 62 個の特徴量を利用して SVM による識別器を生成し、LOSO での検証では 88.2%の識別精度を得た。一方で、62 個の特徴量から生成した 3,968 次元のデータの学習は計算コストがかかる。本節では 4.1.3 項と同様にシステムの高速化に向けて、識別器生成に使用する特徴量の削減による識別精度を検証する。

これまで使用してきた 62 個の特徴量のうち、60 個の特徴量は顔上の各追跡点の x, y 軸方向の座標値であるため、運動中にこれらは一律に動いている可能性が高い。そこで 60 個の特徴量を x, y 軸方向で分け、軸ごとに 30 カ所の追跡点に対して PCA を実行した。その結果、第 1 主成分の寄与率は 90%を超えたが、第 1 主成分における因子負荷量の大きさは全追跡点で同一となったため、各軸方向の追跡点で識別精度に大きく寄与する点が存在しないことが

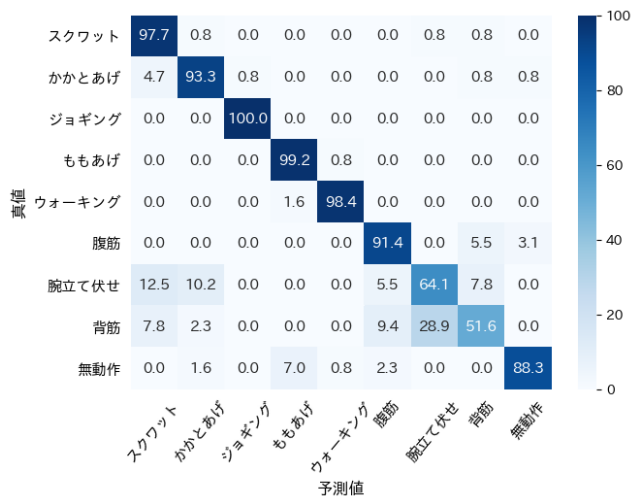


図 8 4 個の特徴量から生成した識別器による精度検証結果/%  
Fig. 8 Result of classification with four features(%).

確認できた。これをふまえ、本項では新たに 4 個の特徴量を用いて識別精度を検証した。そのうち 2 個は 4.1.2 項で算出した z 軸方向の特徴量である 2 つの距離値である。残りの 2 つは、x, y 軸方向ごとに算出した平均座標とした。4.1.2 項での考察や運動中のユーザの瞬きによる影響を考慮し、平均座標の算出では口と目の追跡点 (28 番~59 番) を除外した。よって、4 個の特徴量に対しそれぞれ 64 次元の周波数成分を抽出し、合計で 256 次元のデータとなった。検証方法は LOSO とし、結果を図 8 に示す。

平均識別精度は 87.1% となり、62 個の特徴量による識別器を利用した 4.1.2 項での結果と比較するとその精度は 1.1% 低下した。一方で、4.1.2 項では運動を 1 回識別する処理を行うのに 6.6 ミリ秒を要していたが、本節では 0.5 ミリ秒と 1/15 になりシステムの高速化を実現できた。リアルタイム識別時に加え、識別器生成時の高速化が期待できる。個別に運動の識別精度を比べると、「背筋」の識別精度が 4.1.2 項での結果より 17.2% 低下しており、床上運動どうしでの誤識別率も増加した。また運動時の状態ごとの平均識別精度を比較すると、立位運動は 96.2% となり 4.1.2 項での結果より精度が 0.5% 向上したのに対し、床上運動は 69.0% となり精度が 4.2% 低下した。これより立位運動時には 4 個の特徴量による識別器を使用の方が計算速度も識別精度も向上するが、床上運動時には特徴量を増やした方が誤識別しにくくなることが判明した。床上運動ではモバイル端末の位置やユーザの運動時の姿勢により顔追跡が困難な場合があり、4 個の特徴量のみでは顔の一部が誤追跡された際に生じる部分的な情報欠落の影響を補えないことが考えられる。

#### 4.2 ユーザとカメラの位置関係による識別精度の検証

ユーザの運動中、カメラ画像の中央につねに顔が映り込むとは限らない。また複数のユーザが同時に映り込むよう

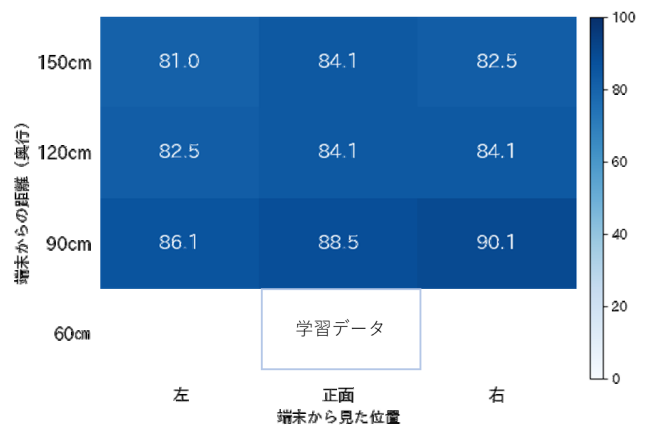


図 9 ユーザのカメラに対する位置と精度検証結果/%  
Fig. 9 Result of classification by changing standing position(%).

な環境での利用では、カメラ画像の複数箇所に顔が存在することになり運動中のユーザの顔が誤追跡される恐れがある。これに対し、ユーザのカメラに対する位置が運動識別に影響しないと仮定すると、複数人の同時計測を実現することで誤追跡の影響を低減できる。そこで本節では提案システムにおいてユーザとカメラの位置関係が運動識別に与える影響を調査した。ただし、カメラの画角を考慮して複数人で同時計測しやすい運動に絞り、本稿では立位運動のみの識別精度を検証した。

これまでの節では立位運動時にはモバイル端末から 60 cm 離れた位置で運動を実施していた。本節ではモバイル端末からの距離を 30 cm 間隔に 60 cm から 150 cm まで変更しそれぞれの識別精度を検証した。また各位置においてカメラの正面、正面から右に 30 cm (右)、正面から左に 30 cm (左) の 3 パターンの位置で検証した。ただし端末からの距離が 60 cm の場合には右と左の位置ではカメラの画角から顔が外れてしまうため除外し、合計 10 カ所の位置で検証を行った。以下、端末からの距離  $\alpha$  cm かつ  $\beta$  の位置を  $\alpha$  cm  $\times$   $\beta$  と表現する。

本実験には 22 歳から 24 歳までの 3 人の実験協力者 (男性: 1 人, 女性: 2 人) が参加した。彼らのうち 2 人は 4.1.1 項の実験に参加していたが、残りの 1 人はそうではなかったため、実験開始前に再び実施する運動について説明した。本実験では実験協力者は 10 カ所の位置で 6 種類の運動をランダムに実施し、4.1.1 項と同様の手法でデータを取得した。この際、位置は「60 cm  $\times$  正面  $\rightarrow$  90 cm  $\times$  右  $\rightarrow$  90 cm  $\times$  正面  $\rightarrow$  90 cm  $\times$  左  $\rightarrow$  120 cm  $\times$  右  $\rightarrow$  120 cm  $\times$  正面  $\rightarrow$  ...」という流れで変更した。

ユーザのカメラに対する位置が識別精度に与える影響を検証するため、60 cm  $\times$  正面での実験協力者のデータのみを識別器生成に使い、それ以外の位置のデータをテストに使用した。識別器生成時には 4.1.4 項で提案した 4 個の特徴量を使用した。結果を図 9 に示す。

結果よりすべての位置で80%以上の識別精度となった。変更した位置の順に識別精度は低下したが、これは実験協力者の疲労によって運動ペースなどが変化したことが原因と考えられる。また、実施順をランダム化しなかったことが関係している可能性もある。加えて、実験協力者が端末から120 cm以上離れた場合、すべての位置で顔追跡が開始時に作動しなかった。そのため、実験協力者は数秒間端末に近づき、顔追跡を作動させる必要があった。同様の位置において運動中に顔追跡が外れてしまうこともあり、実験協力者には同様に対処してもらった。

本システムはユーザが画角内に顔を収めていれば位置に関係なく運動を識別できるため、複数人で同時に計測することも可能である。誰かと一緒に行う対戦ゲームや協力ゲームなどの新たなエンタテインメントの創出が期待できる。

### 5. アプリケーション例

顔追跡やユーザに表示するインタフェースをUnity、運動識別のための機械学習処理をPythonで実装し、パソコン上でHTTP通信を介して実行される計測アプリケーションを作成した(図10a)。アプリケーションには60秒の時間制限があるモード(60s Time Attack)とユーザがいつでも運動計測を終了できるモード(Training)の2つのモードを作成した。両者の主な違いは時間制限機能の有無のみであるため、以下では後者のモードについてのみ説明する。

まずアプリケーションを起動すると、カメラ画面が表示される。ユーザが自身の顔をカメラの画角内に収めると、システムは顔を検出しピンク色のマーカーを付与する

(図10b)。マーカー付与後、ユーザは表1の9種類の運動を任意で実施でき、任意のタイミングで中断することもできる。アプリケーションは4.1.4項で検証された4個の特徴量による識別器を用いてユーザの運動を自動で識別し、運動の種類に応じて回数を計測する。運動終了後、実施結果がユーザに提示される。また直近の7日間のデータが保存され、どの運動を何回実施したかを確認できる(図10c)。

### 6. 制約と議論

本稿ではモバイル端末の内蔵カメラの画像から取得された顔上の追跡点による運動計測について検証を行ったが、システムにはいくつかの制約や課題が存在する。

まず顔追跡の欠落である。一般的な顔追跡ミドルウェアでは追跡点を抽出するために顔全体の画像を必要としており、本システムでも同様である。ユーザの顔が画角から外れてしまう、もしくは顔の正面に遮蔽物が存在する場合に顔を追跡できなくなるため、本システムの利用時ユーザはつねに画角内に顔が収まり、顔全体が映っている状態で運動する必要がある。そのため、「腹筋」や「背筋」といった、モバイル端末に対して前後に動く運動では特に注意して運動しなければならない。一方で、顔追跡が一定時間外れた場合、前の運動が続いていると想定して運動の種類を決定し、決定した運動における運動周期から顔追跡が外れた時間における運動回数を推定できる可能性がある。また、4.1.4項の結果より、顔上の追跡点の位置による識別能力の差がないことが分かっており、顔の一部分の追跡による運動計測が期待できる。COVID-19によりマスクを着けて生活することが一般的となっている状況も考慮すると、眉や顔の輪郭などの部分追跡による運動計測の実装が必要である。

つねに画角内に顔を収めた状態で運動するという制約がある一方で、本システムは日常の行動認識にも活用できる可能性がある。運動中には疲労により自然と俯いてしまうなど、カメラに顔を向け続けられない場面が想定されるが、顔を上げた状態で行われる行動であれば本制約は緩和される。たとえば、歩行動作や掃除機をかけるなどの家事動作があげられる。俯瞰位置からカメラで室内を撮影することでユーザの行動を識別できる可能性がある。今後の展望として、運動に限らないユーザの日常的な行動を顔位置計測によって識別できるか検証したい。

次に環境光の制約である。本システムは顔画像を利用しているため、この計測を保持できる明るい環境での実施が要求される。これまでの実施から、市販の昼光色の室内照明や日中の太陽光などで十分であることが分かっているが、屋内でも直射日光が差し込むような環境においては顔追跡に失敗してしまうことが多い。また自身の腕などによって顔に影が頻繁に生じるような運動では十分に顔追跡できない。加えて、この環境光量はシステムの動作フレームレ

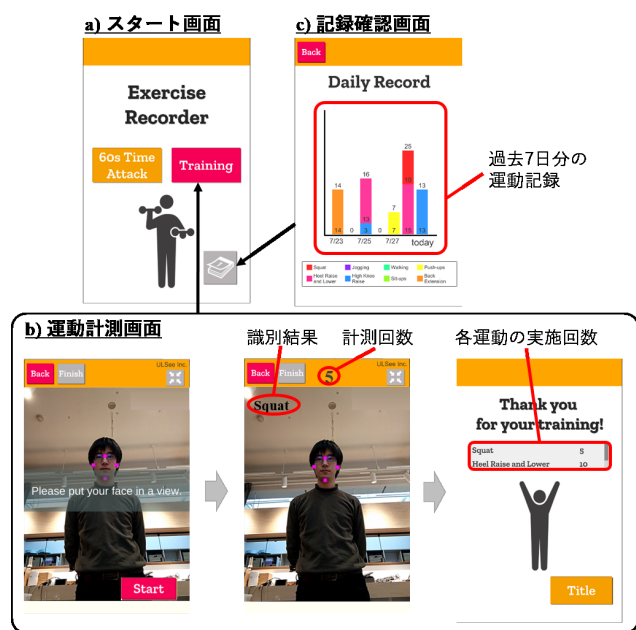


図10 アプリケーション画面  
Fig. 10 Application view.



トにも影響を与えることから、滑らかにデータを取得するためには高フレームレートを維持する必要がある。

本稿では LOO と LOSO の 2 種類の検証を行った。LOO では実験協力者自身のデータを用いて識別器を生成し、LOSO では他の実験協力者のデータを用いて識別器を生成した。前者に比べると後者は精度が低下し、特に床運動の精度の低下は顕著であった。床運動実施時は実験協力者間で運動ペースや運動強度における個人差が大きかったことがあげられる。また本実験では実験協力者が若年層に偏っているが、年齢や運動習慣の違いは運動ペースや運動強度に影響を与える恐れがある。これに対し、より多くのユーザのデータによる個人差を吸収できる識別器や年齢別・男女別の識別器、個人に特化した識別器などを生成することで精度を改善できると考える。運動時の疲労もユーザの運動ペースなどに変化を与える恐れがあるため、疲労を考慮した識別器生成を行いたい。また運動とは適した運動ペースや運動強度で行うことで効果を増幅し怪我を負うリスクを低減できるため、ユーザごとの個人差は少ないことが望ましい。今後は、運動ペースや運動強度を推定するシステムの構築や正しく実施された運動データの収集により、ユーザの個人差を指摘し改善できるシステムを目指す。

本システムは一般的な家庭や施設での利用を目指し、生活に身近なモバイル端末のみを使用してユーザの運動を識別・計測した。特にスマートフォンは多くの人々が日常的に持ち歩くものなので、ユーザの移動に合わせて自然と持ち運ぶことができる。また、従来のカメラベースの運動計測システムと比較して利用時の環境や服装の制約が少ないため、移動先などでも継続的に運動計測システムを利用できる可能性がある。現状では運動識別のソフトウェアは Python で実装しているため、本稿ではノートパソコンを使用して運動の識別を行っている。顔追跡が可能なスマートフォンやタブレット端末が増えてきている一方で、Python での処理に対応していないものは多い。日常生活での利用を目指し、スマートフォンやタブレット端末上で動作する機械学習手法の実装に取り組む必要がある。

また実際の利用シーンを想定すると、個人での計測にとどまらず、複数人で一緒に取り組むことによる運動モチベーションの増加が期待できる。たとえば、運動計測システムを基にした競争や協力が可能なゲームを作成すれば、複数人で楽しみながら運動に取り組める。市販の運動ゲーム [19] で協力プレイが可能なものも存在するがゲーム端末を保持している必要があり、導入障壁となる恐れもある。これに対し、本システムではすでに保持しているモバイル端末を利用できる利点がある。本システムで使用した顔追跡ミドルウェアでは複数の顔を同時計測できない。また、ユーザがカメラから離れすぎると顔追跡が時折外れてしまう。一方で、4.2 節より本システムによる運動識別がユーザのカメラに対する位置に依存しないことが分かっている

ため、同時に複数の顔を追跡できるシステムを実装することで複数人の同時運動計測が実現できる。また複数のモバイル端末間で通信できる仕組みを導入することで、離れた場所でも複数人で同時に計測できる可能性がある。このように、身近なモバイル端末を利用した複数人プレイ可能な運動ゲームの実現が期待できる。

## 7. おわりに

本稿では顔画像情報を利用した複数の運動の識別方法を提案し、それに対する検証を行った。運動時のユーザの顔画像をモバイル端末の内蔵カメラより取得し、顔追跡ミドルウェア [20] により画像から 30 カ所の追跡点を検出した。追跡点より算出した 62 個の特徴量に対し、SVM で識別器を生成し 9 種類の運動を識別した。検証の結果、識別精度は LOO では 97.2%、LOSO では 88.2% となった。加えてシステムの識別速度を向上するためにウィンドウサイズや特徴量削減の検討を行ったところ、最適なウィンドウサイズはユーザの運動周期に近い値になることや 4 個の特徴量による識別器を使用した場合でも 87.1% の精度で識別できることが判明した。また運動時のモバイル端末とユーザの位置関係の変化にも頑健であることが確認できた。今後は、複数人を同時計測するための顔追跡システムの実装や運動ペースや運動強度の推定などを行う。

謝辞 本研究は、JST PRESTO 課題番号 JPMJPR17J4、および、JST AIP-PRISM 課題番号 JPMJCR18Y2 の支援を受けたものです。

## 参考文献

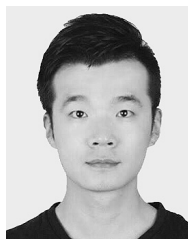
- [1] Al-Mashaqbeh, G.A.: Computers and e-Health: Roles and new applications, *2012 International Conference on Computer Systems and Industrial Informatics*, pp.1–6 (online), DOI: 10.1109/ICCSII.2012.6454498 (2012).
- [2] Ar, I. and Akgul, Y.: A Computerized Recognition System for the Home-Based Physiotherapy Exercises Using an RGBD Camera, *IEEE Trans. Neural Systems and Rehabilitation Engineering*, Vol.22 (online), DOI: 10.1109/TNSRE.2014.2326254 (2014).
- [3] Colberg, S., Sigal, R., Yardley, J., Riddell, M., Dunstan, D., Dempsey, P., Horton, E., Castorino, K. and Tate, D.: Physical Activity/Exercise and Diabetes: A Position Statement of the American Diabetes Association, *Diabetes Care*, Vol.39, pp.2065–2079 (online), DOI: 10.2337/dc16-1728 (2016).
- [4] Crema, C., Depari, A., Flammmini, A., Sisinni, E., Haslwanter, T. and Salzmann, S.: IMU-based solution for automatic detection and classification of exercises in the fitness scenario, pp.1–6 (online), DOI: 10.1109/SAS.2017.7894068 (2017).
- [5] Depari, A., Ferrari, P., Flammmini, A., Rinaldi, S. and Sisinni, E.: Lightweight Machine Learning-Based Approach for Supervision of Fitness Workout, pp.1–6 (online), DOI: 10.1109/SAS.2019.8706106 (2019).
- [6] Deponti, D. and Maggiorini, D.: DroidGlove: An android-based application for wrist rehabilitation, pp.1–7 (online), DOI: 10.1109/ICUMT.2009.5345442 (2009).

- [7] Duzel, E., Praag, H. and Sendtner, M.: Can physical exercise in old age improve memory and hippocampal function?, *Brain*, Vol.139 (online), DOI: 10.1093/brain/awv407 (2016).
- [8] Farrance, C., Tsofliou, F. and Clark, C.: Adherence to community based group exercise interventions for older people: A mixed-methods systematic review, *Preventive Medicine*, Vol.87, pp.155–166 (online), DOI: 10.1016/j.ypmed.2016.02.037 (2016).
- [9] Gandomkar, M., Sarang, R. and Gandomkar, Z.: TrainingPal: An Algorithm for Recognition and Counting Popular Exercises Using Smartphone Sensors, *ICEE 2018*, pp.1471–1476 (online), DOI: 10.1109/ICEE.2018.8472444 (2018).
- [10] HarvardHealthPublishing: The 4 most important types of exercise - Harvard Health, available from (<http://www.health.harvard.edu/exercise-and-fitness/the-4-most-important-types-of-exercise>) (accessed 2020-09-21).
- [11] Hashizume, H., Yoshimura, N., Nagata, K., Miyazaki, N., Ishimoto, Y., Nishiyama, R., Oka, H., Yamada, H. and Yoshida, M.: Development and evaluation of a video exercise program for locomotive syndrome in the elderly, *Modern Rheumatology*, Vol.24, No.2, pp.250–257 (online), DOI: 10.3109/14397595.2013.854063 (2014).
- [12] KaiaHealth: Digital Therapeutics to Manage Chronic Diseases, available from (<https://www.kaihealth.com/>) (accessed 2020-07-17).
- [13] Khurana, R., Ahuja, K., Yu, Z., Mankoff, J., Harrison, C. and Goel, M.: GymCam: Detecting, Recognizing and Tracking Simultaneous Exercises in Unconstrained Scenes, *Proc. ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol.2, pp.1–17 (online), DOI: 10.1145/3287063 (2018).
- [14] NationalStatistics: Health state life expectancies, UK - Office for National Statistics, available from (<http://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/bulletins/healthstatelifeexpectanciesuk/2014to2016>) (accessed 2020-07-16).
- [15] Ockendon, M. and Gilbert, R.: Validation of a Novel Smartphone Accelerometer-Based Knee Goniometer, *The Journal of Knee Surgery*, Vol.25, pp.341–346 (online), DOI: 10.1055/s-0031-1299669 (2012).
- [16] Pernek, I., Kurillo, G., Stiglic, G. and Bajcsy, R.: Recognizing the Intensity of Strength Training Exercises with Wearable Sensors, *Journal of Biomedical Informatics*, Vol.58 (online), DOI: 10.1016/j.jbi.2015.09.020 (2015).
- [17] Prabhu, G., Ahmadi, A., Noel O' Connor and Moran, K.: Activity Recognition of Local Muscular Endurance (LME) Exercises Using an Inertial Sensor, pp.35–47 (online), DOI: 10.1007/978-3-319-67846-7\_4 (2018).
- [18] Stubbs, B., Koyanagi, A., Hallgren, M., Firth, J., Richards, J., Schuch, F., Rosenbaum, S., Mugisha, J., Veronese, N., Lahti, J. and Vancampfort, D.: Physical activity and anxiety: A perspective from the World Health Survey, *Journal of Affective Disorders*, Vol.208 (online), DOI: 10.1016/j.jad.2016.10.028 (2016).
- [19] Ubisoft: Just Dance 2020, available from (<https://www.ubisoft.co.jp/jd2020/>) (accessed 2020-01-05).
- [20] ULSeeInc.: Single Face Tracker Plugin (Lite Version - 30 Face Tracking Points) - Asset Store, available from (<http://assetstore.unity.com/packages/tools/integration/single-face-tracker-plugin-lite-version-30-face-tracking-points-90212>) (accessed 2020-01-13).
- [21] UnitedNations: World Population Prospects Population Division, available from (<http://population.un.org/wpp/Graphs/DemographicProfiles/Line/900>) (accessed 2020-07-16).
- [22] VAYSports: Real-Time Fitness Coaching App, available from (<https://www.vay-sports.com/>) (accessed 2020-07-17).
- [23] VimoLabs: Motion Gesture Recognition Technology, available from (<http://www.vimo.co/>) (accessed 2020-07-17).
- [24] Vox, J. and Wallhoff, F.: Recognition of human motion exercises using skeleton data and SVM for rehabilitative purposes, pp.266–269 (online), DOI: 10.1109/LSC.2017.8268194 (2017).
- [25] WorldHealthOrganization: WORLD HEALTH STATISTICS 2020, available from (<http://apps.who.int/iris/bitstream/handle/10665/332070/9789240005105-eng.pdf?ua=1>) (accessed 2020-07-16).
- [26] Zhang, H. and Ye, C.: RGB-D Camera Based Walking Pattern Recognition by Support Vector Machines for a Smart Rollator, *International Journal of Intelligent Robotics and Applications*, Vol.1 (online), DOI: 10.1007/s41315-016-0002-6 (2017).
- [27] Zulkifley, M., Mohamed, N. and Zulkifley, N.: Squat Angle Assessment Through Tracking Body Movements, *IEEE Access*, Vol.7, pp.48635–48644 (online), DOI: 10.1109/ACCESS.2019.2910297 (2019).
- [28] 稲葉このみ, 清野 諭, 遠峰結衣, 杉浦裕太: タブレット端末内蔵カメラによる踵上げ動作トレーニング支援, エンタテインメントコンピューティングシンポジウム 2019 論文集, Vol.2019, pp.393–397 (2019) (オンライン), 入手先 (<http://id.nii.ac.jp/1001/00199381/>).
- [29] 国立健康・栄養研究所: 寿命について, 入手先 ([http://www.nibiohn.go.jp/eiken/kenkounippon21/download\\_files/other/document\\_en.01.pdf](http://www.nibiohn.go.jp/eiken/kenkounippon21/download_files/other/document_en.01.pdf)) (参照 2020-07-16).



**加藤 花歩**

2019 年慶應義塾大学理工学部情報工学科卒業。2021 年同大学大学院修士課程を修了。



**夏 成碩**

2016 年中国の西安理工大学卒業。2018 年と 2019 年にイギリスのラフバラー大学大学院と中国の西安电子科技大学大学院にて修士課程修了。現在、慶應義塾大学大学院にて博士課程に在籍。ヒューマンコンピュータインタラクションとウェアラブルセンサアプリケーションに関する研究に従事。



杉浦 裕太 (正会員)

2013年慶應義塾大学大学院メディアデザイン研究科博士後期課程修了。博士(メディアデザイン学)。同大学特任助教、産業技術総合研究所特別研究員等を経て、2016年慶應義塾大学工学部情報工学科助教に就任。2018年

同専任講師、2020年同准教授。ヒューマンコンピュータインタラクションに関する研究に従事。