

ウェアラブルカメラを用いた人間の歩行状態推定

A Method of Human Gait Estimation Using a Wearable Camera

畑中 哲生† 渡辺 義浩† 小室 孝† 石川 正俊†

Tetsuo Hatanaka†, Yoshihiro Watanabe†, Takashi Komuro†, Masatoshi Ishikawa†

†東京大学 情報理工学系研究科

†Graduate School of Information Science and Technology, University of Tokyo

E-mail: Tetsuo.Hatanaka@ipc.i.u-tokyo.ac.jp

Abstract

実環境下における人間の位置・姿勢推定に関して、様々な研究が行われてきた。これまでの研究では、GPSによる位置のみの推定や、設置カメラによる姿勢推定が行われていた。しかし、このような手法は、屋内や、障害物がある環境下、群衆中において適用することが困難であると考えられる。そこで、本論文では、ウェアラブルカメラによって、周囲の環境に影響されず、歩行中の人間の位置と姿勢を取得する新たな計測システムを提案する。本論文では、センシングシステムの構成、歩行表現とその解析手法、階層的な方向推定手法の3つを提案する。また、実験によって、人間に装着された単眼カメラからの観測画像の変動から、歩行中の全身の動作が推定可能であることを示した。

1 はじめに

近年、日常生活の中で、いつでもどこでも、人間の位置・姿勢を推定可能なシステムのニーズが高まっている。あらゆる場所で、障害物の有無に左右されない位置・姿勢推定が実現されることで、スポーツ選手の運動解析や、見た目に自然なCGの作成、個人の健康状態の詳細把握、屋外・屋内での経路案内、集団・群衆行動の統制に向けた最適な行動支援などの、様々なアプリケーションが実行可能になる。また、人間とロボットが共生することが見込める将来の社会では、ロボットが人間の位置・姿勢を常に認識することで、屋内・屋外など、様々な環境状況下でサービスを提供することが可能になると期待できる。

これまでに行われてきた大域的な位置推定には Global Positioning System(GPS)があるが、電波の届かない地下などにおいては利用が困難である。また、同技術は、受信機が複数の衛星からの電波を受け取ることで位置の推定を行うが、僅かな計測の誤差から大きな推定誤差が生じてしまうという欠点を持つ。そのため、GPS

の受信機を複数を使用した姿勢推定は困難であると考えられる。詳細な姿勢推定に関しては、モーションキャプチャ[1]があるが、煩雑なマーカーを身体に取り付ける必要があり、手軽なシステムとは言えない。この他、マーカーを用いない設置カメラによる監視型のアプローチ[2]が考えられるが、障害物がある環境下や群衆中では、死角の影響により情報の取得が困難である。このように、従来のシステムは、いつでもどこでも人体の位置・姿勢を推定するには不向きであった。

そこで本論文では、このような問題が致命的となる新しいアプリケーションのための人間の位置・姿勢推定に着目する。

2 先行研究

様々な環境におけるモーションキャプチャの需要が高まり、それに伴って、設置カメラを必要としないシステムの研究が行われている。

Vlasicらは、日常生活におけるモーションキャプチャの重要性に着目し、加速度センサ、ジャイロスコープ、超音波発生装置、マイクロフォンを使用したモーションキャプチャシステムを構築した[3]。これは、加速度センサ、ジャイロセンサの出力と、超音波のTime-of-flightを計測することによって、人間の姿勢を推定する原理のものである。このシステムの実証として、サイクリングやスキー、卓球をしている人間の姿勢を推定することが可能であることが示されている。しかし、このシステムは数十個ものセンサを体に取り付ける必要があり、煩雑なシステムとなっている。また、人間の姿勢を推定することに重点を置いており、大域的な位置推定は行われていない。さらに、超音波を用いるため、多人数での同時使用には限界があると考えられる。

濱口らは、磁気センサと足裏にマイクロスイッチを設置した靴を使用して、人間の歩行中の位置推定を行った[4]。このシステムでは、磁気センサを使用して、腰の絶対位置と両足の腰からの相対位置を計測する。歩行中に両足が接地すると、両足の足裏に取り付けられ

ているマイクロスイッチが押され、その時に計測されている腰と両足の位置が時々刻々加算され、歩行軌道が生成される。このシステムでは、センサ数は少ないが、両足が地面に付いている時の移動距離を積分して位置推定をする手法を採用しているため、累積誤差の影響が避けられない。また、走行時は両足が同時に接地する瞬間が無いため、このシステムでは位置推定が困難である。

Moletらは、磁気センサを使用し、人間の関節の角度を測ることによってモーションキャプチャを行うシステムを提案した [5]。このシステムは、ユーザは全身に14個の磁気センサを取り付け、それぞれのセンサが得る位置・角度の情報を基に、人間の動作を推定するものである。この手法では、磁場発生装置の付近でしか正しい推定を行うことができず、また、全身にセンサを取り付けるため、煩雑なシステムとなっている。

3 着目する課題と目的

これまでの設置カメラを必要としないモーションキャプチャシステムでは、ユーザの全身にセンサが取り付けられており、煩雑なシステムとなっていた。また、ウェアラブルカメラを用いた研究は、人間の位置のみが推定され、姿勢は推定されていなかった。そこで、本論文では、ウェアラブルカメラを1つだけ取り付けた構成で、歩行中の人間の位置と姿勢を推定する技術の実現を目指す。

本論文は、以下の3つの要素から構成されている。ウェアラブルカメラを用いたセンシングシステムの提案、システムが保有する歩行データと歩行形態分類、歩行状態推定アルゴリズムである。本論文では、歩行中の位置・姿勢を歩行状態と呼ぶ。

本システムは、人間の動作によって生じる変動情報をウェアラブルカメラで捉え、その情報からオリジナルの動作を復元するアプローチを導入した技術である。また、人間の歩行の特徴的な成分を抽出した、パラメトリック表現に基づくコンパクトな歩行データによって、人間の歩行を記述し、歩行形態の分類を行う。さらに、1つのウェアラブルカメラの画像という限られた情報から歩行状態を推定する技術の確立を目指す。本手法では、歩行形態の切り替え推定手法と、Extended Kalman Filter(EKF)を用いた歩行状態推定手法を使用する。

4 ウェアラブルカメラを用いたセンシングシステム

4.1 システムの提案

本論文では、ウェアラブルカメラによって撮影された画像と、事前知識として保持される人間の歩行データによって、センサの限られた情報から歩行状態を推

定する技術の確立を目指す。本システムでは、図1のように、ウェアラブルカメラを1台のみ、ユーザの大腿部に下向きに装着する構成をとる。この状態で歩行することで、撮影された画像から得られる歩行動作によるシーンの変動情報と、人間の歩行データを併用することにより、歩行状態の推定が可能になると考えられる。



図1 システムの概観

4.2 本システムの実現に向けた要件

本論文で提案するシステムに使用するウェアラブルカメラには、他のセンサに比べて、複数の利点を備えている。加速度センサなどは、位置推定の累積誤差が生じるが、カメラを使用することで、大域的な位置推定が可能である。これは、カメラが周囲の環境情報を元に、自己位置の修正を行うことが可能であるためである。さらに、近年、センサの小型化、低コスト化が進み、広く普及している携帯電話にも、カメラや加速度センサなどが組み込まれており、このようなセンサを身につけることが現実的となっている。高機能な情報取得を携帯機器によって取得することが可能となった背景からも、このようなアプローチは有効なものと考えられる。

5 簡易的歩行データを用いた歩行形態表現

5.1 歩行形態表現の課題

ロボティクスやコンピュータグラフィックによるアニメーションの分野で、人間の歩行に関して、多くの表現方法が提案されてきた。人間の歩行を表すモデルとして広く利用されているシンプルなものには倒立振り子モデルが挙げられる [6]。このモデルでは、人間の重心位置を (x, y, z) とし、重心が水平に移動すると仮定し、ダイナミクスを以下のように表す。

$$\ddot{x} = \frac{g}{z}x \quad (1)$$

$$\ddot{y} = \frac{g}{z}y \quad (2)$$

ここで、 z 軸は重力方向であり、人間は xy 平面上を歩行しているとする。 g は重力加速度である。このような物理学的見地に基づいた歩行モデルは、主にロボットの歩行パターン生成などに用いられており、人間の独特な動きに対応させることが困難であると考えられる。

一方、アニメーションの分野では、観測データから人間の歩行を表現する方法が主流である [7]。特に、限られたデータから、類似する行動パターンを生成するなどの手法が発達しており、有用性が高い。同表現は、状態推定のための動作表現としても有効である見込みが高い。そこで、本手法では観測によって得たユーザの歩行データから、直接歩行姿勢を記述する表現を採用する。

5.2 歩行データのパラメトリック表現の作成

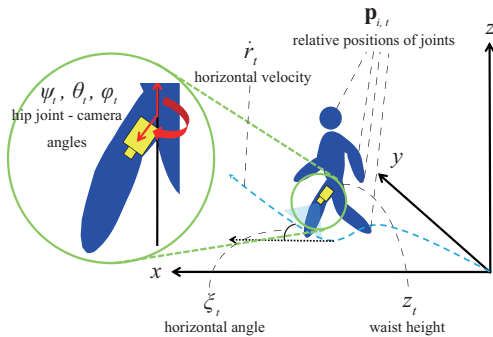


図 2 システムのパラメータ定義

本節では、本手法で使用する歩行データのパラメトリック表現の作成方法について説明する。本手法では、一つの歩行データは、歩行者の腰の高さ z_t 、カメラのロール・ピッチ・ヨー角 $(\psi_t, \theta_t, \phi_t)$ 、腰の進行方向の速度 r_t 、腰の進行方向の角速度 $\dot{\xi}_t$ 、体の各関節の腰からの相対座標 $\mathbf{p}_{i,t} = (p_x, p_y, p_z)^T_{i,t}$ によって記述される。ここで t は時刻を表し、 i は各関節のインデックスを表す。これらのパラメータの定義を図 2 に示す。

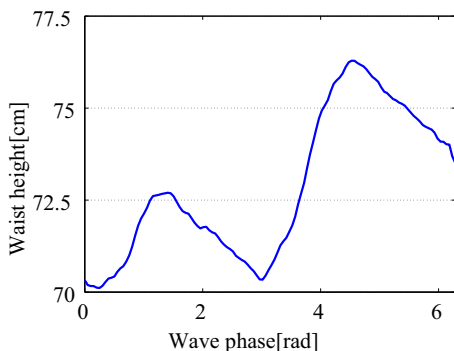


図 3 腰の高さ z_t

例として、ユーザの腰の高さ z_t の 1 周期分の観測デー

タを図 3 に示す。しかし、計測された歩行データはノイズによる高周波成分を含んでおり、流動的な歩行を表す表現としては適当であるとは言えない。そこで、これらの歩行データを低周波成分のみから表すものとした。人間の歩行に関する先行研究 [8] によると、歩行動作における各関節の運動は第 2 次までの低い周波数成分からなる正弦波だけで再現できることが示されている。本論文では、フーリエ級数展開を用いて、式 (3) のように、第 5 次までの周波数成分で歩行データを記述する。

$$q_t = \sum_{n=0}^5 a_{n,q} \sin(n\eta_t + b_{n,q})$$

for all $q(q = z, \psi, \theta, \phi, \dot{r}, \dot{\xi}, p_{xi}, p_{yi}, p_{zi})$ (3)

η_t は時刻 t での歩行の位相である。式中の全 $a_{n,q}, b_{n,q}$ からなるベクトル $\mathbf{f} = (a_{n,q}, b_{n,q})$ をパラメトリック表現された歩行データとして登録する。これ以降、歩行データとは、このパラメトリック表現された歩行データを意味する。

時刻毎に歩行の位相と、歩行者の腰部の位置 $\mathbf{r}_t = (x_t, y_t, z_t)^T$ 、進行方向の角度 ξ_t から、下式に示されるように歩行状態が決定できる。

$$\mathbf{r}_{i,t} = \mathbf{r}_t + \mathbf{R}(0, 0, \xi_t) \mathbf{p}_{i,t} \quad (4)$$

ここで、 $\mathbf{R}(0, 0, \xi_t)$ は引数である進行方向の角度に関する回転行列である。

5.3 歩行形態の分類

本節では、人間の歩行を複数の歩行形態に分類する手法について述べる。人間の歩行には、「ゆっくり歩く」、「歩く」、「走る」、「曲がる」、「坂道を登る」、「階段を降りる」など様々な歩行形態がある。しかし、人間は通常、歩行速度の変化に伴う歩行形態の遷移時を除いては、有限個の歩行形態で歩いていると考えられる。従って、前節で述べた歩行データによって各歩行形態を表すことができる。そこで、本論文では、有限個の歩行データで人間の歩行形態を表し、どの歩行形態で歩行しているかを、時刻ごとに推定することによって、歩行状態推定の精度の向上を目指す。

また、各歩行形態間の接続関係を記述した歩行形態遷移図を作成する。図 4 に歩行形態遷移図の一例を示す。歩行形態遷移図を導入することで、ある時刻で推定された歩行形態から、行動上繋がりが得ない歩行形態が次時刻で推定されてしまうことを防ぐことが可能となる。また、一時刻前に推定された歩行形態から変化する歩行形態の歩行データだけを、次の時刻での推定に使用する手法を採用することで、計算時間を削減

できるという点でも有用性が高い．具体的な歩行形態の分類と，その遷移関係の生成手法を下記に述べる．

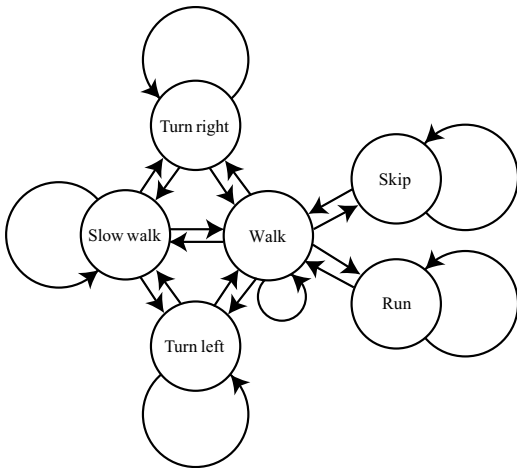


図 4 歩行形態遷移図

歩行データの収集

まず、「ゆっくり歩く」、「走る」などの様々な歩行形態の1周期(2歩)分の生の歩行データを準備する．これらの歩行データを，3.1節で説明したパラメトリックな表現を用いて， $\mathbf{f}_i (i = 1, 2, \dots, N)$ と記述する．ここで， N は準備した全歩行データの数である．

歩行データの特徴量空間への射影

収集した歩行データ \mathbf{f}_i は，歩行一周分分の身体動作の周波数成分情報を有する．全歩行データ $(\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N)$ について，主成分分析 [9] を行う．1つの歩行データを，特徴量空間上の一点 $\mathbf{f}_{\text{PCA},i} (i = 1, 2, \dots, N)$ で表す．

$$\mathbf{f}_{\text{PCA},i} = \mathbf{U}(\mathbf{f}_i - \bar{\mathbf{f}}) \quad (5)$$

ここで， \mathbf{U} は主成分分析に使用する正規直交基底であり， $\bar{\mathbf{f}}$ は全歩行データの周波数成分の平均ベクトルである．

クラスタリングを用いた歩行形態の分類

特徴量空間上で距離が近い2つの歩行データは，同じ歩行形態か，似通った歩行形態であると考えられることができる．そこで，これらの歩行データに関して k-means クラスタリング [10] を行い，複数のクラスタに分割する．各クラスタのセントロイドに最も近い点を，当クラスタ内の代表歩行データとし，1つの歩行形態として登録する．

歩行形態の接続

分類された歩行形態の内，連続して繋がり得る歩行形態同士は接続される必要がある．使用した歩行データが時系列データであった場合は，時系列に沿って入力された2つの歩行データが属する歩行形態同士を接続する．また，経験的知識を用いて接続を作成することも可能である．例えば、「ゆっくり歩く」と「歩く」は

接続し得るが、「ゆっくり歩く」と「走る」は接続しない，と決定づけることが可能である．本論文では，後者の方法で，各歩行形態を接続した．

6 歩行形態の切り替えと Extended Kalman Filter を用いた歩行状態推定手法

6.1 歩行状態推定手法の概要

本節では，歩行状態推定の手法の概要について述べる．各時刻で，歩行データの位相 η_t と，歩行者の腰部の位置 $\mathbf{r}_t = (x_t, y_t, z_t)^T$ ，進行方向の角度 ξ_t が定めれば，式 (4) に示されるように歩行状態が決定できる．そこで，時刻 t での人間の歩行状態を次式で再定義する．

$$\boldsymbol{\mu}_t = \begin{pmatrix} x_t \\ y_t \\ \xi_t \\ \eta_t \\ \dot{\eta}_t \end{pmatrix} \quad (6)$$

ここで， $\dot{\eta}_t$ は現在の歩行データの位相遷移速度を表す．

提案手法では，この歩行状態を毎時刻推定し続けるために，「歩行状態仮説の生成」，「歩行状態の予測」，「歩行形態の尤度計算」，「歩行状態のリサンプリング」，「歩行状態の更新」の5ステップを繰り返す．図5にその概要を示す．提案手法では，歩行状態 $\boldsymbol{\mu}_t$ の仮説を複数用意し，それぞれの仮説に対してこの5ステップを適用する．図の w_t は現在の歩行状態仮説の尤度を表す．毎時刻の推定ではこの尤度が閾値より大きい幾つかの仮説が残り，次時刻の推定に使用される．

4.2節では歩行状態推定に使用する5ステップの概要を述べ，4.3節では，「歩行状態の予測」，「歩行形態の尤度計算」，「歩行状態の更新」の具体的な計算方法について述べる．

6.2 歩行状態推定手法の手続き

ある歩行状態の仮説が，前の時刻に歩行データ $\mathbf{f}_i (i = 1, 2, \dots, n_f)$ で動作しており，仮説のインデックスを $m (m = 1, 2, \dots, n_c)$ とすると，時刻 t における歩行状態の仮説は $(\boldsymbol{\mu}_{i,m,t}, w_{i,m,t})$ と書くことができる．ここで， $w_{i,m,t}$ はこの仮説の尤度である．推定では，5ステップの手続きを行う．各手続きの詳細を下記に示す．

クラスタリングを用いた歩行形態の分類

図5の緑のボックスで示される Hypothesis creation では，歩行形態遷移図に基づいて，仮説 $(\boldsymbol{\mu}_{i,m,t}, w_{i,m,t})$ の歩行形態(歩行データ \mathbf{f}_i) から遷移し得る全ての歩行形態に歩行状態の仮説を生成する．ここでは遷移し得る歩行形態の歩行データを $\mathbf{f}_i, \mathbf{f}_j, \dots, \mathbf{f}_k$ とする．

歩行状態の予測

図5の Predict では， $(\boldsymbol{\mu}_{i,m,t-1}, w_{i,m,t-1})$ から各歩行データによって予測された仮説 $(\bar{\boldsymbol{\mu}}_{i,m,t}, w_{i,m,t}), (\bar{\boldsymbol{\mu}}_{j,m,t}, w_{j,m,t}), \dots, (\bar{\boldsymbol{\mu}}_{k,m,t}, w_{k,m,t})$ を

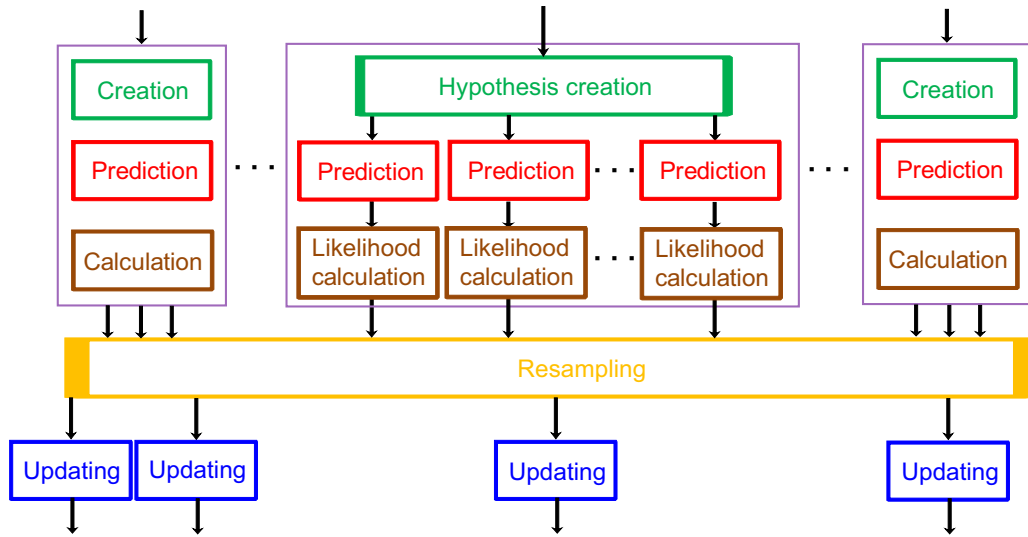


図5 歩行状態推定手法の概要

それぞれ生成する．具体的な計算については6.3節で述べる．

歩行形態の尤度計算

生成された各仮説に対して，その仮説が選択している歩行形態の尤度 $(\lambda_{i,m,t}, \lambda_{j,m,t}, \dots, \lambda_{k,m,t})$ を求める．尤度の具体的な計算方法については6.3節で述べる．計算した歩行形態の尤度と，各仮説の尤度を掛け合わせて，仮説の尤度を更新する．ここで，次の時刻では前の時刻で仮説が使用していた歩行形態と同じ形態で動作する確率が高くなると考えられるため，使用する動作形態が前の時刻と同じ場合は，さらに定数 C を掛け合わせる．

$$w_{i,m,t} = w_{i,m,t-1} \times \lambda_{i,m,t} \times C \quad (7)$$

$$w_{\alpha,m,t} = w_{\alpha,m,t-1} \times \lambda_{\alpha,m,t} \quad (8)$$

for all $\alpha (\alpha = j, \dots, k)$

歩行状態のリサンプリング

仮説の尤度 $w_{i,m,t}$ が閾値 w_{TH} 以上の仮説を残し，他の仮説を切り捨てる．尚，このときに尤度が最も大きい仮説の歩行形態を，現在の歩行形態として選択する．

歩行状態の更新

歩行形態の尤度に従って，残った各仮説の歩行状態を更新する．本論文では，EKFを用いた．具体的な計算方法は6.3節で述べる．残った仮説に対してこの5ステップを繰り返し適用する．

6.3 歩行状態推定における計算の詳細

本節では，観測される外界のワールド座標が全て既知である状況下において，歩行状態を推定する手法を提案する．また，外界のワールド座標が未知である場合でも，Simultaneous Localization and Mapping(SLAM) [11] による手法の拡張が可能であり，本論文で提案する手法はそのまま適用できると考えられる．

本手法は，図5のHypothesis creation部分で使用した各歩行形態 i と，Resampling部分で選ばれた各仮説 m に対して共通の処理であるため，各歩行データと各仮説に関する添え字 i, m の表記は省略する．

本システムの状態方程式と観測方程式は下式で表される．

$$\tilde{\boldsymbol{\mu}}_t = \mathbf{g}(\boldsymbol{\mu}_{t-1}) + \boldsymbol{\epsilon}_t \quad (9)$$

$$\mathbf{u}_t = \mathbf{h}(\tilde{\boldsymbol{\mu}}_t) + \boldsymbol{\delta}_t \quad (10)$$

ここで， $\tilde{\boldsymbol{\mu}}_t$ は時刻 t での実際の人間の歩行状態であり， \mathbf{u}_t は観測値である．また， $\boldsymbol{\epsilon}_t, \boldsymbol{\delta}_t$ はそれぞれプロセスノイズと観測ノイズである． \mathbf{g} は動作モデルであり，時刻 $t-1$ の歩行状態 $\boldsymbol{\mu}_{t-1}$ が歩行データに基づいて次の時刻 t にどのように遷移するかを表す．

$$\begin{aligned} \bar{\boldsymbol{\mu}}_t &= \mathbf{g}(\boldsymbol{\mu}_{t-1}) \\ &= \begin{pmatrix} \bar{x}_t \\ \bar{y}_t \\ \bar{\xi}_t \\ \bar{\eta}_t \\ \bar{\dot{\eta}}_t \end{pmatrix} \\ &= \begin{pmatrix} x_{t-1} + \dot{r}(\eta_{t-1}) \cos \xi_{t-1} \\ y_{t-1} + \dot{r}(\eta_{t-1}) \sin \xi_{t-1} \\ \xi_{t-1} + \dot{\xi}(\eta_{t-1}) \\ \eta_{t-1} + \dot{\eta}_{t-1} \\ \dot{\eta}_{t-1} \end{pmatrix} \quad (11) \end{aligned}$$

ここで， $\bar{\boldsymbol{\mu}}_t$ は，プロセスノイズがのらず，動作モデ

ルのみで予測された歩行状態を表す。\$r, \xi\$ は歩行データ \$\mathbf{f}\$ と、時刻 \$t-1\$ での歩行の位相、\$\eta_{t-1}\$ から算出される。

また、\$\mathbf{h}\$ は時刻 \$t\$ で予測された歩行状態 \$\bar{\boldsymbol{\mu}}_t\$ に対して、観測される画像上の特徴点の位置 \$\mathbf{u}_t\$ を算出する観測モデルである。ユーザの腰の予測位置を \$\mathbf{r}_t = (\bar{x}_t, \bar{y}_t, z(\bar{\eta}_t))\$、ウェアラブルカメラの3次元位置を \$\mathbf{r}_{ct} = (x_{ct}, y_{ct}, z_{ct})^T\$、地面の \$i\$ 番目の特徴点を \$\mathbf{X}_i = (X_i, Y_i, 0, 1)^T\$、\$i\$ 番目の特徴点のカメラ座標を \$\mathbf{c}_{i,t} = (c_{i,x,t}, c_{i,y,t}, c_{i,z,t})^T\$、腰部とカメラの距離を \$l\$ とすると \$\mathbf{h}\$ は以下のように表わせる。

$$\mathbf{u}_t = \mathbf{h}(\bar{\boldsymbol{\mu}}_t) = \mathbf{S} \begin{pmatrix} c_{i,x,t} & c_{i,y,t} \\ c_{i,z,t} & c_{i,z,t} \end{pmatrix}^T + \mathbf{d} \quad (12)$$

$$\mathbf{c}_{i,t} = (\mathbf{R}(\psi(\bar{\eta}_t), \theta(\bar{\eta}_t), \phi(\bar{\eta}_t)), \quad (13)$$

$$-\mathbf{R}(\psi(\bar{\eta}_t), \theta(\bar{\eta}_t), \phi(\bar{\eta}_t))\mathbf{r}_{ct})\mathbf{X}_i \quad (14)$$

$$\mathbf{r}_{ct} = \mathbf{r}_t - \mathbf{R}(0, 0, \xi_t)\mathbf{R}(\psi(\bar{\eta}_t), \theta(\bar{\eta}_t), \phi(\bar{\eta}_t))\mathbf{l} \quad (15)$$

\$\mathbf{l} = (0, 0, l)^T\$ である。\$\mathbf{S}\$ はカメラの焦点距離を含むパラメータである。また、\$\mathbf{R}(0, 0, \xi_t)\$ は進行方向に関する回転行列であり、\$\mathbf{R}(\psi(\bar{\eta}_t), \theta(\bar{\eta}_t), \phi(\bar{\eta}_t))\$ はワールド座標からカメラ座標への回転行列である。

ここで、実際の撮像画像の特徴点 \$\mathbf{u}_t\$ と、予測されている歩行状態から導かれる特徴点 \$\mathbf{h}(\bar{\boldsymbol{\mu}}_t)\$ の差分が小さければ小さいほど、現在選択されている歩行形態の尤もらしさが高くなる。ここでは予測特徴点 \$\mathbf{h}(\bar{\boldsymbol{\mu}}_t)\$ と最も距離の近い画像上の特徴点を、対応する特徴点として扱う。また、予測特徴点付近に画像特徴点が検出されない場合は、尤度が計算できないため、予測に使用された歩行形態の尤度は小さく設定することにした。予測特徴点と画像特徴点との差分は、実際の歩行が各歩行形態の歩行データ通りではないことに起因するもので、ガウス分布で記述することができる。このため、4.2.3節の各歩行形態の尤度は次式のように計算される。

$$\lambda_t = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(|\mathbf{u}_t - \mathbf{h}(\bar{\boldsymbol{\mu}}_t)| - \nu)^2}{2\sigma^2}\right\} \quad (16)$$

\$\nu\$ は差分の平均値であり、\$\sigma^2\$ は分散である。本論文では \$\nu = 0, \sigma^2 = 10\$ に設定した。この歩行形態の尤度 \$\lambda_t\$ と仮説の尤度 \$w_t\$ の積が大きい幾つかの仮説が、次時刻での推定に使用される。

ここで、時刻 \$t\$ における実際の歩行状態 \$\tilde{\boldsymbol{\mu}}_t\$ と、歩行データによって予測された歩行状態 \$\bar{\boldsymbol{\mu}}_t\$ との差分を解消するために EKF を使用する。本手法における EKF を使用した歩行状態推定は下記ようになる。

$$\bar{\boldsymbol{\mu}}_t = \mathbf{g}(\boldsymbol{\mu}_{t-1}) \quad (17)$$

$$\bar{\boldsymbol{\Sigma}}_t = \mathbf{G}_t\boldsymbol{\Sigma}_{t-1}\mathbf{G}_t^T + \mathbf{P}_t \quad (18)$$

$$\mathbf{K}_t = \bar{\boldsymbol{\Sigma}}_t\mathbf{H}_t^T(\mathbf{H}_t\bar{\boldsymbol{\Sigma}}_t\mathbf{H}_t^T + \mathbf{Q}_t)^{-1} \quad (19)$$

$$\boldsymbol{\mu}_t = \bar{\boldsymbol{\mu}}_t + \mathbf{K}_t(\mathbf{u}_t - \mathbf{h}(\bar{\boldsymbol{\mu}}_t)) \quad (20)$$

$$\boldsymbol{\Sigma}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\bar{\boldsymbol{\Sigma}}_t \quad (21)$$

\$\mathbf{G}_t, \mathbf{H}_t\$ はそれぞれ時刻 \$t\$ における \$\mathbf{g}, \mathbf{h}\$ のヤコビアンである。\$\mathbf{P}_t, \mathbf{Q}_t\$ はそれぞれプロセスノイズと観測ノイズの共分散行列である。\$\boldsymbol{\Sigma}_t\$ はシステムの共分散行列で、時刻 \$t\$ での状態推定の精度を表す。

EKF のアルゴリズムの流れを述べる。式 (17) では、動作モデルに基づいて、1 時刻後の歩行状態を予測している。予測は、実際の歩行では無視することのできないプロセスノイズを考慮せず、動作モデルに忠実に行う。式 (18) では、動作モデルに基づく状態推定の共分散行列を計算している。予測された状態から導かれる時刻 \$t\$ での画像上の特徴点位置は \$\mathbf{h}(\bar{\boldsymbol{\mu}}_t)\$ である。式 (20) では、この予測された特徴点 \$\mathbf{h}(\bar{\boldsymbol{\mu}}_t)\$ とカメラが撮像した実際の特徴点 \$\mathbf{u}_t\$ の差分をとり、式 (19) から算出されるカルマンゲイン \$\mathbf{K}_t\$ を掛けることで、状態変数 \$\boldsymbol{\mu}_t\$ の更新を行っている。式 (21) では、式 (20) での更新による共分散行列を計算している。本手法は、歩行状態を、モデルとのずれを観測量 \$\mathbf{u}_t\$ によって修正しながらロバストに推定するものとなっている。

7 実画像を用いた歩行状態推定の検証

本システムの実用性を検証するため、実画像を用いた検証実験を行った。まず、モーションキャプチャを行い、「ゆっくり歩く」、「ゆっくり左折」、「ゆっくり右折」、「歩く」、「左折」、「右折」の6つの歩行データをそれぞれ3種類ずつ(計18種類)取得した。モーションキャプチャシステムは NaturalPoint 社 OptiTrack を使用した。

各歩行形態を5節の方法で分類したものが図6である。18種類の歩行データが6つの歩行形態に正しく分類できていることがわかる。各クラスに所属する歩行データの内、各セントロイドに最も近い歩行データを、そのクラスの代表歩行データとして登録した。

次に、カメラを脚に装着した状態でチェッカーボードの上を歩行し、システムの入力となる撮像画像系列を取得した。ここで、チェッカーボードの各正方形の1辺の長さは15.15cmであり、格子点のワールド座標は既知とした。本実験で使用したカメラは Basler 社 Pioneer で、フレームレートは250fpsとした。また、撮像画像からの特徴点の検出は Harris 検出器を利用した。本実験では、「ゆっくり歩く」、「ゆっくり左折」、「ゆっくり歩く」の3歩行形態が繋がっている1000フレームの歩行

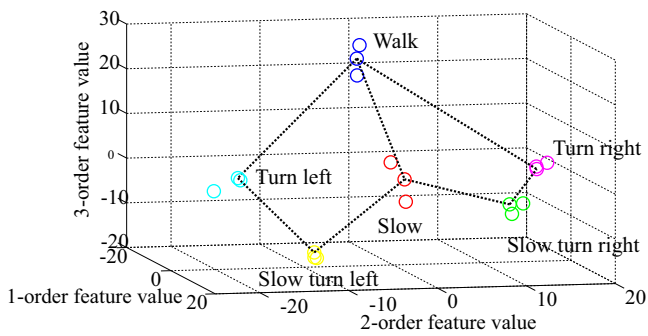


図 6 18 種類の歩行データの分類 (6 つの歩行形態)

を入力歩行として使用した。「ゆっくり歩く」は最初の 2 歩 (320 フレーム), 「ゆっくり左折」は続く 2 歩 (340 フレーム), 2 回目の「ゆっくり歩く」は最後の 2 歩 (340 フレーム) である。

図 7 は，歩行中の写真と歩行状態推定結果の歩行を CG で表したものを比較したものである。両者を比較してみると，推定結果は，入力歩行を再現していると考えられる。

図 8 は 1000 フレームの推定の中に撮影された地面の特徴点の画像座標と，カメラの位置・姿勢から再投影した画像上の特徴点座標との誤差を示すグラフである。推定の間，再投影された特徴点に対応するチェッカーボード上の特徴点が撮像画像中に見つからず，歩行状態推定が正確に行われていない瞬間があった。これは，320 フレーム近辺，及び 910 フレーム近辺で確認された。これは，推定に用いられた全ての歩行形態が実際の動作と合致せず，EKF の動作モデルによる予測歩行状態 $\hat{\mu}_t$ に大きな誤差が乗ってしまったことが原因として考えられる。しかし，この場合以降も，推定が破綻することなく，継続して正常な動作が行われていることが確認された。

また，779 フレーム目で，再投影誤差が最大となった。図 9 に，同フレームにおける特徴点の再投影の様子を表す。図中の青丸はチェッカーボード上に検出された特徴点であり，赤丸は 799 フレーム目のカメラの位置・姿勢から再投影した特徴点である。このフレームでは，チェッカーボード上の特徴点と，再投影した特徴点が平行にずれて現れている。この原因としては，EKF のアルゴリズムの中で，再投影誤差を使用したカメラ位置・姿勢の更新よりも，動作モデルによる歩行状態予測に，システムがより強く追従した可能性が挙げられる。EKF では，パラメータの値により，更新と予測のどちらに比重を置くかを調整することが可能である。しかし，この問題を解決するためにカメラ位置・姿勢の更

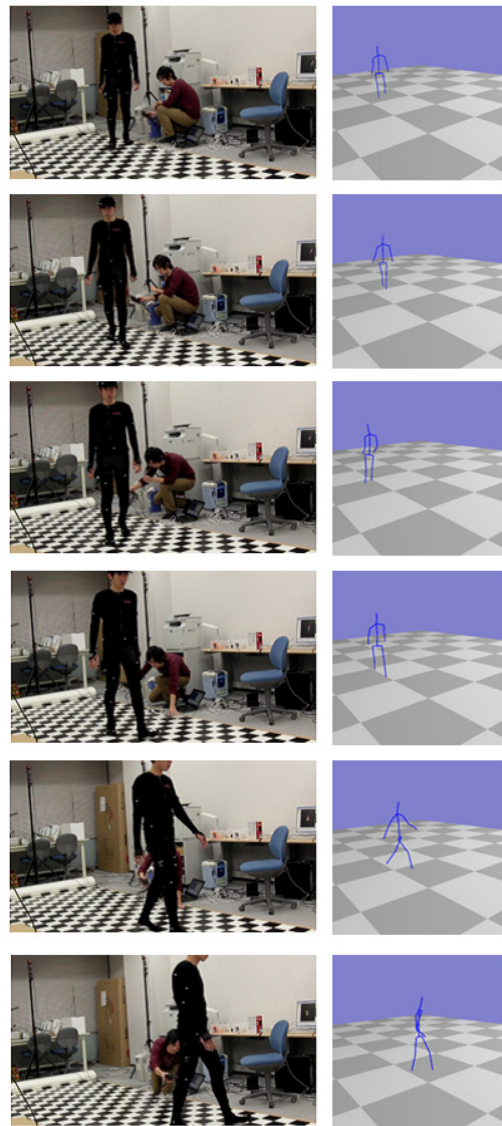


図 7 歩行中の写真 (左) と歩行状態推定結果 (右)

新を優先すると，歩行形態の推定が失敗してしまう恐れがあり，EKF の最適なパラメータ調整は大きな課題である。

8 おわりに

本論文では，ウェアラブルカメラを 1 つだけ装着した構成で，歩行中の新たな人間の位置と姿勢を推定するシステムを提案した。

本システムでは，ウェアラブルカメラの情報を利用して，人間の重要な動作である，「歩行」の状態推定を行った。また，人間の歩行の特徴的な成分を抽出した，パラメトリック表現に基づくコンパクトな歩行データによって，人間の歩行を記述し，歩行形態の分類を行った。さらに，1 つのウェアラブルカメラの画像という限られた情報から歩行状態を推定する手法を提案した。提案した手法は，歩行形態の切り替えと，EKF を用いた

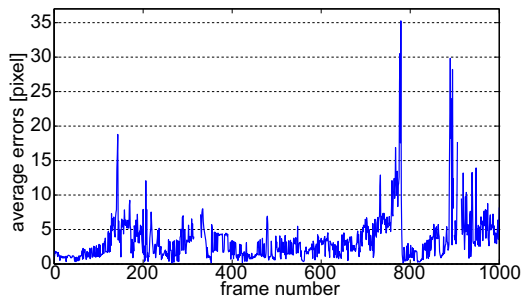


図 8 1000 フレーム間の再投影誤差

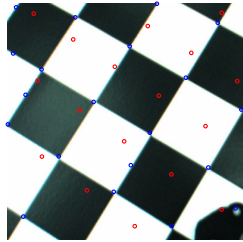


図 9 779 フレーム目における特徴点の再投影

歩行状態推定を同時に行うものである。また、実画像を用いて、地面の特徴点の世界座標が既知である環境下において、歩行状態が推定可能であることを確認した。

本論文では、周期的な歩行について歩行状態推定を行ったが、実際の人間の動作には、歩き始める、立ち止まる、転ぶ、などの非周期的なものがあり、このような動作にも対応できるシステムを構築する必要がある。

また、今回は地面の特徴点の世界座標が既知であるとして手法の提案、実験を行ったが、実際には、そのような環境は稀であり、周囲の特徴点の世界座標が未知の環境下でも歩行状態推定が可能になるようなシステムの実現が望まれる。そのためには、近年盛んに研究が行われている SLAM の枠組みを利用することが有効である。具体的には、本論文で提案した EKF の状態変数の中に、環境中の特徴点の世界座標を時々刻々と登録していくことによって、人間の歩行状態と、特徴点の世界座標の両方を推定することが可能になると考えられる。

参考文献

- [1] Seth Rosenthal and John Pella, "The Process of Motion Capture: Dealing with the Data," *Computer Animation and Simulation*, vol. 97, pp. 3-18, 1997.
- [2] Thomas B. Moeslund and Erik Granum, "A survey of computer vision-based human motion cap-

ture," *Computer Vision and Image Understanding*, vol. 81, no.3, pp. 231-268, 2001.

- [3] Daniel Vlasic, Rolf Adelsberger, Giovanni Vanucci, John Barnwell, Markus Gross, Wojciech Matusik, and Jovan Popovic, "Practical Motion Capture in Everyday Surroundings," *ACM Transactions on Graphics*, vol. 26, no. 3, 2007.
- [4] Akihiro Hamaguchi, Masayuki Kanbara and Naokazu Yokoya, "A Localization Method Using Wearable 3D Position and Orientation Sensors," *The Institute of Electronics, Information and Communication Engineers*, vol. 105, no. 566, pp. 19-24, 2006.
- [5] Tom Molet, Ronan Boulic and Daniel Thalmann, "Human Motion Capture Driven by Orientation Measurements," *Presence: Teleoperators and Virtual Environments*, vol. 8, no. 2, pp. 187-203, MIT Press, 1999.
- [6] 梶田秀司, "ヒューマノイドロボット," オーム社, 2005.
- [7] Munetoshi Unuma, Ken Anjyo and Ryoza Takeuchi, "Fourier Principles for Emotion-based Human Figure Animation," *Proceedings of the 22nd Annual Conference on Computer graphics and interactive techniques*, pp. 91-96, 1995.
- [8] Munetoshi Unuma and Ryoza Takeuchi, "Generation of Human Walking Motion with Emotion for Computer Animation," *The transactions of the Institute of Electronics, Information and Communication Engineers*, vol. J76-D-II, pp. 1822-1831, 1993.
- [9] Svant Word, "Principal Component Analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1, pp. 37-52, 1987.
- [10] Hartigan John A, "A k-means clustering algorithm," *Journal of the Royal Statistical Society*, vol. 28, pp. 100-108, 1979.
- [11] Sebastian Thrun, "Simultaneous Localization and Mapping," *Robotics and Cognitive Approaches to Spatial Mapping*, vol. 38, pp. 13-41, Springer, 2008.