

Boostingを用いたボリュームデータからの人体姿勢推定

中澤 篤志[†] 平 亮介^{††} 竹村 治雄[†]

[†] 大阪大学 サイバーメディアセンター 〒560-0043 豊中市待兼山町1-32

^{††} 大阪大学 情報科学研究科 〒560-0043 豊中市待兼山町1-32

E-mail: [†]{nakazawa,takemura}@cmc.osaka-u.ac.jp, ^{††}taira.ryosuke@ist.osaka-u.ac.jp

あらまし 画像からの人体姿勢推定問題では、関節角の初期パラメータをどう得るかが重要な問題である。この解決法として、人体の姿勢に関するさまざまなサンプル（事例）を準備し、それらと入力データを比較することで解決する手法が提案されている。本論文では、人体のボリュームデータからの姿勢推定をこのアプローチにより解決する手法を提案している。まず学習データとなる姿勢サンプルとして、複数のカメラから復元された人体のボリュームデータを得る。次に、これらを任意のクラスに分類・正規化し、画像特徴である Rectangle Feature を3次元ボリュームデータに拡張した Volumetric Feature を求める。この特徴ベクトル群を Boosting により学習してクラス分類器を作成し、入力データに対して適用することで、姿勢の初期姿勢パラメータを獲得する。得られた姿勢パラメータから収束法等を用いることで、安定かつ正確な姿勢推定を行うことができる。

キーワード 姿勢推定, ボリュームデータ, Volumetric Feature, Boosting

Human pose estimation using volumetric features and boosting approach

Atsushi NAKAZAWA[†], Ryosuke TAIRA^{††}, and Haruo TAKEMURA[†]

[†] Cybermedia Center, Osaka University 1-32 Machikaneyama, Toyonaka, Osaka, 560-0043 Japan

^{††} Cybermedia Center, Osaka University 1-32 Machikaneyama, Toyonaka, Osaka, 560-0043 Japan

E-mail: [†]{nakazawa,takemura}@cmc.osaka-u.ac.jp, ^{††}taira.ryosuke@ist.osaka-u.ac.jp

Abstract Recently, considerable research has been conducted on markerless human-body motion capture using volume data. Most studies have used articulated body models that consist of primitives such as cylinders or ellipsoids. However, the methods that use such models require very good initial parameters. This paper proposes a method to find the human posture parameter without any knowledge of previous frames, just from single frame data. We first prepare sample volume data which takes various postures; and then they are clustered according to the direction of body links. The volumetric feature vector is acquired from every sample and they are learned by AdaBoost algorithm. The resulting classifier is used for estimating input volume data and the parameter of matched class can be used as an initial parameter of tracking based methods.

Key words Posture Estimation, Volume Data, Volumetric Feature, Boosting

1. はじめに

画像からの人体姿勢推定問題は様々な手法が提案されているが、一般には人体を多関節モデルとして近似し、入力画像との相関を表す関数を収束法などで最小化し、多次元の関節パラメータを得る手順が一般的である [1] [2] [3]。しかし、収束法を適用するに当たっては良い初期値が必要であるため、これをどう得るかが大きな問題である。

この解決法として、人体の姿勢に関するさまざまなサンプル（事例）を準備し、それらと入力データを比較する手法が提案されている [4] [5]。人体の姿勢サンプルを画像あるいはモーショ

ンキャプチャなどで取得し、入力データ（画像）と、画像特徴やテンプレートなどで比較し、最も近いものを出力する。基本的に各フレームでの処理は独立しているため、初期値問題を回避することができる。

一方で、この種の手法において問題となるのが、サンプルとの比較コストである。ここでは、以下の2種のコストが問題となる。

- (1) 1サンプル毎の比較（マッチング）コスト
- (2) 大量のサンプル群からの最適探索コスト

前者は、入力データと教師データの比較に関するコストであり、後者は大量にある事例から最適な1つの事例を選び出

すためのコストである。この問題に対し岡田らは、モーションキャプチャデータから生成したすべてのシルエット画像を、Tree-based Filtering によって分類し、階層的に探索する手法を提案している [4]。また、画像のマッチングに対して Chamfer Distance を用いることで、コストの削減とマッチングの柔軟性を確保している。一方で、この手法はシルエット画像を入力とするため、隠蔽等により誤った結果を導くことがある。

本論文では人体のボリュームデータからの姿勢推定問題を解決するために、様々な姿勢サンプルデータをクラスタリングし、学習に基づくクラス識別器を作成することで、入力とする人体ボリュームデータの姿勢推定を行う手法を提案する。ボリュームデータを入力データとすれば、画像（シルエット）を用いるよりもより正確な関節角度を得ることが可能だが、一方でパラメータの自由度が多く、単純な探索では多くの事例および計算コストが必要である。これに対し、入力データを人体姿勢に応じてクラスタリングすることで比較対象数を減らし、また、各クラスのマッチングには矩形特徴 (Rectangle Feature) を 3 次元体積空間に拡張した体積特徴 (Volumetric Feature) を用いた線形分類器を用い、計算コストを削減する。また、この線形分類器の作成は AdaBoosting による学習手法を用いることで、識別の一般性および識別性能を高めている。

2. 関連研究

人体の姿勢推定問題は様々な手法が提案されている。

Mikić らの手法 [6] は、人体の各部位を楕円球や円筒等の CAD 図形で近似し、体全体を多関節モデルとして表す。関節のパラメータを変化させながら、モデルとボリュームデータをフィッティングする。また、胴体に腕が密着した姿勢や両足の密着した状態に対しても、階層的な複数のモデルを用いることで対応できる。しかしこの階層的モデルは木構造をしているため、頭に体が触れた場合や両手を接触させた場合などのループのある姿勢には対応できない。またボリュームデータとモデルのフィッティングには大きな計算コストがかかり、追跡に失敗した場合の復帰が難しい。Caillette らの手法 [7] は、ボリュームデータを色を基準に blob 群に近似し、それらを木構造を持つ人体モデルにフィッティングし姿勢を推定する。特に体と四肢が接近する姿勢において、色情報を用いることによって正しく姿勢推定を行えることが示されている。一方で前者と同様、モデルの運動自由度が大きいためフィッティングにコストがかかり、初期値設定の問題や追跡の失敗に対する復帰が困難であるという問題がある。

Chu らの手法 [8] は、ボリュームデータの各点を Isomap 空間に投影する。Isomap 空間は、データの隣接性を考慮した測地距離を用いてデータを別の空間に投影するため、屈曲した腕や脚も Isomap 空間中では直線状となる。これを利用して、Isomap 空間中でデータを分割して体の各部位を求め、それらを元の空間で再分割することで関節モデルを得ている。この手法は、各フレームの処理がほぼ独立しており、またボリュームから直接多関節構造を見つけるボトムアップアプローチをとっているため、与えられた多関節モデルとマッチングする手法に比

べ、初期値の設定が必要なく、動作の安定性や追跡失敗からの復帰についても優れているといえる。一方で、ボリュームデータの Isomap 空間への変換に時間がかかり、またデータの隣接性を利用しているため、手と手、手や体等、体部位同士が繋がった場合（トポロジ変化）に対応できないという問題がある。

Tanaka らの手法 [9] は、ボリュームデータを細線化処理してグラフ構造に変換し、事前に準備した人体部位を記述したグラフデータベースとマッチングすることで入力データの部位を判別、屈曲点を解析することで関節を抽出する手法である。この手法も Chu らの手法と同様、各フレームでの処理がほぼ独立している。またループ等、人体のトポロジが変化した場合でも対応できるという長所がある。一方で、人体部分が完全に融合してしまった場合は、復元を行うことが困難であるという問題点がある。

3. 提案手法の概要

図 1 に本手法の概要を示す。本手法は、教師サンプルからのクラス識別器の学習ステップと、入力データの認識ステップからなる。また、認識ステップは体の方向を認識するステップと、姿勢を認識するステップの 2 段階で構成されている。これは、方向と姿勢の組み合わせをすべて別のクラスとすると、膨大なクラス数が必要であり、認識精度および速度が低下する可能性があるためである。教師データとして、同じ方向を向き、様々な姿勢をとった人体のボリュームデータを用意する。まず、体方向を認識するためのデータとして、教師データをランダムに選択、水平面に対して 20 度ごとに回転し、18 クラスの教師データ群を得る。次に、体姿勢を認識するための教師データとして、サンプルを人体姿勢に応じてクラスタリングする。ここでは、ボリュームデータを既存手法 [9] を用いて関節データに変換し、体全体のリンク方向の差異を距離関数として、K-means アルゴリズムを用いてクラスタリングを行う。

以上のようにクラス分けされたデータを高さ方向によって正規化し、多次元特徴ベクトルを得る。ここでは、画像特徴として知られる矩形特徴 (Rectangle Feature) [10] を 3 次元ボリュームデータに拡張した体積特徴 (Volumetric Feature) を提案する。Volumetric Feature は、入力となるボリュームデータから様々なサイズの部分領域を取り出し、 x, y, z の各軸方向に分割した差分値である。こうして得られる多次元特徴ベクトルを、各クラス毎に Adaboost アルゴリズムで学習することで、体方向の識別器、体姿勢の識別器を生成できる。

姿勢推定フェーズでは、入力データを正規化して Volumetric Feature を求め、まず体方向を学習した識別器によって体方向を求める。求められた体方向から入力ボリュームデータの方向を正規化、再度特徴ベクトルを抽出し、体の姿勢を姿勢の識別器によって推定する。以降では、体方向および姿勢認識のための教師データの作成、ボリュームデータから Volumetric Feature への変換、Boosting による識別器の学習、および姿勢認識手法について順次述べる。

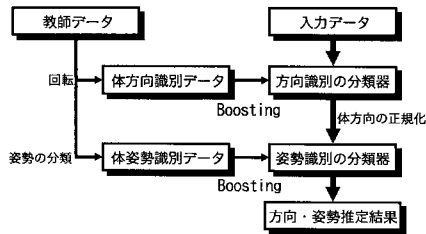


図1 提案手法の流れ

4. 教師データ

教師データとして、複数のカメラを用いて人体のボリュームデータを取得する。ユーザーは環境中の一方向を正面ときめ、その方向を向いて様々な姿勢をとる。この画像を天井に固定された複数のカメラから撮影し、視体積交差法によって体積データを求める。各カメラは、平面パターンを用いてレンズの非線形歪みを校正し、次に既知の大きさのキャリブレーションボックスを用いて、カメラパラメータを求めておく [11], [12]。入力画像から色相を考慮した背景差分を用いることで対象の領域を抜き出し、カメラパラメータを用いて対象の三次元ボリュームデータを 2cm 立方の解像度で復元する。最後に、ユーザーによる体の大きさの影響を防ぐため、データの正規化を行う。床面からの鉛直方向を Z 軸、人体領域の重心を中心を (cx, cy, cz) 、入力データの高さの最大値を z_{max} 、出力データの高さを Z_{out} とした場合、元のデータ (X, Y, Z) を以下の式で正規化する。

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \frac{Z_{out}}{z_{max}} & 0 & 0 & -cx \\ 0 & \frac{Z_{out}}{z_{max}} & 0 & -cy \\ 0 & 0 & \frac{Z_{out}}{z_{max}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

4.1 体方向識別のための教師データ

次に、人体の方向を識別するための教師データを生成する。ここでは、水平方向の回転を 20 度以内の分解能で求めるために、正面を向いて撮影したサンプルを水平方向に回転し、方向を識別するための教師データ群を生成する。すなわち、正面方向から 20 度毎に回転した 18 クラスのデータ群が得られることになる (図 x)

まず、撮影した元のボリュームデータからランダムにサンプルを選び、水平方向に $10 \cdot (n-1) + 20 \cdot rand$ 度 (rad) の回転を行う。ただし $n = \{0, 1, 2, \dots, 17\}$ であり回転に対するクラスの番号を示し、rand は 0~1 までの乱数である。

4.2 姿勢認識のための教師データ

人体の姿勢を識別するための教師データを生成する。ここでは、入力するボリュームデータを、姿勢に応じて任意の数のクラスに分類し、姿勢を表す教師データとする。

まず、入力ボリュームデータを我々が既に提案した手法 [9] にて、人体の関節データに変換する。この手法では、ボリュームデータで体部位が融合した場合などは正しい推定は出来ないため、一部手動にて関節データを与える。用いる人体の関節モデル

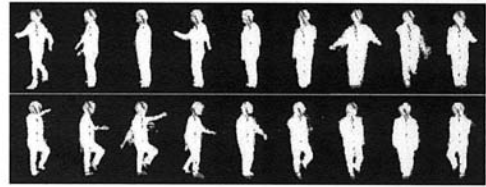


図2 体方向識別のための教師データ。ランダムなサンプルを選び、水平面に対して 20 度ごとに回転させたデータ群を作成する。

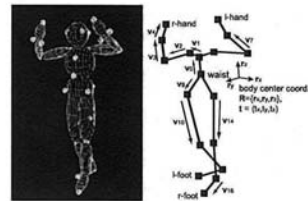


図3 人体の関節モデル

ルを図3に示す。

次に、得られた関節データ間の距離を、体リンク間の角度の線形和で表した以下の式で評価し、K-means 法にてクラスタリングを行う。

$$D(l, m) = \sum_i \beta_i \cdot \arccos(\mathbf{v}_i(l) \cdot \mathbf{v}_i(m)) \quad (2)$$

ただし l, m は入力データの番号を示し、 \mathbf{v}_i は図 x に示された体リンクベクトルを表している。K-means クラスタリングでは、まず入力データの中からランダムに K 個のサンプルを選んでクラスタリングの種 (seed) とし、次に全フレームに対して 2 式を用いて seed となる姿勢データとの距離を計算、もっとも近いクラスに属するとみなす。分類された各クラスの平均姿勢を、体リンクのベクトルの平均として求め、新たなクラスタの seed とする。以上の処理を繰り返して、入力姿勢を決められた数のクラスタに分類することができる。

5. Volumetric Feature

クラス分けされたボリュームデータ (教師データ) を、多次元特徴量として表すために、本研究では、二次元画像の特徴量として Viola and Jones らが提案した矩形特徴 (Rectangle Feature) を 3 次元体積空間に拡張した体積特徴 (Volumetric Feature) を用いる。Rectangle Feature は、画像領域の差分値を用いるシンプルな画像局所特徴であるが、これを多数用いることで高い認識率が得られることが知られている。また、Integral Image を用いた計算手法により、多くの特徴値を高速に計算することも知られている。提案する Volumetric Feature は、この Rectangle Feature を三次元に拡張したものであり、特徴値の計算は、二次元画像における Integral Image と同様、あらかじめ積分画像 (Integral Volume) を計算することによって高速に求めることができる。

Volumetric Feature は、ボリュームデータ内の部分空間とな

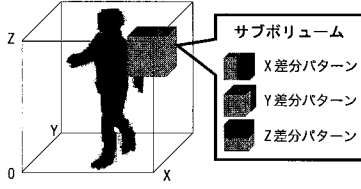


図4 Volumetric Feature. 本研究では, X,Y,Zの3方向の差分値を用いる。

る直方体領域内に含まれるボクセル値の差として計算できる。パターンとり方は様々なものが考えられるが, 本実験では x, y, z 各軸方向に同サイズのボクセルが隣接する3種類を用いている(図4)。

a) Integral Volume

Integral Volume はサブボリューム内に含まれるボクセル値を中間的表現を用いて Volumetric Feature を高速に計算する手法である。まず, ボリュームデータのある点を原点とした座標空間を考え, 空間内の座標 x, y, z に関して各座標の値を以下のように定義する。

$$iv(x, y, z) = \sum_{x' \leq x} \sum_{y' \leq y} \sum_{z' \leq z} i(x', y', z') \quad (3)$$

$iv(x, y, z)$ は Integral Volume で, $i(x, y, z)$ は座標上のボクセル値である(図を参照)。計算には以下の方法で再帰的に計算する。

$$l(x, y, z) = l(x, y, z-1) + i(x, y, z) \quad (4)$$

$$s(x, y, z) = s(x, y-1, z) + l(x, y, z) \quad (5)$$

$$iv(x, y, z) = iv(x-1, y, z) + s(x, y, z) \quad (6)$$

ここで $l(x, y, z)$ とは累積的な列の和, $s(x, y, z)$ は累積的な平面の和, $l(x, y, -1) = 0, s(x, -1, z) = 0, iv(-1, y, z) = 0$ である。Integral Volume は入力ボリュームデータに対して一回の走査で計算できる。この Integral Volume を使うことで, サブボリューム内の直方体領域に含まれるボクセル値を8つの要素の参照で計算することができる。すなわち, (x_0, y_0, z_0) から (x_1, y_1, z_1) を対辺とするサブボリューム内に含まれるボクセル値の和は, $iv(x_1, y_1, z_1) - iv(x_0, y_1, z_1) - iv(x_1, y_0, z_1) + iv(x_0, y_0, z_1) - iv(x_1, y_1, z_0) + iv(x_0, y_1, z_0) + iv(x_1, y_0, z_0) - iv(x_0, y_0, z_0)$ で得られる。

6. Boostingによるクラス識別器の学習

AdaBoost を用いて, 体方向および姿勢を認識する識別器を学習する。ある方向を示すデータをポジティブサンプル群, それ以外の方向を表す17クラスのデータをネガティブサンプル群として学習し, 18個の識別器を作成する。姿勢を認識する識別器も同様に, 認識対象のクラスのデータをポジティブサンプル, それ以外のデータをネガティブサンプルとし学習を行う。以下では, 特徴ベクトル \mathbf{x} に対するあるクラス k の分類器 $H^k(\mathbf{x})$ を, AdaBoost で識別する学習する場合について記述する。

6.1 分類器の概略

分類器は, Volumetric Feature の特徴ベクトルの各要素を閾値によって識別する弱分類器を準備し, その出力を重み付き線形和で組み合わせたものである。弱分類器 $h_j(x), (j = 1, \dots, J)$ は, 入力データに対して一つの閾値 θ_j を用いた二値関数である。

$$h_j(x_j) = \begin{cases} 1 & \text{if } x_j < \theta_j \text{ or } x_j > \theta_j \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

ただし x_j は教師データの j 番目の要素を表す。サンプルがそのクラスに属している場合には1を, 属していない場合には0を出力する。この弱分類器を組み合わせて, あるクラス k を識別する学習済み分類器 $H^k(x)$ は以下の関数で表される。

$$H^k(\mathbf{x}) = \sum_{j=1}^J \alpha_j h_j(x) \quad (8)$$

α_j は Volumetric Feature の要素 j の分類器 $h_j(x)$ の重みである。分類器 H の出力が大きいくほど, データがそのクラスに属する可能性が高いといえる。AdaBoost による学習は, 対象クラスのデータとそれ以外のデータを最もよく分離できる各分類器 $h_j(x)$ の閾値とその重みを得ることが目的である。

6.2 AdaBoostの学習アルゴリズム

教師データとして, 対象とするクラスのデータ(ポジティブサンプル)およびそれ以外のデータ(ネガティブサンプル)を用意する。すなわち, 教師データ \mathbf{x} とその属性 y の組 $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_{m+n}, y_{m+n})$ とした場合, クラス k に属しているサンプルをポジティブサンプル (m 個), 属していないサンプル (n 個) がネガティブサンプルである。属性値 y はポジティブサンプルなら1, ネガティブサンプルなら0である。また, \mathbf{x} の j 番目の要素を x^j とする。AdaBoost の学習アルゴリズムを以下に示す。

入力 教師データ \mathbf{x} とその属性 y の組 $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_{m+n}, y_{m+n})$
出力 $H^k(x)$ の各分類器 $h_j(x)$ の閾値 θ_j , 重み α_j

Step 1 各サンプルの重み $w_i (i = 1, \dots, m+n)$ を初期化する。 $y = 1$ の場合は $w_i = \frac{1}{2m}$, $y = 0$ の場合は $w_i = \frac{1}{2n}$ とする。

Step 2 各弱分類器の閾値 θ_j を探索する。すべての $j = 0, \dots, J$ に対し, $\epsilon_j = \sum_{i=1}^{m+n} w_{j,i} \times |h_j(x_i) - y_i|$ が最小となる h_j の閾値 θ_j を求める。

Step 3 分類器の重み付きエラー ϵ_j を以下の式で求める。

$$\epsilon_j = \sum_{i=1}^{m+n} w_{j,i} \times |h_j(x_i) - y_i|$$

Step 4 分類器の重み α_j を以下の式で求める。

$$\alpha_j \leftarrow -\log(\beta_j) \quad (\beta_j = \frac{\epsilon_j}{1-\epsilon_j})$$

Step 5 サンプルの重みを更新し正規化する。すなわち, $h_j(x_i) > y_i$ の場合, $w_{j+1,i} \leftarrow \beta_j \times w_{j,i}$ とする。

Step 6 Step 2-5 を, 決められた回数繰り返す。

Step 1 では学習サンプルに対し重みを設定する。Step 2 では特徴ベクトル \mathbf{x}_i の各要素に対し, 最適な分類器の閾値を探索する。最適な閾値は各サンプルの重みを考慮し, 重み付きエラーが最小になる値を選択する。Step 3-4 では Step 2 で求めた閾

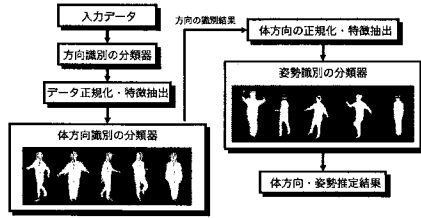


図5 体方向・姿勢認識処理の流れ

値を用いて分類器 h_j の重み付きエラーを求め、求めた重み付きエラーから分類器の重みを決定する。これにより、精度の高い分類器ほど重み α_j が大きくなる

Step 5 では Step 3 で求めた重み付きエラーを元に、各サンプルの重みを更新する。正解したサンプルの重みを軽くし、誤認識されたサンプルに対して重みを加えることで、次の学習で間違ったサンプルもうまく学習できるようにしている。以上の処理で得られた出力結果を、各クラスの識別パラメータとして使用する。

7. 識別器を用いた方向および姿勢認識

入力データを AdaBoost による学習済みのクラス分類器で評価し、どのクラスに分類されるかを決定する。認識処理の流れを図5に示す。入力データは教師データと同様に正規化および特徴ベクトル抽出が行われ、まず、体方向の分類器で体方向を認識する。認識結果を用いて、元の入力データを正面方向に回転し再度特徴ベクトルを抽出、姿勢の分類器で姿勢の認識を行う。

方向および姿勢の認識は、複数の分類器によって構成されているので、入力データから得られた特徴ベクトルを全分類器で評価し、その中で最も大きいの出力した分類器のクラスに属しとするとみなす。

8. 実験

提案手法の有効性を確認するための実験を行った。教師データ、および認識を行うための入力データとして、前述の8台のカメラから構成されるスタジオを用いた。被験者はスタジオ内の一方向を正面と決め、歩行や体操などの様々な動作をとる。この映像を同期カメラで撮影し、PC クラスタを用いて三次元ボリュームで他の復元、特徴ベクトルへの変換を行った。

被験者は2名であり、それぞれ被験者A(身長166cm)および被験者B(身長173cm)である。教師データは被験者Aの計測結果である756フレームを用い、3種の歩行パターン(362フレーム)、様々なポーズ(394フレーム)からなる。正規化したボリュームデータのサイズは $64 \times 64 \times 64$ とし、この中から $16 \times 16 \times 16 \sim 54 \times 54 \times 54$ までのサブボリュームをとり、74089次元の Volumetric Feature を得た。

8.1 姿勢のクラスタリング

被験者Aから得られた教師データを既存手法で処理して関節角度を抽出、関節方向に従って K-means アルゴリズムで分類

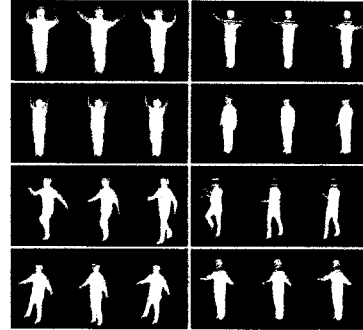


図6 姿勢のクラスタリング結果

した結果を図6に示す。クラスタ数は10とした。歩行データから選び出された右ステップ、左ステップが正しく分類されていることが確認できる。そのほかの姿勢においても、立位、手を挙げる動作などが数パターンに分類され、おおむね人の直感に近い分類が実現出来ていることが確認できた。全756フレームのボリュームデータに対し、最も少ないクラスは45フレーム、多いクラスは172フレームであった。

8.2 識別器の学習

人体方向および姿勢を認識するための識別器の学習を行った。人体方向の学習は、様々な姿勢データからランダムにポジティブサンプル100個選び、ネガティブサンプルを毎クラス(方向)50個選択し、計950個のサンプルを1クラスの学習に用いた。学習時間は、PC(Pentium4, 4GB Memory)で、1クラスあたり276[sec]である。

姿勢の識別器にはポジティブサンプルを50個、ネガティブサンプルを30個をランダムに選び出した。各クラスに属するサンプル数が異なるため、一部のクラスでは同じサンプルが含まれている可能性がある。学習時間は同じ環境で154[sec]であった。

8.3 実験結果と考察

体方向および姿勢認識実験として、被験者Aが歩き回るシーケンス(99 frames)、被験者Bが歩き回るシーケンス(167 frames)に対して実験を行った。なお、被験者Aのシーケンスは学習に用いたものとは別の動作列である。図7~8に、被験者A、Bの歩行軌跡と本手法で認識された体方向と姿勢パターンをプロットした結果を示す。なお、歩行軌跡は人体領域の重心位置をトレースしたものであり、この歩行軌跡に沿って人は移動しているため、進行方向を体の正面方向とみなすことができる。実験結果から、体方向に関してはおおむね良好に認識を行うことが出来ているが、一部で誤りがあるフレームも見られた。これらに関しては、誤って認識された結果が正しい方向からほぼ180度戻った結果であることから、人体姿勢の前後判定に失敗した結果であることがわかる。また、教師データと体の大きさ等が異なる被験者Bの実験結果から、被験者Aと変わらない程度の結果が得られていることがわかる。これにより、教師データの違いによる個人差は、人体方向の認識に関しては

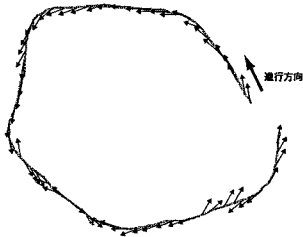


図7 被験者 A の歩行実験。歩行軌跡 (灰色実線) と識別された人体方向 (矢印)。

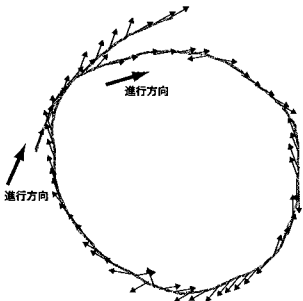


図8 被験者 B の歩行実験。歩行軌跡 (灰色実線) と識別された人体方向 (矢印)。

それほど影響が無いことがわかる。

姿勢推定の結果を表 1 に示す。対象は歩行動作なので、姿勢は左ステップ・右ステップに分離できる。姿勢のクラスタリングによって、それぞれの姿勢クラスが分類されているので、入力フレームがそれぞれのクラスに正しく分類されているかを評価した。実験の結果、被験者による認識率の差異はほとんど見られず 70% 程度の認識率であり、体方向推定とほぼ同様の結果であることがわかった。この結果を、正しく体方向推定が行われた群と、誤った群に分けて評価したところ、ほとんどの姿勢推定エラーは、体方向推定が誤った場合に生じており、両者が密接に関連していることがわかった。一方で、正しく方向を推定できていた場合には、姿勢も高い確率で正しい結果が得られており、方向推定の精度が重要であることがわかる。このことから、今後の改良点として、前出の 180 度裏返った間違いに対する対処を行うことが重要であることがわかった。識別時間に関しては、学習と同様の PC を用いて 1 識別器あたり約 1.7 秒、方向および姿勢を認識する全処理において 25 秒で処理を行うことが可能だった。今回はカスケード型識別器の特徴である、一部の特徴のみを使うという高速化を行っていないが、重みの大きい識別器だけを取り出して認識することによって、より高速な姿勢認識が可能になると考えられる。

9. まとめと課題

本論文では、ボリュウムデータからの人体の姿勢推定問題を解決するために、人体の姿勢に関するさまざまな事例を準備

	全正答率	方向推定が誤った場合の正当率	方向推定が正しい場合の正当率
被験者 A	69.7%(69/99)	0% (25/25)	93.2% (69/74)
被験者 B	70.7%(118/167)	0% (25/25)	94.4% (118/125)

表 1 歩行実験における姿勢認識結果

し、それらと入力データを比較する推定手法を提案した。教師データとなる姿勢サンプルとして、複数のカメラから復元された人体のボリュウムデータを用い、これらを任意のクラスに分類・正規化し、画像特徴である Rectangle Feature を 3 次元ボリュウムデータに拡張した Volumetric Feature を求める。この特徴ベクトル群を Boosting により学習してクラス分類器を作成し、入力データに対して適用することで、体の方向および姿勢を推定する。実験では、体方向および姿勢についておおむね良好な推定結果を得ることが出来たが、一方で、体が「裏返る」状況に対しての認識精度が弱く、これを解決することが必要であることが示唆された。今後の課題として、体方向の識別として時系列的な認識を行うこと、サンプルデータとして CG によるシミュレーションデータを用いること、識別器のチューニングによる高速化、取束法との組み合わせによる姿勢推定の高精度化などが挙げられる。

文 献

- [1] D. M. Gavrilu. The visual analysis of human movement: A survey. *Computer Vision and Image Understanding*, Vol. 73, No. 1, pp. 82-98, 1999.
- [2] J. K. Aggarwal and Q. Cai. Human motion analysis: A review. *Computer Vision and Image Understanding*, Vol. 72, No. 3, pp. 428-440, 1999.
- [3] T. B. Moeslund and E. Granum. A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, Vol. 81, No. 3, pp. 231-268, 2001.
- [4] 岡田隆三, シュテンガ ビヨンシルエットを用いた Tree-Based Filtering による人体の姿勢推定画像の認識・理解シンポジウム講演論文集, pp.64-79, 2006.
- [5] Liu Ren, Gregory Shakhnarovich, Jessica K. Hodgins, Hanspeter Pfister, Paul Viola. Learning Silhouette Features for Control of Human Motion ACM Transactions on Graphics, Vol. 24, No. 4, pp. 1303-1331, 2005.
- [6] I. Mikić, M. Trivedi, E. Hunter, and P. Cosman. Human body model acquisition and tracking using voxel data. *IJCV*, Vol. 53, No. 3, pp. 199-223, 2003.
- [7] F. Caillette and T. Howard. Real-Time Markerless Human Body Tracking with Multi-View 3-D Voxel Reconstruction. In *Proc. of British Machine Vision Conference*, Vol. 2, pp. 597-606, 2004.
- [8] C. W. Chu, O. C. Jenkins, and M. J. Mataric. Markerless kinematic model and motion capture from volume sequences. In *Proc. of CVPR'03*, Vol. 2, pp. 475-482, 2003.
- [9] 田中秀典, 中澤篤志, 町田貴史, 竹村治雄. ボリュウムデータの細線化とグラフマッチングを用いた事例ベース人体姿勢推定画像の認識・理解シンポジウム講演論文集, pp.70-77, 2006.
- [10] Paul Viola, Michael Jones. Rapid Object Detection using a Boosted Cascade of Simple Features In *Proc. of CVPR'01*, Vol. 1, pp. 511-518, 2001.
- [11] 加藤博一, M. Billinghamurst, 浅野浩一, 橋啓八郎: マーカー追跡に基づく拡張現実システムとそのキャリブレーション', 日本バーチャルリアリティ学会論文集, Vol. 4, No. 4, pp. 607-616, 1999.
- [12] 井口征士, 佐藤宏介: 三次元画像計測', 昭見堂, 1990.