

全身によるアクロバティックな武道動作に対する 知識ベースの階層的記述と認識

宇野 友季子^{1,a)} 佐藤 啓宏¹ 工藤 俊亮² 池内 克史¹

概要：モーションキャプチャが容易になりつつある現在，様々な全身運動に対して網羅的な記述方法による意味付けが有用となる．本稿では，三次元的な攻防を特徴とする「躰道」という武道を対象として，動作の知識を元にしたトップダウンな記述・認識モデルを提案する．人体を階層的円筒モデルで表現し，全身の回転や移動など大まかな動作から，手先・足先の細かい所作へと注目していく．このような階層的表現によって，多様な動作を直感的に理解でき，かつ効率的な記述を行うことができる．また，実際にモーションキャプチャデータを用いて認識を試みた．

1. はじめに

コンピュータビジョンの分野において，人間が「何を」行っているかを機械に理解させることが重要視され，人体の動作認識について様々なアプローチが試みられてきた．ゲーム操作，監視カメラによる危険行為の検出，スポーツ実況，ヒューマンマシンインターフェースなど，実社会での様々なアプリケーションにおける実用化や他の研究分野への応用が期待されている．

近年まで，容易に撮影・入手できる 2D 動画像からの人体姿勢，手指姿勢の推定が盛んに研究されてきた [1][2]．動作認識の分野においても，2D 画像・動画を入力として学習を行うものが多い．特に歩行などの日常的な反復動作やジャンプ，投球などのスポーツ動作の認識 [3]，あるいは手によるジェスチャーの認識を試みる研究などが数多く行われている．こういった手法においては，2D 画像から得られる情報の撮影視点による変化に耐えるために，大量の学習データが必要になることが多い．

しかし，Kinect などの登場によって空間の三次元情報が容易に得られるようになり，人体姿勢の推定もより手軽かつ正確に行えるようになってきた [4]．そのため，2D 画像ではなく，3D 情報から得られた姿勢，つまり関節角度の時系列データを入力とした動作認識を行うことができるようになった．

そこで，与えられた動作が何を意味するものなのかを，

日常動作に限らず，より網羅的に認識することが求められるようになった．しかし，人体が構造上取りうる全ての全身動作の一つ一つについて，それぞれが持つ意味を記述することは困難である．なぜなら，全身動作には細かい所作の異なる膨大なパターンがある上に，認識すべき結果は動作の行われた環境などによって異なるためである．

Marr ら [5] は，人間は物体の形状を認識する際，全体の大まかな状態から末端へと階層的に捉えて行くと考え，図 1 のような人体の階層的円筒モデルを提唱した．本稿ではこのモデルを動作に拡張し，多様な全身動作の階層的な記述方法を提案する．動作の知識を用いた直感的で必要十分な記述を一通り行えば，認識結果と認識過程を見るだけで動作に関する知識を人や機械に伝達することができる．ダンスや武道など一つ一つの動作に意味を持ったものが認識対象として適すると考え，「躰道」[11] という武道を選択した．躰道は全身運動によるアクロバティックな技を特徴とし，技のすべては体幹の動きの性質によって大きく 5 つに分類されているため，階層的な認識に適すると考えられる．そこで，躰道の動作について階層的認識器を生成し，モーションキャプチャデータを用いて認識を試みた．

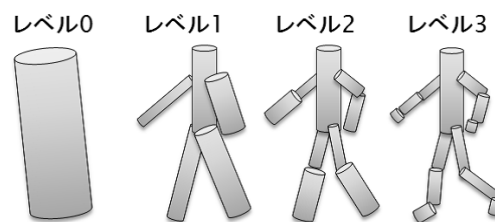


図 1 階層的円筒モデル

Fig. 1 Hierarchical Cylinder Model.

¹ 東京大学
The University of Tokyo

² 電気通信大学
The University of Electro-Communications

a) uno@cvtl.iis.u-tokyo.ac.jp

2. 関連研究

一連の全身動作を分解し記述する試みが、舞踊などの分野について行われている．池内ら [6][7] は “Learning from Observation (LFO)” のパラダイム [8] とモーションキャプチャデータに基づき、上半身の止め動作と下半身の動作モデルによる伝統舞踊の記述と、ヒューマノイドロボットによる再現を行った．八村ら [10] は図形記号による舞踊記述法 Labanotation をモーションキャプチャデータに基づいて生成し、伝統舞踊の保存と解析に役立てるという試みを行っている．また、専門知識を必要とする殺陣シーンの構成を効率化するため、剣戟アクションを分解した基本動作のモーションデータの収録と、それらを CG 上で再構成する手法が考案されている [9]．

これらの研究は、関節角度という量的なデータを人間にも直感的に理解できる質的な表現に変換し、対象とした舞踊などに関する専門知識がなくても動作の内容を保存・伝達可能にしている．

3. 躰道について

本稿において動作記述・認識の対象としている躰道は、創始者である祝嶺顕献が空手の一派「玄制流」を母胎とし、1965 年に体系化した新しい武道である．空手と同様に突き・蹴りによる攻撃が主体であるが、床に鉛直な姿勢を保つことは求められておらず、人体の取りうるあらゆる三次元的動作が取り入れられている．体幹の姿勢を大きく変えることで移動や相手の攻撃の回避を行うと同時に、攻撃に威力を乗せることができる．その複雑な動作の数々は考案・伝承の過程で体系化され、図 2 に、示すような 5 種類に分類されている．順に、

旋技 独楽のように旋回する動作

例：旋回からの突き、回し蹴り

運技 跳躍や膝の引き上げによる前進と上下動の勢いを利用する動作

例：飛び蹴り、前進しながらの前蹴り

変技 体を素早く倒す動作

例：横に倒れながらの蹴り、後に倒れながら斜め上への蹴り

捻技 腰を捻り回転を用いて相手を巻き込む動作

例：伏せた姿勢からの真上への蹴り、相手を両腿で挟んで絡み倒す技

転技 前転・側転・宙返りなどの縦回転を用いた動作

例：前転からの突き、バック転からの後方への蹴り

である．これらの体操作を、単眼カメラで撮影した 2D 動画を用いて分類する研究も行われている [12]．

同じタイプに分類された技の中では、体操作の方向や攻撃に使う部位などによってさらに細かく、個々の技として区別されている．例えば同じ旋技や運技の中に蹴り技、突

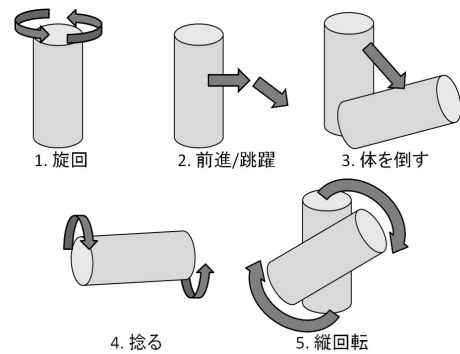


図 2 躰道の体操作 5 種

Fig. 2 5 kinds of Body Motions in Taïdo.

き技、貫手の技（掌を平らにして揃えた指先で相手を突く）がある．また、変技や転技の中には体を倒す方向や回転する方向によって異なる技が存在する．

4. 全身動作の階層的記述法

本章では、図 1 の階層的円筒モデルに基づき、具体的な躰道の動作を階層的に認識する手法を述べる．

4.1 動作の定義

本稿では、各円筒の形状・位置・姿勢に注目し、フレーム間での変化量に基づいて一連の動作を静止状態と動作状態の繰り返しと見なした．さらに、一つの動作を、ある静止状態（フレーム f_s ）から次の静止状態（フレーム f_e ）への変化と定義する．

4.2 レベル 0

全身を一つの円筒と見なした際の大まかな動作分類について、図 3 のような座標系において

- 動作直前の円筒の傾き: $\angle(\vec{z}_{f_s}, \vec{Z})$
- 動作中の一回転以上の縦回転量: $\sum_{f=f_s}^{f_e-1} \Delta \vec{z}_f$
- 重心の上下運動量: $\max(o_z(f)) - \min(o_z(f))$
- 円筒の長さの変化量: $\max(L(f)) - \min(L(f))$
- 動作直後の円筒の傾き: $\angle(\vec{z}_{f_e}, \vec{Z})$

に注目し、各値に対して適切に設定した閾値との比較もとに図 4 のように分類する．縦方向に回転・跳躍・伸縮の有無、横方向に動作前後のポーズで場合分けをしている．

例えば、鉛直姿勢から縦回転を伴って、跳躍・伸縮せずに鉛直姿勢に至る動作は背転技（前転・後転）である．

4.3 レベル 1: 胴体部分

四肢を除く胴体部分について、レベル 0 より細かい分類を行う．具体的には、静止状態の円筒の傾きは 3 パターンに分け、動作については図 3 における体の正面方向 $\vec{y}$ に注目し、

- $\vec{y}$ まわりの旋回量: $\sum_{f=f_s}^{f_e-1} \angle(\vec{y}_f, R_z \cdot \vec{y}_{f-1})$
- $\vec{y}$ まわりの回転量（縦回転の回転方向（前/側/後方））

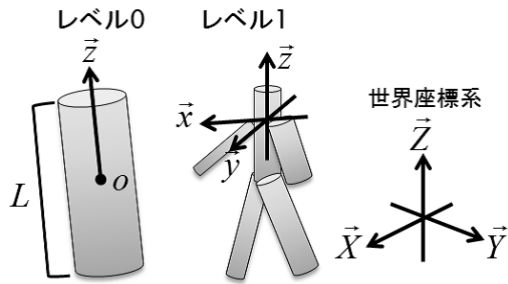


図 3 レベル 0, 1(胴体) の座標系

Fig. 3 The Coordinate System for Level 0, 1(Torso).

回転	跳躍	伸縮	ポーズ 1		ポーズ 2	
			鉛直	傾斜	鉛直	傾斜
○	○	○	宙返り	宙返り	飛び込み転技	宙返り
		×	伸身転技	伸身転技	-	伸身転技
	×	○	背転技	背転技	-	背転技
		×	腕転技	腕転技	-	腕転技
×	○	○	二段蹴り	二段蹴り	二段蹴り	-
		×	二段蹴り	二段蹴り	二段蹴り	-
	×	○	(旋回技)	(伏せ技)	(復帰)	-
		×	(移動技)	(伏せ技)	(復帰)	(捻り技)

図 4 レベル 0(全身) の動作分類

Fig. 4 Motion Classification in Level 1 (Whole Body).

$$: \sum_{f=f_s}^{f_e-1} \angle(\vec{v}_f, R_x \cdot \vec{v}_{f-1})$$

で分ける。 R_z は $\vec{v}_f \times \vec{v}_{f-1}$ まわりに $\angle(\vec{v}_f, \vec{v}_{f-1})$ だけ回転を行う回転行列である。図 5 に示すように、レベル 0 と同様、縦方向に動作のタイプ (縦回転の有無、縦回転の方向、旋回の有無)、横方向に動作前後のポーズで場合分けをしている。

例えば、鉛直姿勢から縦回転・旋回を経ずに下向きの姿勢に至る動作は海老蹴りである。

4.4 レベル 1~3: 四肢部分

四肢については静止状態のみを分類する (図 7, 8)。各レベルの円筒の形状情報 (伸縮) と、図 6 の座標系における姿勢情報 ($\vec{v}$: 円筒の軸, $\vec{v}_f$: 円筒の正面) の一部に注目し、全体から末端へとポーズを限定していく。なお、図 7, 8 中の「外」とは右手足なら右、左手足なら左、「内」とはその反対方向である。図 9, 10 に、剣道の技の途中で見られるポーズの外観一覧を示す。姿勢情報は、人間にも意識しやすい以下のものを用いた。手足の姿勢のみで判断できないものは、胴体部分の姿勢によって決定できる。

腕

- レベル 1: 軸方向 $\vec{v}$ (肩→指先)

ポーズ 1	ポーズ 2	回転方向	上向き		横向き		下向き	
			上向き	横向き	下向き	上向き	上向き	下向き
○	後	○	後方転技	-	-	-	-	-
		×	後方捻り転技	-	後方転技	-	-	後方転技
	横	×	側転	側転斜状蹴り	-	-	-	-
		○	前方転技	-	前方転技 前方捻り転技 前転海老蹴り	-	-	-
×	前	○	前方転技	-	-	-	飛び込み前転	-
		×	(旋回技)	旋体斜状蹴り	-	(復帰)	(復帰)	絡み技
×	-	×	(移動技)	蹴り	海老蹴り	(復帰)	(復帰)	半月当て

図 5 レベル 1(胴体) の動作分類

Fig. 5 Motion Classification in Level 0 (Torso).

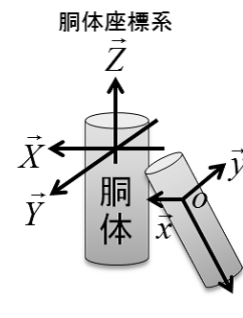


図 6 四肢の座標系

Fig. 6 The Coordinate System for Arms and Legs.

- レベル 1: 正面方向 $\vec{v}_f$ (肘を向ける方向)
- レベル 3: 軸方向 $\vec{v}$ (手首→指先)
- レベル 3: 正面方向 $\vec{v}_f$ (掌を向ける方向)

脚

- レベル 1: 軸方向 $\vec{v}$ (腰→足先)
- レベル 1: 正面方向 $\vec{v}_f$ (膝を向ける方向)
- レベル 3: 軸方向 $\vec{v}$ (踵→足先)

問題点として、個人差や誤差が原因で静止姿勢がどのポーズにも分類されない可能性がある。これを解決するため、左右の組み合わせを用いる。左右の手または足がともに静止状態だった場合で、一方のポーズのみが決定されていて他方の姿勢が決まらない場合、決定されているポーズとの組み合わせを取りうるポーズの中から最も近いものに決定する。例えば、右足が後屈立ちの後ろ足だった場合は左足が後屈立ちの前足、右手が前方貫手だった場合は左手が前方貫手または引き拳となる。

また、フレームレートによっては短時間の静止姿勢が抽出されない場合がある。そこで、レベル 1 の円筒の長さに注目し、急な伸縮を伴う動作の前後も静止姿勢の候補として判定の対象に加えた。

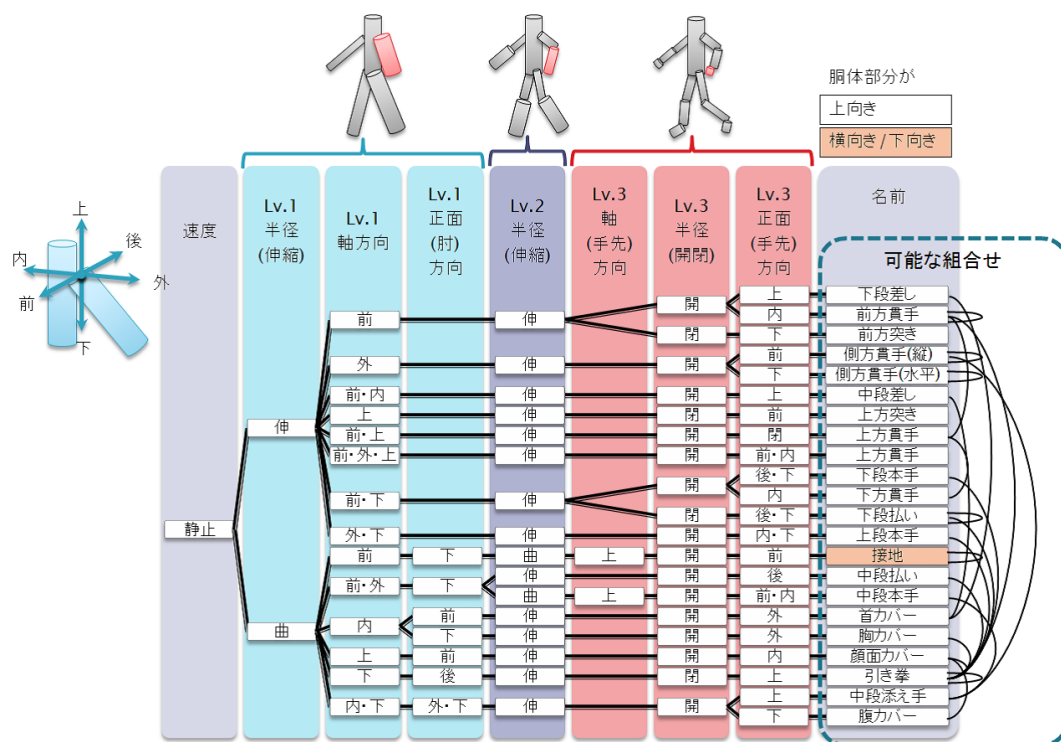


図 7 腕のポーズ分類

Fig. 7 Pose Classification of Arms.

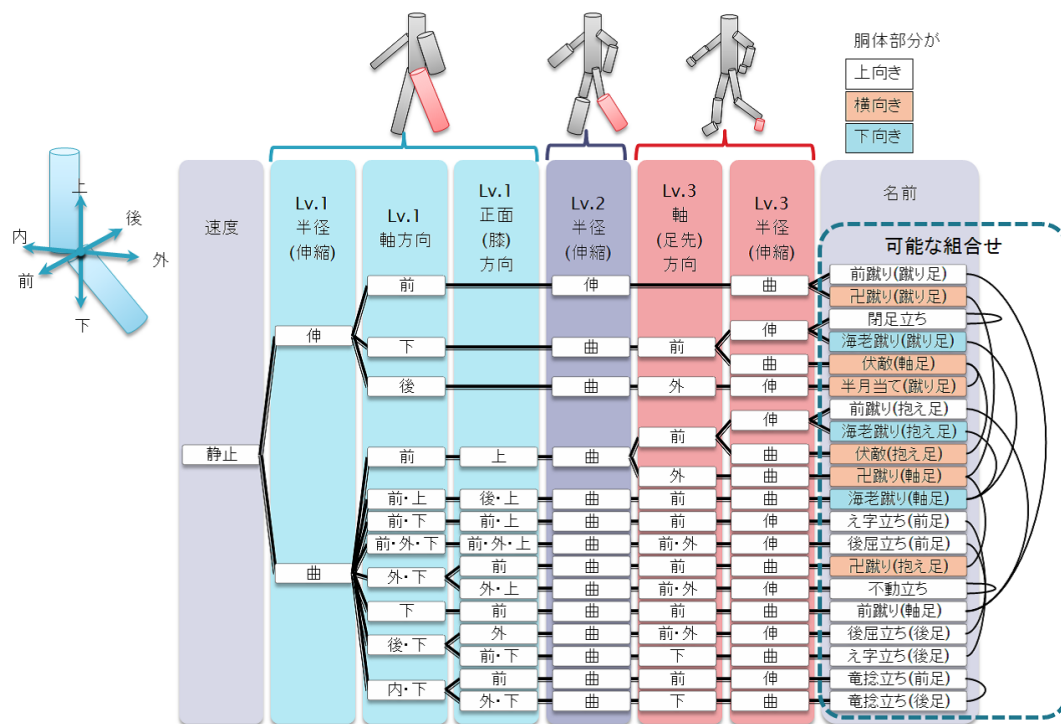


図 8 脚のポーズ分類

Fig. 8 Pose Classification of Legs.

以上の認識器を用いて認識を行う．まずレベル 0 の全身の大まかな動きを決定し，詳細な体捌きと，それに含まれる細かい手足の所作までを段階的に決定していく．最終的

に単一の技名が得られ，認識の過程を見ることで，剣道の知識がなくても簡単な技の特徴を知ることができる．

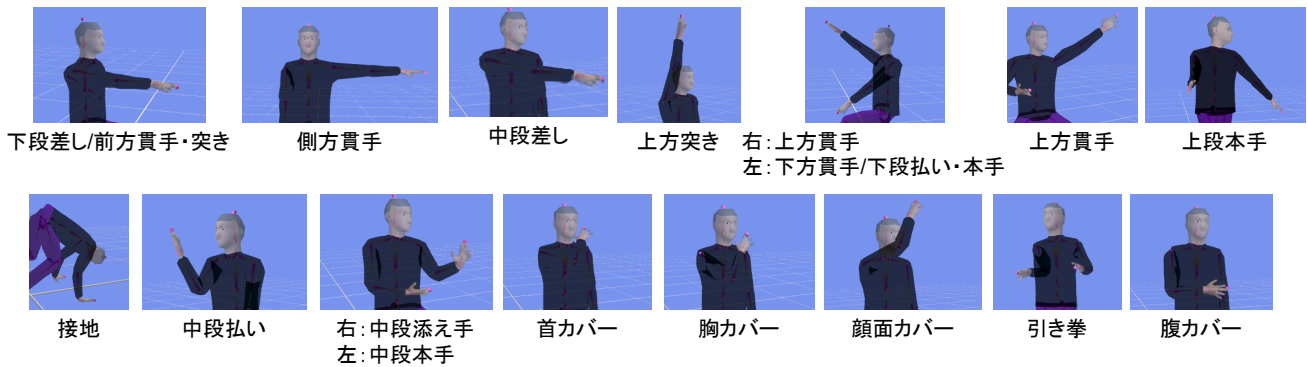


図 9 腕のポーズ外観
Fig. 9 Appearances of Arm poses.

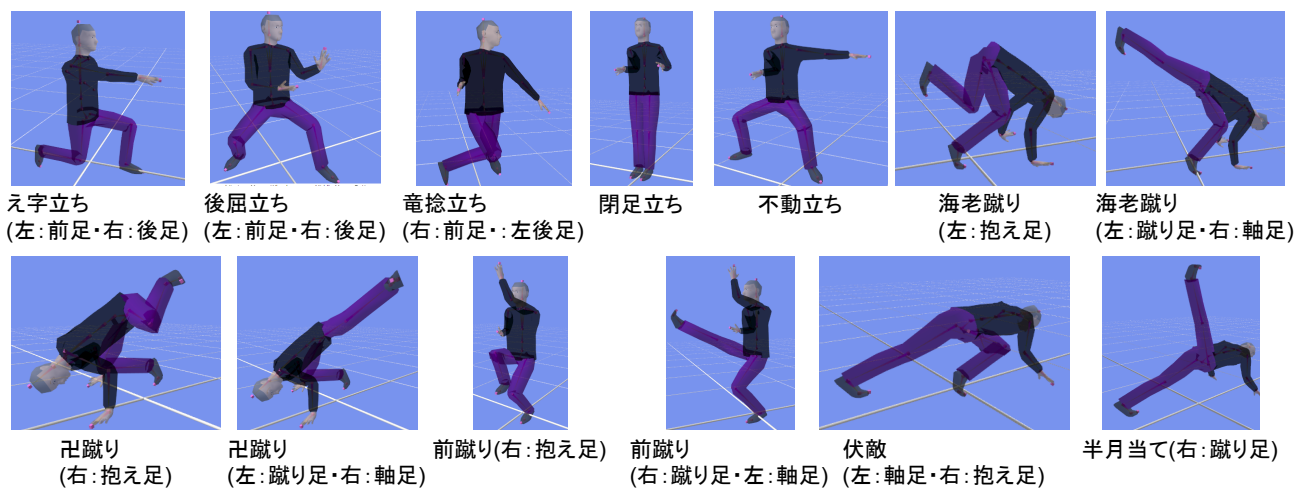


図 10 脚のポーズ外観
Fig. 10 Appearances of Leg Poses.

5. 実装と実験

5.1 モーションキャプチャデータの取得

剣道の動作を取得するため、2 台の Kinect で動作を撮影し、モーションキャプチャソフト”iPiSoft”[13] を用いて処理を行った。図 11 に、得られるボーンモデルの構造を示す。

5.2 階層的円筒モデルの生成

ボーンモデルから図 1 のような円筒モデルを次の手順で計算する。

各円筒の計算に用いたボーンの範囲を図 11 に示す。まず、ボーンを等間隔に配置した点の集合と見なし、重心座標を求める。

次に、円筒の軸を計算する。レベル 0 の全身、またはそれ以上のレベルでの胴体部分については、点の集合を用いて主成分分析で軸となるベクトルを求める。四肢部分については、起点から末端を結ぶ線との平行であり、重心を起

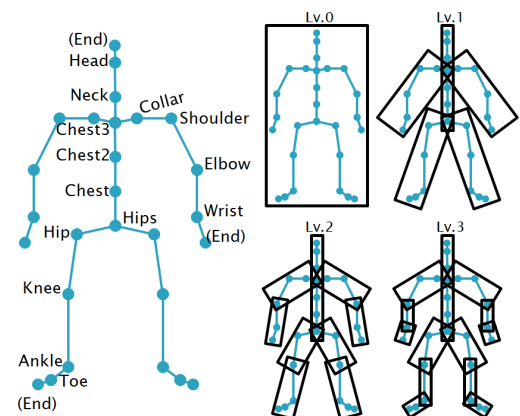


図 11 iPi Soft におけるボーン構造と円筒に含む関節
Fig. 11 The Bone Structure Output from iPi Soft and Joints Included in Cylinders.

点とするベクトルを軸とする。

円筒の半径は各点から軸への距離の平均、円筒の長さは各点から軸へ下ろした垂線の足の存在範囲とする。

また、円筒の正面 (図 3,6 での $\vec{y}$) を決める。全身・胴

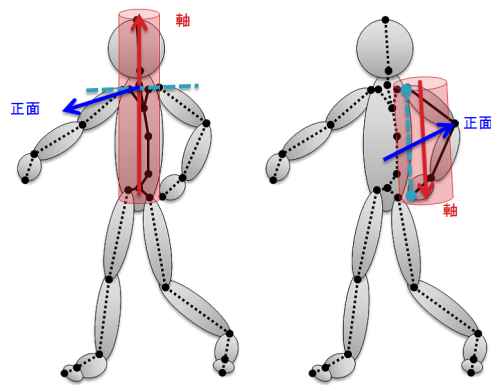


図 12 レベル 1 の軸計算 (左: 胴体部分, 右: 左腕部分)

Fig. 12 Axis Calculation in Level 1
(Left: Torso, Right: Left Arm).

体部分については左右の肩を結ぶ線と円筒の軸の両方に垂直な方向とする。四肢については、円筒に含まれる関節をすべて伸ばしたときに最も中間に近い関節点を通り、軸に垂直な方向とする。

胴体部分以外については、各座標やベクトルは体座標系上で計算する。図 12 にレベル 1 の円筒モデルについての軸を示す。

5.3 認識結果

旋回からの突き、蹴り（横に倒れながらの前蹴り）、バック宙のモーションデータについて、前節の分類手法を用いて動作・ポーズの認識を行った結果を図 13 に示す。なお、モーションキャプチャシステムの仕様上、手首・足首より先の関節角度は精度が得られなかったため省略し、レベル 2 までの認識を行った。結果、現れた静止姿勢のほとんどを抽出・判定することができたが、伸縮が少なく静止時間も短い顔面カバーなどは抽出に失敗することがあった。これを解決するには、円筒の伸縮だけでなく、軸の角度や重心位置の遷移についても考慮することが考えられる。

6. 結論

本稿では、剣道におけるアクロバティックな動作の階層的表現方法を提案し、モーションキャプチャデータにおける認識を試みた。結果、具体的な技や手足の所作の分類と共に、動作に対する直感的な理解が得られた。他の競技やダンスなどについても、動作に関する知識を元に同様の認識器を構築することができる。

本稿では一人の動作を対象に記述を行ったが、今後は対戦における動作のモデル化や、小型ヒューマノイドロボットによる実演を行う予定である。

謝辞 本研究は科研費 MEXT/JSPS 23240026 の助成を受けたものである。

参考文献

- [1] Y. Yang, D. Ramanan: *Articulated pose estimation with flexible mixtures-of-parts*, CVPR, 2011, pp.1385-1392.
- [2] A. Fossati, M. Dimitrijevic, V. Lepetit, et al.: *From canonical poses to 3D motion capture using a single camera*, IEEE TPAMI, vol.32, no.7, pp.1165-1181 (2010).
- [3] J. Liu, B. Kuipers, S. Savarese: *Recognizing Human Actions by Attributes*, CVPR 2011, pp.3337-3344.
- [4] J. Shotton, A. Fitzgibbon, M. Cook, et al.: *Real-time human pose recognition in parts from single depth images*, CVPR 2011, pp.1297-1304.
- [5] D. Marr, H. Nishihara: *Representation and recognition of the spatial organization of three-dimensional shapes*, Proc. R. Soc. Lond. B 200, pp.269-294 (1978).
- [6] S. Nakaoka, A. Nakazawa, F. Kanehiro, et al.: *Learning from Observation Paradigm: Leg Task Models for Enabling a Biped Humanoid Robot to Imitate Human Dances*, International Journal of Robotics Research, Vol.26, No.8, pp.829-844 (2007).
- [7] Manoj Perera, 工藤俊亮, 白鳥貴亮, 池内克史: キーポーズを用いた舞踊動作の低次元化表現, 情報処理学会研究報告 グラフィクスと CAD 研究会報告, No.109, pp.19-24 (2008).
- [8] K. Ikeuchi: *Toward an Assembly Plan from Observation Part I: Task Recognition With Polyhedral Objects*, IEEE Transaction on Robotics and Automation, Vol.10, No.3, pp.368-385 (1994).
- [9] 天目隆平, 柴田史久, 田村秀行: 剣戟アクションの基本要素への分解と連続動作の合成 - コンピュータ殺陣学事始め, 情報処理学会論文誌, Vol.50, No.12, pp.2894-2899 (2009).
- [10] 中村美奈子, 八村広三郎: ラバノーテーション Labanotation とコンピュータテクノロジー-モーションキャプチャの舞踊教育と舞踊分析への応用, 舞踊学, No.24, pp.17-22 (2001).
- [11] 祝嶺正献・近藤光男: 剣道教範, 剣道本院出版局 (2005).
- [12] 宇野友季子, 山崎俊彦, 相澤清晴: 単眼動画画像を用いた大きな姿勢変化を伴う運動の種別判定, 電子情報通信学会総合大会, D-12-82 (2011).
- [13] iPi Soft: 入手先 (<http://www.ipisoft.com/>).

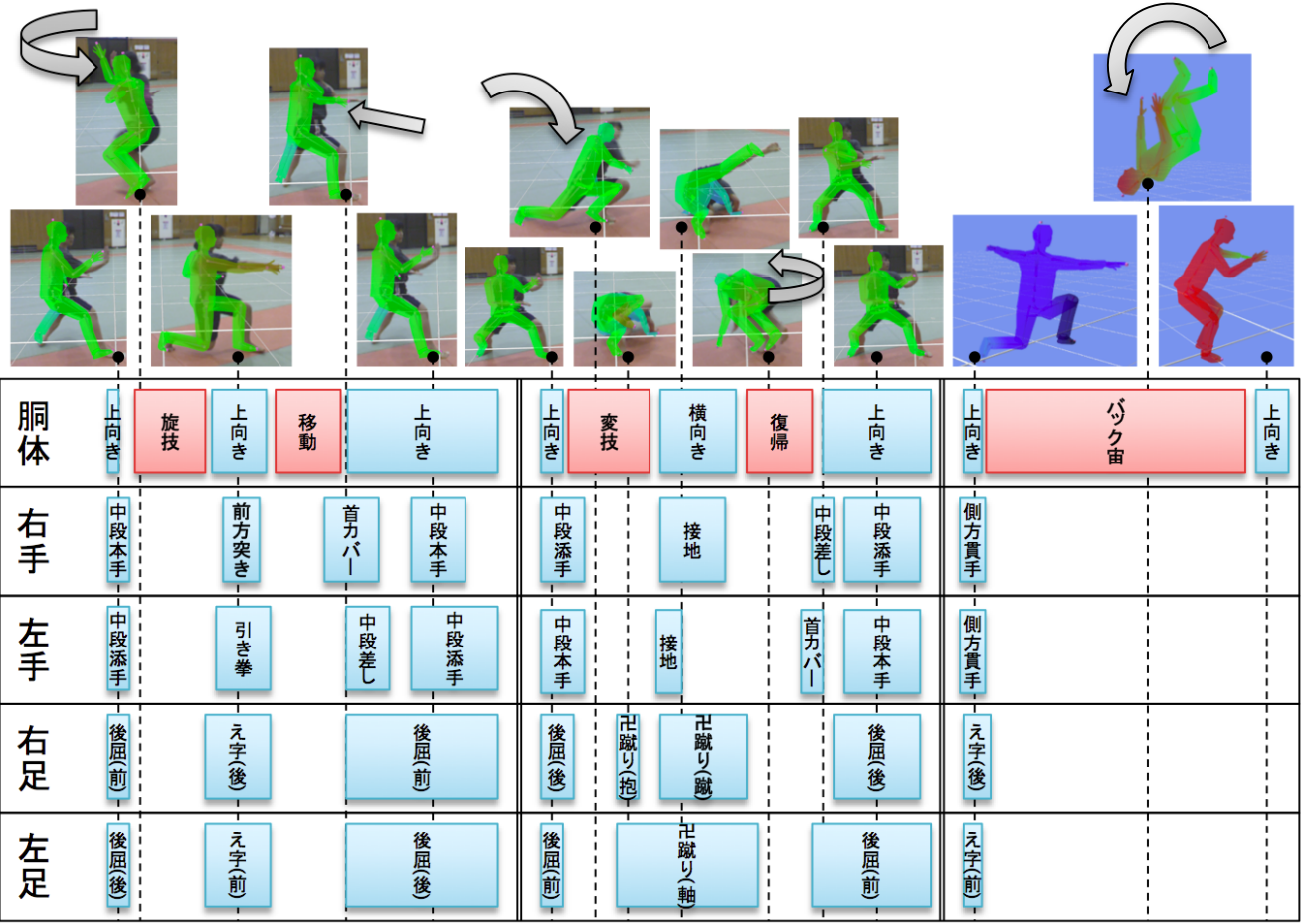


図 13 認識結果例
Fig. 13 Examples of Recognition Results.