

レガシーモーションデータの効率的活用のための 身体骨格構造の自動決定手法

Automatic Skeletally Structure Determination Methods for Efficient Use of Legacy Motion Data

湯川 崇

福島工業高等専門学校 ビジネスコミュニケーション学科

YUKAWA Takashi

National Institute of Technology, Fukushima College, Department of Business Communication

(2022年9月5日受理)

Legacy motion data refers to historically valuable motion data measured by motion capture devices that output obsolete format data without skeletally structure. In this paper, a method for estimating skeletally structure by analyzing motion data is proposed. Experiments on several data sets confirmed that the skeletally structure was properly estimated. The proposed method is expected to make efficient use of legacy motion data.

Key words: motion capture, motion data, human motion analysis, skeletally structure

1. はじめに

少子高齢化やそれに伴う過疎化により、全国各地に伝わる民俗芸能が失伝の危機にさらされている¹⁾。

人間の身体動作を記録する装置としてモーションキャプチャがある。これは人間の動きをセンサやカメラで撮影した画像を利用して位置や角度の時系列データとして記録する装置である。モーションキャプチャを利用した伝統芸能の記録の取り組みが各地で行われている^{2,3)}。

筆者が所属する民俗芸能情報技術研究会では1998年以来、モーションキャプチャを用いた民俗芸能の舞踊の記録・保存、および記録したデータの分析や再利用を目的とした活動を行なっている^{4,5)}。秋田県仙北市を拠点に活動している劇団わらび座を中心に、同劇団が所有するモーションキャプチャ装置を使用して2014年までの15年間で約200種類600テイクの舞踊等の身体動作のモーションデータを記録した。

これらのデータの中には、演技者が亡くなっていたり、高齢化のために記録した時点と同じように動くことができなくなっていたりと、撮り直しが不可能なため歴史的な資料として価値のあるものがあり、そのようなモーションデータを本研究ではレガシーモーションデータと呼ぶ。

現在普及しているモーションキャプチャ装置では、出

力するデータに身体骨格構造の情報も含まれる形式が一般的であるが、初期の装置により記録されたレガシーモーションデータには身体各部に取り付けたセンサの位置と角度の時系列情報のみが記録され、データには身体骨格構造に関する情報が含まれない。

そのため、データを利用するためには、各センサが記録したデータが身体の中のどの部位に対応するかをデータを取得したときの記録を調べたり、記録がない場合には点群のデータとしてディスプレイ上に表示して目視によりどの点がどの部位に対応するかを決定したりといった前処理が必要であり、解析作業の妨げになっている。最新の装置で記録し直すことは現実的ではないため、記録済みのデータを効率よく解析する手法の確立が望まれている。

本研究では、レガシーモーションデータの解析作業の効率化を目的として、センサと身体部位の対応付けを自動的に行うことで自動的に身体骨格構造を推定する手法を提案する。

2. モーションデータ

モーションキャプチャ装置は身体の動きを記録する方式により光学式、慣性式、磁気式、機械式に分けられる。いずれの方式でも人間の動作を、身体各部の位置と角度の時系列データ（モーションデータ）として出力する。

古い形式のモーションデータは、データに身体骨格構造の情報は含まれず、**Fig. 1(a)**に示すような各センサの位置と角度の時系列データが列挙されるのみである。何番目のセンサがどの部位に対応するかには決まりはなく、この形式のデータを利用するためには、手作業によりセンサと部位の対応付けを行い、さらに部位に取り付けたセンサの情報を関節の位置や角度に変換する前処理が必要である。

現在普及している装置ではこれらの処理はモーションキャプチャの受信装置内で行われ、**Fig. 1(b)**に示すような身体骨格構造を含んだBVHやFBXなどの形式でデータが出力される。

本研究では、光学式や慣性式の装置が普及する以前に位置と角度の同時計測が可能な方式として使われていた磁気式の装置が出力するセンサ毎の位置と角度の時系列データを対象とする。前提条件として最大15センサで、上腕または前腕か大腿または脛のいずれかまたは両方が欠落していてもよいこととする。これは、前腕のような2つの部位（上腕と手）に挟まれている部位のデータはそれを挟んでいる部位のデータから推定可能であり⁹⁾、センサ数の制約等により省略されることがあるためである。

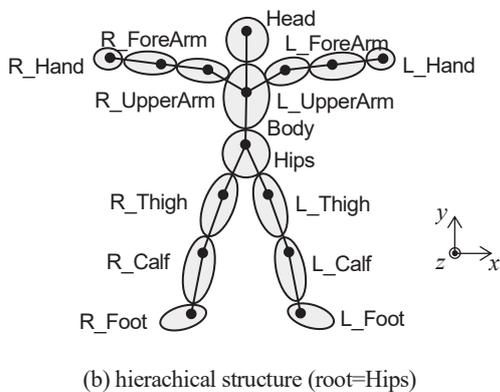
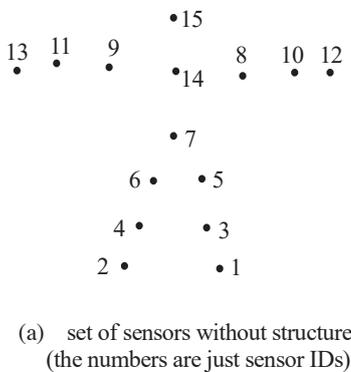


Fig. 1 Hierarchical structure of motion data

3. 提案手法

モーションデータの身体骨格構造となる各センサがどの部位に対応するかの解析は、**Fig. 1(b)**のセンサを表す黒丸をノード、黒丸同士を接続する直線をエッジとするグラフを利用して、次の9つの手順により行う。

手順1 グラフの作成と不要なエッジの削減

各センサをノードとして、全てのノード間を連結した重み付きグラフを作成する。ノード間のエッジの重みはセンサ間の平均距離とする。

ノード数を n とすると、最終的に推定する身体骨格構造が含むエッジ数である $n - 1$ に比べ初期のグラフのエッジ数が $(n^2 - n)/2$ と多いため、明らかに無関係と思われるエッジを削除する。これは、上腕と前腕、前腕と手など隣り合っている部位に取り付けたセンサ間の距離の変動が、それ以外の部位に取り付けたセンサの間との距離の変動より少ないことを利用し、センサ間の距離の分散が全体の分散の平均より大きいエッジを削除することにより行う。

Fig. 2に説明用のサンプルデータを用いて作成した初期のグラフを示す。各センサに対応するノードをS1, S2, ...として表している。なお、**Fig. 8**と**Fig. 10**を除く**Fig. 2**以降のグラフはノード同士の接続関係のみを表しており、ノードの位置やエッジの形状と実際の姿勢とは無関係である。

手順2 頭部の推定

直立した人間の姿勢では頭部が最上部に位置するため、各センサの y 座標の全フレームでの平均値をそれぞれ求め、最も大きな値を持つ（最上部にある）センサに対応するノードを頭部とする。

Fig. 3では黒丸印で表しているノードS13の y 座標の平均値が最も大きいため、頭部と推定している。

手順3 胸部の推定

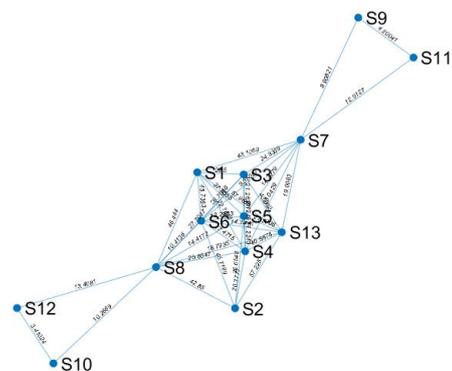


Fig. 2 Graph of initial state

頭部と連結している部位は胴部のみであるため、頭部から最も近い位置にあるノードを探し、それを胴部とする。

Fig. 4 の例では白丸印で表しているノード S6 が頭部に最も近いので、胴部と推定している。

手順 4 腰部の推定

胴部には頭部、左右腕部および腰部の合計 4 つのノードが連結するため、胴部からのエッジのうち平均距離が近い 4 つを残し、それ以外のノードを削除する。その後でそれら 4 つのノードの y 座標の平均をそれぞれ求め、最も小さい（一番下に位置する）ノードを腰部とする。

Fig. 5 の例では胴部のノード S6 には S1, S2, S3, S4, S7, S8 および S13 のノードが隣接しているが、そのうち S5, S8, S7 および S13 が近いので、点線で表されるそれ以外のノードに接続されているエッジを削除する。またノード S5 が一番下に位置するので腰部と推定している。

手順 5 不要なエッジの削除

手順 4 までで身体の根幹にあたる頭部、胴部、腰部が推定された。手足の推定に進む前に Fig. 1(b) に示す身体骨格構造をヒントに不要なエッジを削除する。

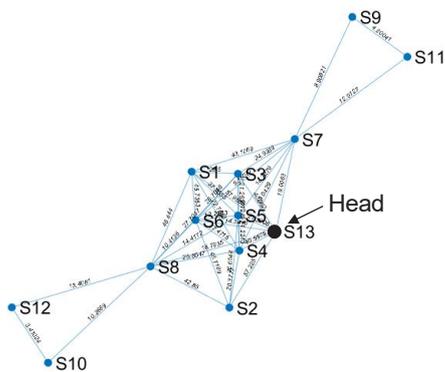


Fig. 3 Estimation of head

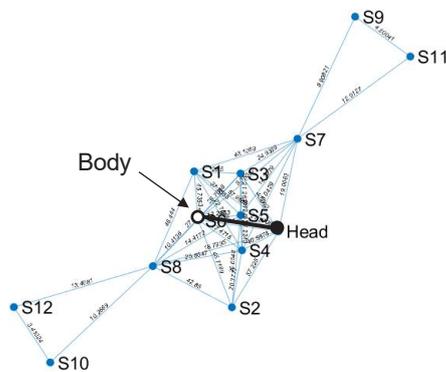


Fig. 4 Estimation of body

(a) 腰部は胴部と左右足部（大腿または脛）のみと連結するため、腰部から出るエッジのうち隣接するノードとの距離が近い 3 つ以外のエッジを削除する。

Fig. 6(a) の例では腰部のノードには胴部につながるエッジを除いて 4 本のエッジがある。これは大腿部または脛部のいずれかが省略されていることを表しており、大腿部が省略されている場合は腰部→脛部→足部の順、脛部が省略されている場合は腰部→大腿部→足部の順にノードが連結される。そのため点線で表す腰部から足部に向かうノード Hip と S1 を連結するエッジおよびノード Hip と S2 を連結するエッジをそれぞれ削除している。

(b) 左右の大腿や足同士をつなぐエッジが残っていることがあるため、腰部より下でかつ次数が 3 以上のノードがある場合、下に位置するノードに連結されているエッジ以外のエッジを削除する。

Fig. 6(b) の例では点線で表しているノード S3 と S4 を連結するエッジが左右の大腿部または脛部同士を連結しているのを削除している。

(c) 身体骨格構造はループを持たないため、三角形を構成するエッジを探し、三辺のうち一番距離の長いエッジを削除する。

Fig. 6(c) の例では、左右の上腕、前腕、手を表すノード S8, S10 および S12 ならびに S7, S9 および S11 で、点線で表している上腕と手を連結している S7 と S11 を連結するエッジおよび S8 と S12 を連結するエッジをそれぞれ削除している。

手順 6 手足の推定

全てのノードの中で次数が 1 のノードは身体の末端にあたる頭部、左右手部および左右足部だけであるので、頭部以外で次数が 1 のノードを求め、高さの平均値を比較し高い方を手部、低い方を足部とする。

Fig. 7 の例では次数が 1 のノードは既に推定済の Head

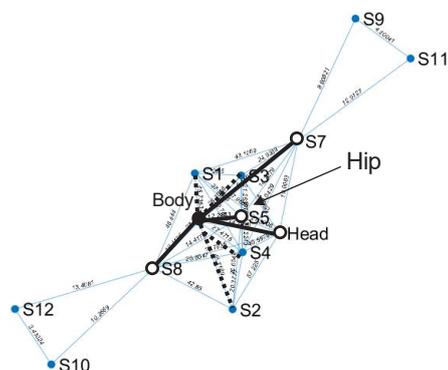
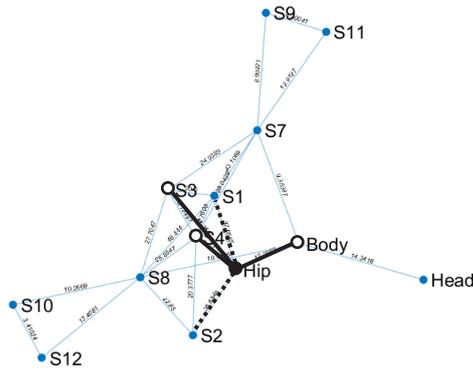


Fig. 5 Estimation of hips

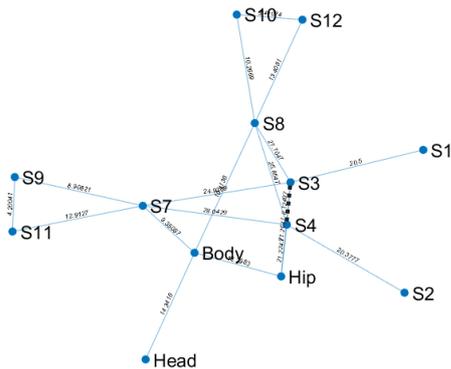
の他, S1, S2, S11 および S12 の 4 つがある。これらは左右の手部と左右の足部に対応するため, 高さの平均を求め, 高い方の S11 と S12 を手部, 低い方の S1 と S2 を足部と推定する。手足の左右は次の手順で推定する。

手順 7 左右の推定

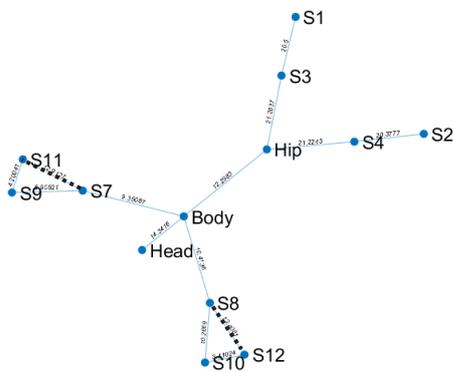
動作中の手や足は平均的に腰よりも前にあると仮定し, 上から見た腰部と 2 つの手部 (足部) との位置関係を調



(a) unnecessary edges around the waist



(b) unnecessary edges connecting left and right sides



(c) unnecessary loops

Fig. 6 Deleting unnecessary edges

べて左右を推定する。これは次の手続きにより行う。

横断面 ($x-z$ 平面) 上での腰の座標を s_0 , 左右が不明な 2 つのノード s_1, s_2 の座標をそれぞれ s_1, s_2 とする。 s_0 を中心に s_1 が x 軸の正方向に一致するように s_2 を回転させた後の s_0 からの相対座標 s'_2 を(1)式により求める。

$$s'_2 = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} (s_2 - s_0) \quad (1)$$

ただし, θ は $s_1 - s_0$ の x 座標を s_x , z 座標を s_z としたとき, 多くのプログラミング言語で標準ライブラリ関数として提供されている atan2 関数を使って $\theta = \text{atan2}(s_z, s_x)$ として求められる値である。

このとき, s'_2 の z 座標の値が 0 より小さいならば s_1 が左で s_2 が右と推定し, 逆に s'_2 の z 座標の値が 0 以上ならば s_1 が右で s_2 が左と推定する。

Fig. 8(a)に s_1 が左で s_2 が右の例を, (b)に s_1 が右で s_2 が左の例を示す。回転前の s_1 と s_2 を黒丸, 回転後の点 s'_1 と s'_2 を白丸でそれぞれ表している。

手順 8 腕部の推定

胴部と手部の間のノード数を調べる。

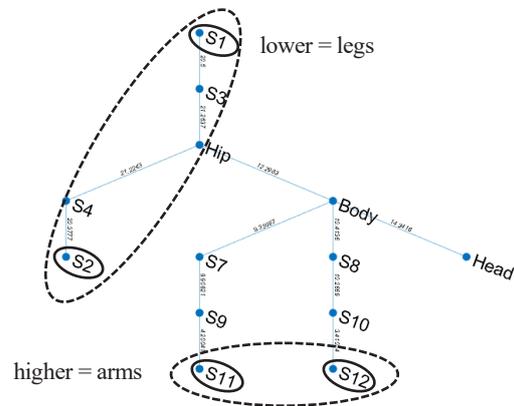
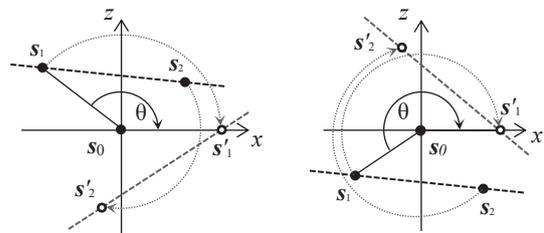


Fig. 7 Estimation of arms and legs



(a) s_1 is left and s_2 is right

(b) s_1 is right and s_2 is left

Fig. 8 Identify right and left

ノードが1つの場合は上腕または前腕のいずれかのノードがないことを表している。そのノードが胴部に近ければ上腕部、手部に近ければ前腕部と推定する。

ノードが2つの場合は胴部に近い方を上腕部、手部に近い方を前腕部と推定する。

手順9 脚部の推定

腕部と同様に腰部と足部の間のノードの数により大腿部と脛部を推定する。

ノードが1つの場合はそのノードが腰部に近ければ大腿部、足部に近ければ脛部と推定する。ノードが2つの場合は腰部に近い方を大腿部、足部に近い方を脛部と推定する。

推定の終了

以上の手順を実行することにより、全てのノードと部位名の対応付けを行うことができる。

Fig. 9 にサンプルデータでの推定結果を示す。このデータは腰部と足部の間のノード数が1であり、脛部のセンサが省略されていることが分かる。

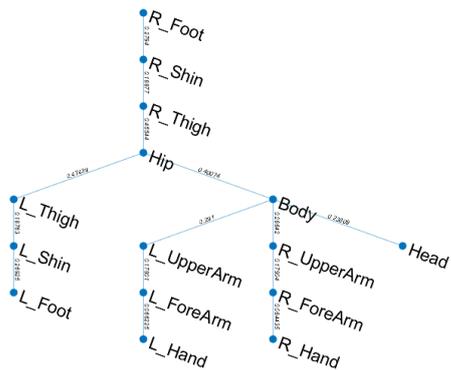


Fig. 9 Estimated body structure results

Table 1 Experimental results

Title	Sensors	Frames	Calculation time [s]
Kakumagawa (A)	11	2,980	0.0506 0.0468
Kakumagawa (B)			0.0523 0.0469
Kakumagawa (C)			0.0492 0.0460
Karate Kata (A)	13	2,702	0.0584 0.0523
Karate Kata (B)			0.0587 0.0525
Karate Kata (C)			0.0587 0.0533
Nishimonai (A)	15	4,576	0.0504 0.0540
Nishimonai (B)			0.0540 0.0498
Nishimonai (C)			0.0503 0.0508

4. 評価

提案する手法を評価するため、記録した日時、演技者そしてデータ中のセンサ数が異なる3種類のモーションデータを用いて実験を行った。実験に使用した機器はApple MacBook Air (CPU: Apple M1 3.2GHz, メモリ: 8GByte, OS: macOS Ventura 13.0.1) である。身体骨格構造推定プログラムは MATLAB R2022b を用いて作成した。

Table 1 に実験に使用したデータの諸元および実験結果を示す。3種類のうち、2種類は秋田県大仙市角間川町と同県雄勝郡羽後町にそれぞれ伝わる盆踊り

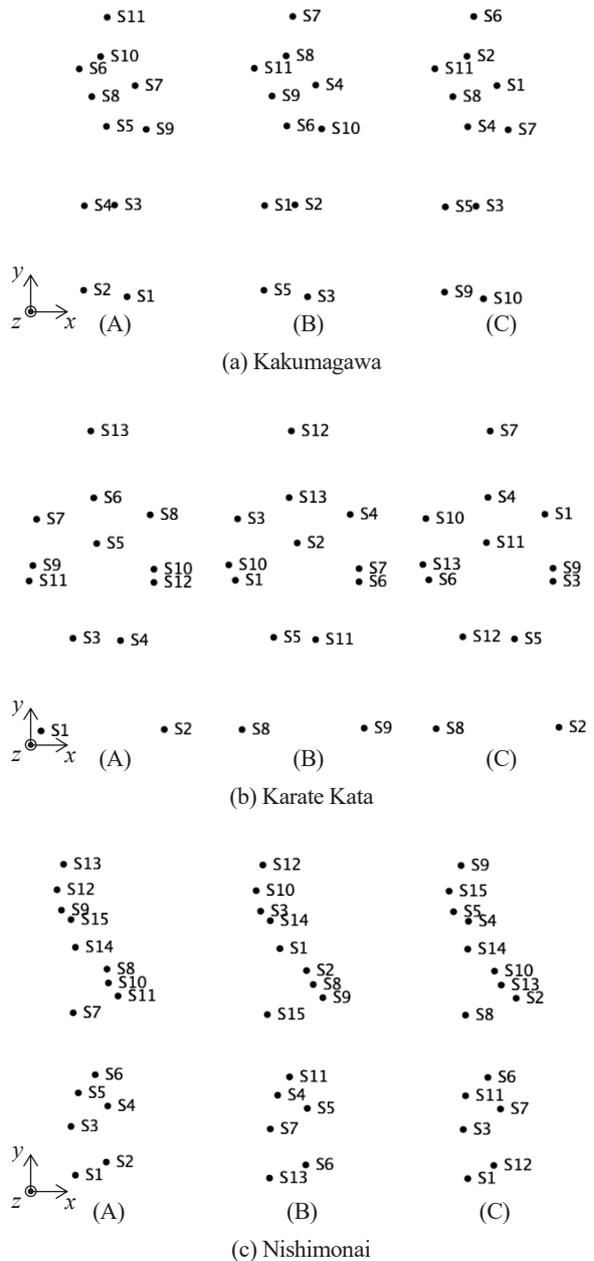


Fig. 10 Placement of sensors in experimental data

(Kakumagawa, Nishimonai と表記) を、もう 1 種類は空手の形の演技 (Karate Kata と表記) を記録したデータである。3 種のデータのそれぞれについて、記録されたオリジナルのデータ (動作名が (A) で終わるもの)、およびデータファイル中のセンサ順を 2 通りにランダムに並べ替えたデータ (動作名が (B), (C) で終わるもの) を使用した。Fig. 10 に Fig. 1 と同じ座標軸を用いて表示した各データのセンサ順の配置図を示す。一連の動作の中でセ

ンサのラベルが重ならずに表示できるような姿勢を選んでいる。

それぞれの動作データについてデータ全体とデータの開始直後の 100 フレームのみの 2 通りについて実験を行った。全てのノードへの部位名の割り当てが終了した後のグラフを Fig. 11~13 に示す。同じ動作データではセンサの出現順が異なっても互いに同形なグラフが生成されており、部位名が適切に割り当てられていることを

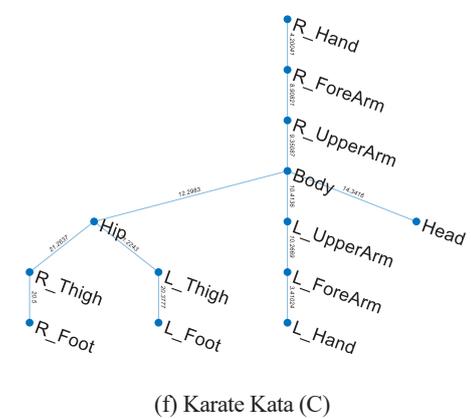
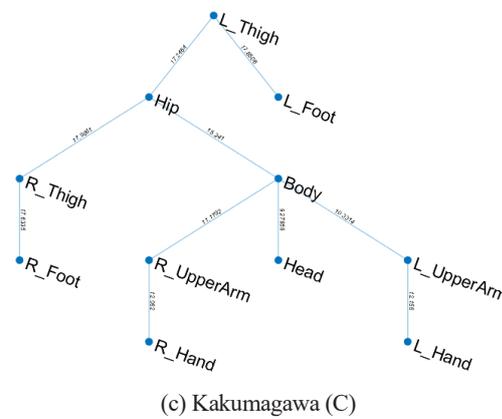
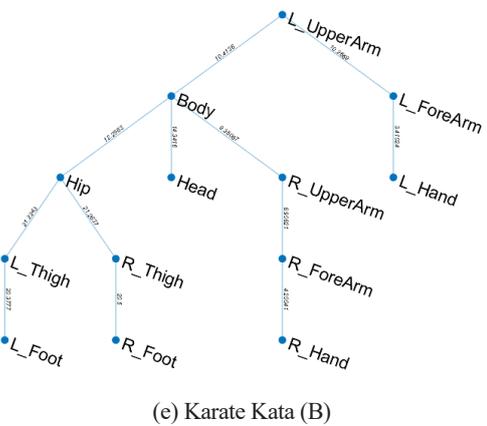
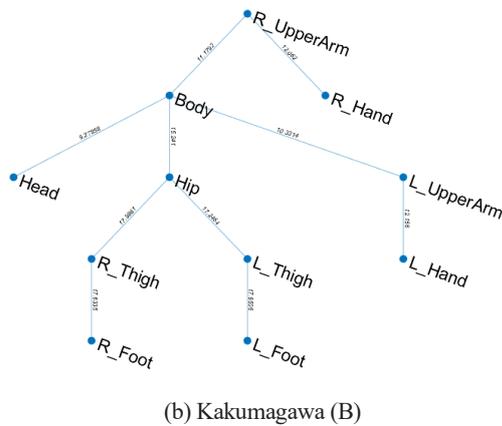
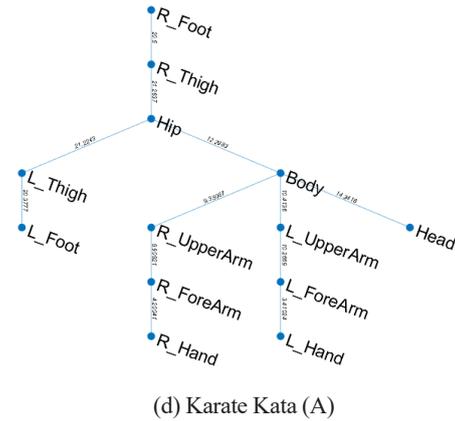
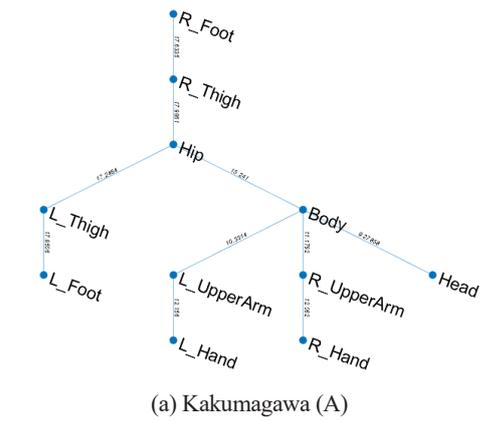
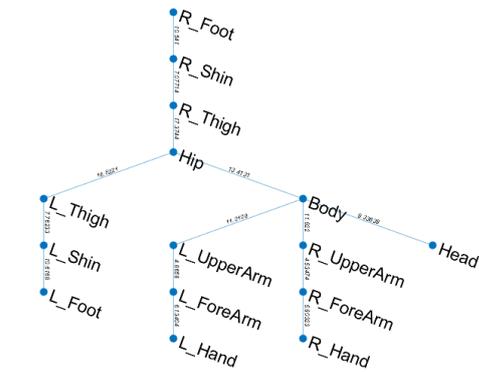


Fig. 11 Graph with part names assigned to all nodes

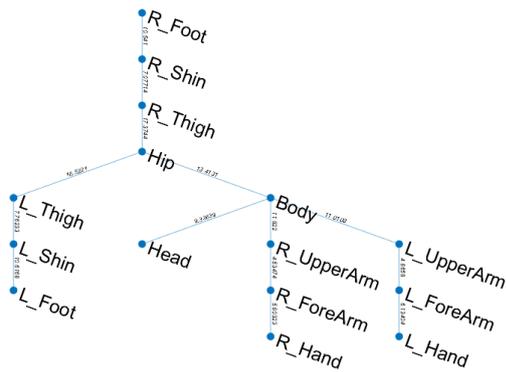
Fig. 12 Continuation of Fig. 11

表している。

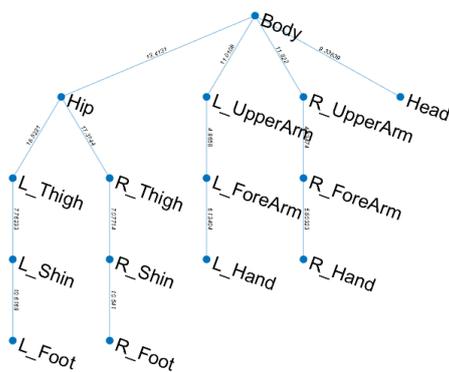
なお、推定に要した時間を Table 1 に示した。| の左側がデータ全体を用いたとき、右側が初めの 100 フレームのみを用いて実行したときの処理時間を表している。いずれも 0.1 秒以内に計算が終了しており、解析用のプロ



(g) Nishimonai (A)



(h) Nishimonai (B)



(i) Nishimonai (C)

Fig. 13 Continuation of Fig. 12

グラムに組み込んで利用しても、データ解析の時間に与える影響は少ないと考えられる。

また、今回実験に使用したデータは全て正立した状態での動作であるが、横臥状態が多い動作であったとしても記録開始時に正立した姿勢の動作が記録されていれば、本手法を適用することができる。

5. おわりに

本研究では、旧式のモーションキャプチャ装置を用いて記録された身体骨格構造を含まないモーションデータを解析し、自動的に身体骨格構造を推定する手法を提案した。実験により、モーションデータファイル中のセンサのデータの並び順に関わらず、適切に部位名を推定できることを示した。

これまで身体骨格構造を含まない古い形式のモーションデータを利用するためには、ファイル中のセンサのデータがどの部位に対応するかを手作業により決定する必要があり、このことがデータ分析作業の障害になっていた。本手法を利用することで、価値のあるレガシーモーションデータの解析の効率化に役立つことが期待できる。

参考文献

- 1) 星野紘: 過疎地の伝統芸能の苦闘, 無形文化遺産研究報告, 国立文化財機構東京文化財研究所(5), pp.29-39, 2011.
- 2) 八村広三郎: モーションキャプチャによる舞踊のデジタルアーカイブ, 情報処理学会研究報告, Vol.2007, No.1, pp. 1-8 2007.
- 3) 玉本英夫: 民俗芸能・伝統芸能をモーションキャプチャで記録する, 文部科学省 私立大学研究ブランディング事業 日本遺産を誇る山形県庄内地方を基盤とした地域文化とIT技術の融合による伝承環境研究の展開, 東北公益文科大学, pp. 23-29, 2019.
- 4) 玉本英夫, 湯川崇, 海賀孝明, 水戸部一孝, 三浦武, 吉村昇: 産学官連携による民俗芸能伝承のためのデジタルコンテンツ制作技術の開発, 電子情報通信学会誌, Vol.91, No.4, pp.303-308, 2008.
- 5) 三浦武: 民俗芸能情報技術研究会の活動, 電気学会誌, Vol.140, No.11, pp.736-736, 2020.
- 6) Takeshi Miura, Takaaki Kaiga, Katsubumi Tajima, Takeshi Shibata, Hideo Tamamoto: Sensor Number Reduction in Skeleton Estimation from Magnetic Motion Capture Data, *Journal of Information Processing*, Vol.23, No.5, pp.704-707, 2015.