

キーモーションの入力による モーションキャプチャデータの検索手法

石川 知一^{1,a)} 川島 基展¹ 柿本 正憲¹ 近藤 邦雄¹

概要: モーションキャプチャ技術は、高品質なアニメーション作成やバーチャルリアリティの表現などに幅広く利用されている。制作過程において、データの再利用を行うことができれば効率的にアニメーション開発を行うことができる。本研究ではデータの再利用の補助として、Kinect を利用した直感的なモーションキャプチャデータの検索方法を提案する。Kinect ではデータの取得時に欠陥部位が存在することがあるが、各部位の信頼度を考慮し検索することで情報の欠陥を補う方法を提案する。この方法は Kinect による取得データに対してのみでなく、ユーザーが予め各部位と演技に時間軸に対して信頼度を設定することで、ユーザーの身体能力や演技力を補うことに利用できる。提案法によって言語化することが難しいデータや、アクロバティックな動作データをもインタラクティブに抽出することを目標とする。

キーワード: モーションキャプチャ, データマイニング, Kinect, ランダムフォレスト

Retrieval Method for Motion Capture Data by Inputting a Key Motion

TOMOKAZU ISHIKAWA^{1,a)} MOTONOBU KAWASHIMA¹ MASANORI KAKIMOTO¹ KUNIO KONDO¹

Abstract: Motion-capture technology is widely utilized in such fields as high-quality animation production and virtual reality representation. If the users could reuse the motion-capture data sets, they would be able to create character animation more efficiently. We propose a method for intuitive search for motion capture data using a Kinect input device in order to provide an auxiliary means for data reuse. While Kinect may produce one or more defective joint sites in capturing motions, our method defines a varying degrees of confidence for each joint node, searches the database considering the set of degrees of confidence, and thus compensates for the deficit. By defining a series of time-variant degree of confidence, it is also possible to use it for compensation for the lack of the performer's physical or acting ability. A goal of our method is to interactively extract data for acrobatic motions or data sets that are hard to verbalize.

Keywords: motion capture, data mining, Kinect, random forest

1. はじめに

高品質なアニメーション制作のために、モーションキャプチャシステムにより測定された人体動作データが利用されている。モーションキャプチャシステムの設備は高価であり、アクターの起用も考えると、容易に撮影できるとは言えない。現在では、代表的な人体動作データが Web 等

を介して提供されている。こうして提供される膨大なデータ群は内容の説明が付与されており、利用者は検索手法には言語を用いることになるが、この検索方法は直感的とは言えない。

また、Kinect [1] の登場により、モーションキャプチャは身近なものになり、個人でも 3D モデルに対してアニメーションを割り当てることが容易になっている。Kinect はゲームに利用するためには十分なスケルトン情報を取得できるが、時間的な分解能が低かったり、取得できる関節の数が少なかったりするため、そのまま別のモデルのアニメーション

¹ 東京工科大学
Katakura-Cho, Hachioji-Shi, Tokyo 192-0914, Japan
^{a)} ishikawatm@stf.teu.ac.jp

メーションに割り当てると不自然な動きが残ってしまう。

そこで本手法では、Kinect によって取得した情報をもとに、高解像度な人体動作データを検索する方法を提案する。これにより、大規模な人体動作データのデータベースから、直感的に必要な人体動作データを検索することができ、Kinect では不足するような人体動作データを取得し割り当てまでの補助を行うことができる。提案法により、個人ユーザーだけでなく、制作現場において、データを再利用する場合の一助となることを目的とする。

キーとなるモーションを入力する場合、検索したい人物が同じような演技ができるとは限らないため、アクロバティックなモーション検索ができないことが問題になる。本手法では、動作が完全に一致するような検索技術を提案するわけではなく、検索時に曖昧にしたい部位や時間を指定することで、検索者の演技力や身体的制約をカバーすることを考える。

以下、2 節で人体動作データの検索手法における既存手法を述べた後、3 節で提案法で採用する特徴量の計算と検索手法について紹介する。次に 4 節で提案手法について説明し、5 節でその結果について述べ、6 節ではまとめと今後の課題について述べる。

2. 既存手法

人体動作データの検索方法には、Web で配信されている [2] [3] ように、テキストデータを用いる方法 [4] も提案されているが、動作の解釈は撮影時に付与したものと、検索者では異なることが考えられるし、直感的な方法ではない。提案する検索方法は、動作から動作を検索する方法であるため、この観点から既存手法を述べる。

2.1 データベース内の検索手法

キーとなる人体動作データを入力として、データベース内を検索する方法は従来から提案されている。Müller らは、関節間の距離や関節の角度、関節の位置のような幾何的な情報から関連性を検索する手法を提案している [5]。浜田らは、この研究を踏まえ、ランダムフォレスト法を用いた検索手法を提案した [6]。この方法により、既存データについて精度良く検索可能であることが示されており、本手法でも検索についてはこの手法を基盤として実装する。しかし、これらの研究では、人体動作データの分類が主な目的である。キーとなる人体動作データはデータベース内に含まれており、新規の人体動作データについては検証されていない。

Chai と Hodgins は、モーションキャプチャ時にマーカの数が少なくても、高解像度の人体動作データを検索できることを示した [7]。これは Kinect のような個人用モーションキャプチャシステムにおいて有用であるが、検索対象は類似データが集まったデータベースに限定されるという欠

点がある。

竹林らは、人体動作データを入力として可動領域の特徴点の分散を考慮し、Müller らとは異なる特徴量の計算を提案した [8]。この特徴量は詳細度を設定することができるため、絞込み検索が可能である。しかし、この手法は部分的な姿勢情報を見ているため、一部動作には検出ミスが発生することがある。

提案法では、Kinect を用いて取得した新規の人体動作データをキーとして、さまざまな人体動作データを含むデータベース内のデータ検索を目標とする。

2.2 入力デバイスによる検索手法

Numaguchi らは、人形デバイスを利用した検索手法を提案した [9]。人形デバイスからは、センサーを利用して関節の角度や速度を計算することができ、これらの情報を利用した直感的な検索手法を提案している。また、検索者の身体能力に依存しない手法として有用で、アクロバティックな動作を検索するためにも有効である。この手法は、デバイスの限界として、平行移動の検出ができないことが挙げられている。同様の人形型入力デバイスに QUMARION があるが [10]、こちらも平行移動についてのデータを取得するのは難しく、連続的な時系列データの取得も困難である。

我々は、検索者が身体動作を行って入力する方法であっても、アクロバティックな動作を検出可能な検索手法を提案する。

3. 検索用数値データの計算と検索手法

我々は Kinect for Xbox 360 を用い、OpenNI [11] により人体動作データの取得を試みる。この節では Kinect によって取得できる情報と、その情報から検索用数値データを計算する方法を述べる。我々は検索精度が良いと報告されている浜田らの方法 [6] を参考に、実装を行う。まず、人体動作データから検索用の数値データの計算方法について述べ、その数値データを用いた検索手法としてランダムフォレスト法について説明する。

3.1 Kinect による人体動作データの取得

OpenNI によって取得できる人体の関節を図 1 に示す。Kinect で得られる情報は、モーションキャプチャスタジオで撮影するデータと比べて、以下のような欠点がある。

- 関節の数が少ない
- 時間的分解能が粗い
- 検出ミスが発生しやすい

検出ミスが発生する原因は self-occlusion であったり、必要な関節がカメラの外に出てしまうことなどが考えられる。この解決策としては、Kinect を複数台設置することが考えられる。本手法では、検出ミスは検索手法の中でアプ

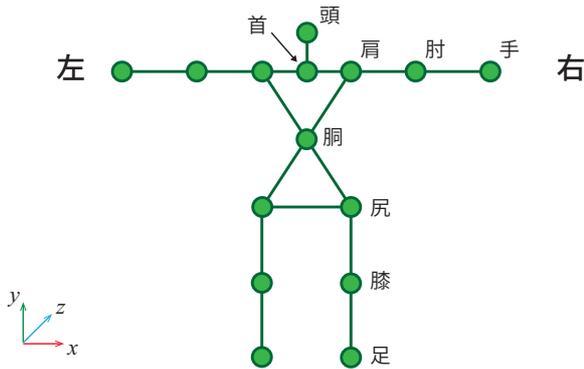


図 1 OpenNI によって取得できる人体の関節

Fig. 1 Joint data of the human body obtained by OpenNI

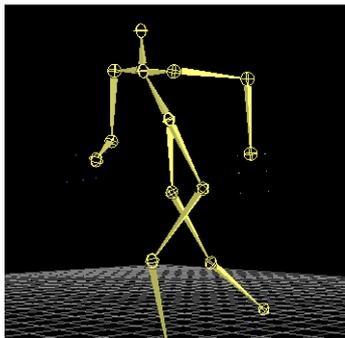


図 2 MotionBuilder によって Kinect と同じ関節を持つスケルトンデータに変換した結果

Fig. 2 A result of the skeleton data converted to the same as Kinect joints by using MotionBuilder

リケーション側でカバーするため、Kinect は一台で実験する。

検出ミスによってボーンが取得できなかつたり、信頼度の低いデータが得られることがある。OpenNI では、ボーンが取得できない場合や、そのボーンの信頼度の情報を取得する関数が提供されている。この情報を利用して、検索する方法を提案する。

Kinect で取得できる関節の位置情報は、グローバル座標系になっている。対して、Web 等で提供されている人体動作データは一般に関節角と関節間の長さの情報であることが多い。また、Kinect で得られる胴体部分の関節は一つのみであるのに対して、MotionBuilder などの商用のアニメーション作成ツールでは二か所を扱うことが多い。本手法では、既存の人体動作データから MotionBuilder を用いて Kinect と同じ関節を持つスケルトンデータに変換したものと、Kinect から取得したデータとを比較する方法を取る。既存の人体動作データをスケルトンに割り当てた結果を図 2 に示す。

3.2 検索用数値データの計算

人体動作データは演技者の体格によって差異が生じるため、Kinect によって取得するデータ、既存の人体動作

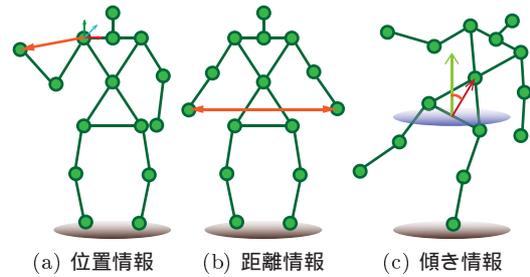


図 3 姿勢表現に用いる相対情報

Fig. 3 Relative variables for pose representation

表 1 本手法で利用する検索用数値データ

Table 1 Numerical data for retrieval in our method

位置情報	左(右)手	左(右)足	頭	
距離情報	左右の手	左右の足	左右の肘	左右の膝
傾き情報	胸	頭	左(右)手	左(右)足

データともに各データの身長を 1 とした長さの正規化を行う [6]。ここで、身長とは直立姿勢における頭の関節と両足の関節の y 座標の差とする (図 1 参照)。

特徴量はある時間間隔 $[t_0, t_1]$ における姿勢情報の平均値をとる [12]。ある時刻 t における i 番目の関節姿勢情報を $p_i(t)$ とすると、平均姿勢情報 \tilde{p}_i は

$$\tilde{p}_i = \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} p_i(t) dt \quad (1)$$

で計算される。

ここで姿勢情報とは、位置情報、距離情報、傾き情報に分類される。図 3 は計算する特徴量の概念図である。

位置情報は、特定の関節の回転中心を原点としたときの、注目する別の関節の x, y, z の成分である。例えば、右肩の関節を中心としたときの、右手の座標の各成分のことである (図 3(a) 参照)。

距離情報は、2 つの関節間のユークリッド距離である。例えば、右手と左手の関節間距離のことである (図 3(b) 参照)。

傾き情報は、 y 軸の正の向きと各関節のなす角度で表現する。図 3(c) では、胸と y 軸のなす角度を示している。

本手法の中では、Kinect で取得できるデータに合わせて、表 1 の情報を検索用数値データとして利用する。

3.3 データベース内の検索手法

検索手法には集団学習法の一つであるランダムフォレスト法を用いる [13]。ランダムフォレスト法を用いることで、高精度かつ高速な検索システムを構築することができる [6]。図 4 は検索手法の概念図である。ランダムフォレスト法における一般的な識別の方法を以下に示す。

Step1. 前処理としてデータ集合 A から、重複を許してランダムに B 個のサブセットを作成する。

Step2. サブセットの数 B だけ決定木を作成する。各決

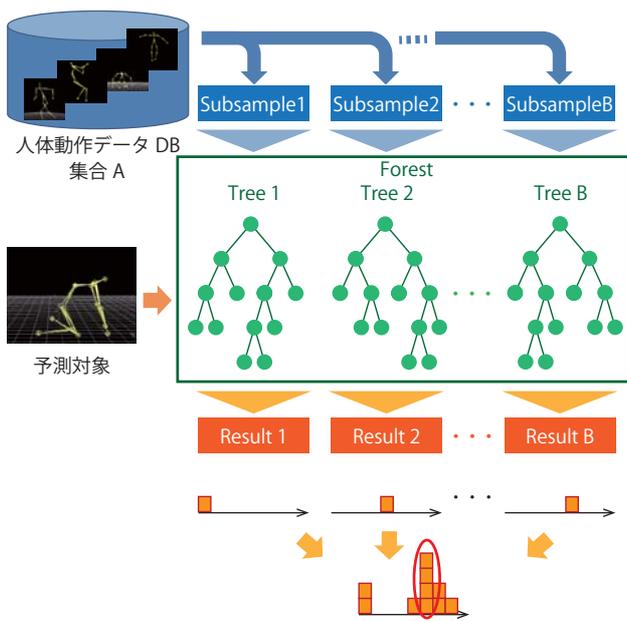


図 4 ランダムフォレスト法による検索
Fig. 4 Retrieval by random forest

決定木を作成する際、分岐ノードの候補を m 個ランダムに選択し、サンプルを分割した後で情報利得を算出する。最も情報利得が高い候補を選択して分割する。情報利得が 0 またはノードが 1 になった場合に末端ノードとする。

Step3. 予測対象を B 本の決定木にかける。それぞれに対する結果を多数決を取ることで識別結果とする。

4. 提案法

我々は、Kinect を用いて類似した人体動作データを検索する方法を提案する。Kinect によって取得できるデータは欠陥部位が含まれていることがあり、信頼度が低いものも含まれるため、取得データを決定木にかけて得られた識別結果をそのまま信用できない。そこで、決定木に対して信頼度を付与した重み付き識別結果を用いて、Kinect による取得データの信頼度に応じた検索を提案する。

この信頼度は、欠陥部位などの Kinect によって検出されるものに限らない。ユーザーが予め各部位の信頼度を設定しておくことにより、人体動作を取得する場合に問題になる検索者の身体能力を補うことができる。例えば、ハイキックの場合、体の柔軟なアクターによって足が高く上がっている人体動作データが既に存在しているかもしれないが、検索者は身体能力により足が高く上がらないケースが考えられる。この場合高く上げる足の信頼度をあらかじめ下げることによって、検索範囲を広げることができる。

信頼度は部位に対してだけでなく、時間軸に対しても設定することで、よりアクロバティックな動作を検索することができる。バク宙の動作は、バク宙に入る前のモーシ

ョンと着地後のモーションを入力とすることで検索結果の候補に含むことができる。この場合、実際に宙返りしている時間をあらかじめ指定し、その時間帯を省くことで検索範囲を広げる。

人体動作を直接取得する方法でも、あえて曖昧を残して検索することで検索の幅を広げることができる方法を提案する。

4.1 前計算

Kinect による人体動作データの取得前に、ランダムフォレスト法で使用する決定木を計算する必要がある。ユーザーは以下の項目を事前に設定することができる。

- 各部位の信頼度
- 時間軸に対する信頼度
- Kinect によってデータを取得する時間

信頼度は $0 \sim 1$ の値で設定し、1 が信頼度が高いものとする。

まず、既存の人体動作データの演技時間は一定ではないため、「Kinect によってデータを取得する時間」の長さによってデータをトリミングする。トリミングの開始地点はランダムとして、長さ L のデータから指定した時間 T を分割する数 N は、以下の式で決定する。

$$N = \alpha \left\lceil \frac{L}{T} \right\rceil \quad (2)$$

ここで、 $\lceil \cdot \rceil$ はガウス記号で、その値を超えない最大の整数を表す。また、 α はユーザー指定の分割定数で、1 以上の整数値とする。

次に、一定間隔に分割した人体動作データから、平均姿勢情報を計算する。ここで「時間軸に対する信頼度」を考慮する。図 5 のように、分割後の時間間隔 $[t_0, t_1]$ 中に信頼度 w_j の時間区間 $[s_{2j}, s_{2j+1}]$ が k 個含まれているとすると、平均姿勢情報 \tilde{p}_i は以下の式で計算される。

$$\tilde{p}_i = \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} p_i(t) dt - \sum_{j=0}^{k-1} \frac{1 - w_j}{s_{2j+1} - s_{2j}} \int_{s_{2j}}^{s_{2j+1}} p_i(t) dt \quad (3)$$

ここで、 $p_i(t)$ は時刻 t における i 番目の関節姿勢情報である。各部位の信頼度で、事前に設定されたものは決定木を作成する段階で考慮に入れる。すなわち、3.3 節のランダムフォレスト法の中で、説明変数をランダムにサンプリングするプロセスで、信頼度に応じたサンプリングを行う。

4.2 Kinect で取得した姿勢情報の信頼度

Kinect の精度の問題によって取得データには欠陥部位や信頼度の低いデータが散在することが考えられる。検索者があらかじめ設定した信頼度とは別に、Kinect で人体動作データを取得する時に、入力各部位の信頼度を累計する。欠陥部位である場合は該当部位の信頼度を 0 とする。 i 番

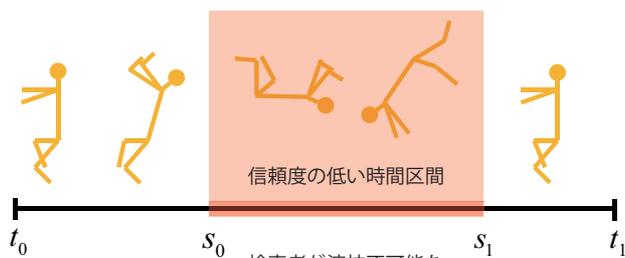


図 5 時間軸に対する信頼度
Fig. 5 Confidence in a timeline

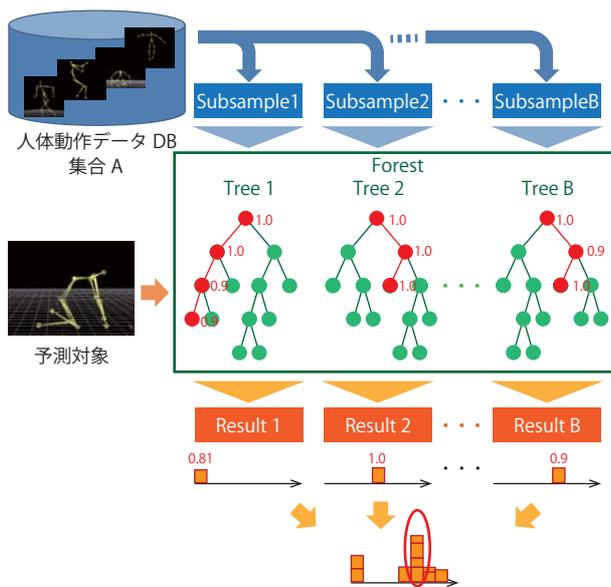


図 6 トラバース時に各ノードで記述する情報の信頼度を考慮
Fig. 6 Take into account the confidence of the information that describes each node when traversing

目の部位について、ある時刻 t における信頼度を $c_i(t)$ として、Kinect によって取得された時間間隔 $[t_0, t_1]$ 全体の信頼度 \tilde{c}_i は

$$\tilde{c}_i = \int_{t_0}^{t_1} c_i(t) dt \quad (4)$$

によって計算される。

4.3 信頼度を考慮した検索結果

3.3 節のランダムフォレスト法で述べたように、識別結果は多数決による。ここで、信頼度を加味することで、Kinect によって取得したデータの信頼度を考慮する。予測対象をトラバースする際に、各ノードを記述する情報の部位の信頼度を乗算しながら末端ノードまで遷移する(図 6 参照)。全決定木において信頼度の乗算の総和を計算し、得点が最も高いものを識別結果とする。

5. 検索実験と結果

実験は、Intel(R) Core(TM) i7-2640M 2.80GHz CPU、メモリ 8.00GB を搭載した PC で行った。決定木を作成す

表 2 実験で利用したデータ

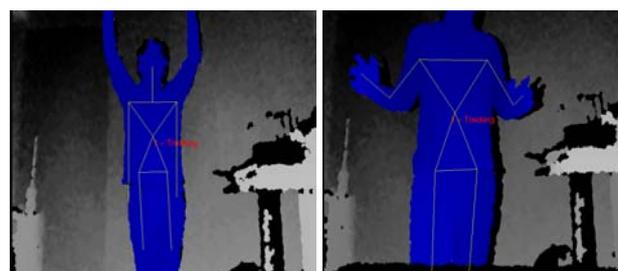
Table 2 A dataset used in the experiments

	ファイル数	総動作時間 [sec]	データ容量 [MB]
データ集合	21	527	11.7

表 3 人体動作データの演技内容

Table 3 Contents of the performance of human motion data

歩く	飛び降りる	ジャンプ	スコップで掘る
走る	倒れる	ドアを開ける	ベッドに入る
書類整理	待ち合わせ	祈りをささげる	坂を下る
立食パーティー	診察する	立って朗読	小銃を構える
しゃがむ	棒にぶら下がる	ハシゴに登る	携帯で電話する



(a) ボーンの検出ミス (両腕がとれていない)
(b) 関節情報の未取得 (頭がとれていない)

図 7 欠陥を含む Kinect による取得データ

Fig. 7 Obtained data by Kinect including defects

るまでに数秒程度、Kinect によるデータの取得、平均姿勢情報の計算、検索の実行はほぼリアルタイムに動作した。実験で利用したデータの数值情報を表 2 に示し、演技内容を表 3 にまとめる。

Kinect で取得したキャプチャデータに欠陥がある場合の例を図 7 に示す。このように明らかにキャプチャによるボーンの検出にミスがある場合にも、各部位の信頼度を導入することによって期待される検索結果を得ることができる。検索に成功する例としては、「歩く」や「走る」のような単純な動作が挙げられる。Kinect を利用して「歩く」動作を提案法による信頼度付きランダムフォレスト法にかけた場合のスコアの順位を表 4 にまとめる。このスコアがそのまま検索順位や類似度を表すわけではないが、データベース内の「歩く」人体動作データをキーとして同じ決定木にかけた場合に、ほぼ同じ順位が得られることから、提案法の結果が意図通りであることがわかる。このように Kinect で取得したデータを入力として、意図通りの検索結果を得ることができることを確認できた。

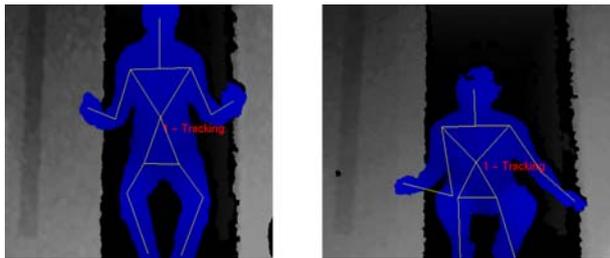
しかしながら、所望の動作がデータベースに存在しない場合には検索結果は意図しないものが出現する。各部位に対して信頼度を考慮することで、検索に有効に働く場合は、主に Kinect で取得したデータに欠陥がある場合である。

また、時間軸に対する信頼度としてアクロバティックな「飛び降りる」動作を検索することにより実験した。図 8(a) に示すように、着地後のポーズのみ Kinect による取得デー

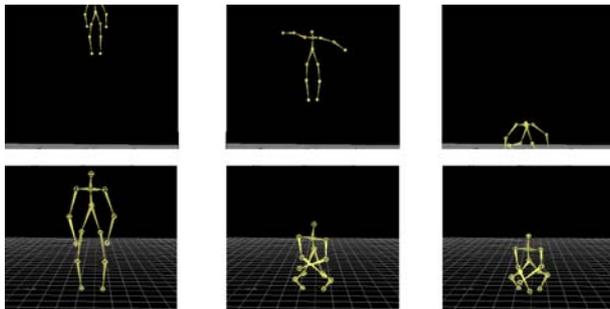
表 4 「歩く」動作の検索結果 (ランダムフォレスト法で重み付き投票の総合スコアの順位)

Table 4 Retrieval results for walk behavior (Order of overall score of weighted voting in a random forest method)

スコア順位	提案法	「歩く」をキー
1位	歩く	歩く
2位	走る	走る
3位	立食パーティー	ドアを開ける
4位	ドアを開ける	ベッドに入る
5位	ベッドに入る	立食パーティー



(a) Kinect による取得データ



(b) 検索結果 (上段)「飛び降りる」, (下段)「しゃがむ」

図 8 時間軸に対する信頼度を利用した検索結果

Fig. 8 Relative result using a confidence in a timeline

タを利用し、その前は信頼度が低い時間帯として設定した。検索結果でスコアが高いものを図 8(b) に示す。「飛び降りる」動作と、「しゃがむ」動作の一部が検出された。この検索結果は、Kinect による取得データによって変わったり、決定木の作成によって変わったりするため、一回で検出することは難しい。

6. まとめと今後の課題

Kinect のように低解像度で欠陥を含む可能性があるモーションキャプチャデータをキーとして、人体動作データのデータベースを検索する手法を提案した。提案法によって、一部分に欠陥を含むような人体動作データからも類似動作データを検索することができる。

今回使用したデータベースには類似した人体動作データが少ないため、容易に識別できたことも考えられる。今後、Web 等でも入手できるデータも検索の対象として、人体動作データの数や種類を増やした実験をすることを予定している。

時間軸に対して信頼度を設定して、アクロバティックな動作を検索する方法は、今後も改良の余地がある。提案法のように、あらかじめ決められた時間軸に合わせて動作することは、検索者に対して余計な負担を強いることになる。Kinect では指の部分まで検出できないため、検索者に簡単なスイッチを握らせて、時間軸に対する信頼度はインタラクティブにオフ・オンする方法が考えられる。しかし、この方法はインタラクティブな検索となるため、決定木を作成するという前計算が必要なランダムフォレスト法は適切でないことが考えられる。ランダムフォレスト法以外の方法で、同程度の検索精度があり、インタラクティブに検索可能な手法を検討する必要がある。

参考文献

- [1] Microsoft Kinect. <http://www.xbox.com/Kinect> (2012年7月1日アクセス)。
- [2] mocapdata.com. <http://mocapdata.com> (2012年7月1日アクセス)。
- [3] CMU Graphics Lab Motion Capture Database. <http://mocap.cs.cmu.edu/> (2012年7月1日アクセス)。
- [4] 川島 基展, 近藤 邦雄, 金子 満. パフォーマンススクラップブックの提案. 情報処理学会研究報告. グラフィクスとCAD 研究会報告, 2008(109):13-18, 2008-10-31.
- [5] Meinard Müller, Tido Röder, Michael Clausen. Efficient content-based retrieval of motion capture data. *ACM Trans. Graph.*, 24(3):677-685, July 2005.
- [6] 浜田 祐一, 栗山 繁, 向井 智彦. ランダムフォレスト法を用いた動作検索. 電子情報通信学会論文誌. D, 情報・システム, 93(11):2516-2524, 2010-11-01.
- [7] Jinxiang Chai, Jessica K. Hodgins. Performance animation from low-dimensional control signals. *ACM Trans. Graph.*, 24(3):686-696, July 2005.
- [8] 竹林 佑介, 西尾 孝治, 小堀 研一. モーションキャプチャデータの類似検索に関する一手法. 映像情報メディア学会誌: 映像情報メディア, 62(9):1420-1426, 2008-09-01.
- [9] Naoki Numaguchi, Atsushi Nakazawa, Takaaki Shiratori, Jessica K. Hodgins. A puppet interface for retrieval of motion capture data. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '11, pages 157-166, New York, NY, USA, 2011. ACM.
- [10] QUMARION. <http://www.clip-studio.com/quma/> (2012年7月1日アクセス)。
- [11] OpenNL. <http://75.98.78.94/> (2012年7月1日アクセス)。
- [12] 浜田 祐一, 栗山 繁, 向井 智彦. 姿勢変位量の特徴抽出に基づく動作データの高速検出. 情報科学技術フォーラム講演論文集, 7(3):303-304, 2008-08-20.
- [13] Tan Pang-Ning, Steinbach Michael, Kumar Vipin. *Introduction to Data Mining*. Addison-Wesley, 2006.