

# 関節の順序関係を利用した 3D ポーズ検索システム

井原真一郎<sup>†1</sup> 金森由博<sup>†1</sup> 佐々木竜生<sup>†2</sup> 徳永隆治<sup>†1</sup>

概要：2Dイラストなどの参考用に特定のポーズをとらせた3D人物モデルが用いられる。しかし3Dツールに不慣れたユーザにとって3Dモデルのポージングは難しい。本研究では、作成済みの3Dポーズデータセットから、簡単な入力によって所望の3Dポーズを検索できるシステムを提案する。様々な視点から見た3Dポーズを検索できるように、視点ごとに主要な関節間の順序関係をビット列で表現し、特徴量とする。ユーザが2D上でポーズの骨格構造を入力すると特徴量が計算され、データセット内の特徴量とビット演算により高速に照合される。1万程度のデータセットに対しリアルタイムに検索結果が得られることを示す。

## A 3D pose retrieval system using joint-order relationships

SHINICHIRO IHARA<sup>†1</sup> YOSHIHIRO KANAMORI<sup>†1</sup> TATSUO SASAKI<sup>†2</sup> RYUJI TOKUNAGA<sup>†1</sup>

3D character models in specific poses are often used as reference when drawing 2D illustrations. However, posing a 3D model is difficult for users unfamiliar with 3D tools. In this study, we propose a system for searching for a pose from a preset database easily by querying a 2D stick figure as input. To support a pose viewed from any direction, our system calculates feature values for various views per 3D pose. Our feature value is a bit sequence that encodes the large-or-small relationship for each pair of major joints' 2D coordinates projected from a certain view. Every time the user specifies a part of a stick figure, the system compares feature values very fast using bit operations. We demonstrate that our system shows the retrieval results in real time for a database with roughly ten thousand feature values.

### 1. はじめに

特定のポーズをとった人物のイラストを何も見ずに描くのは、普段イラストを描き慣れている人であっても容易ではない。これまではイラストを描く際の参照用に、デッサン人形と呼ばれる関節構造を持った人形が用いられてきたが、3DCGによるデッサン人形も用いられている(図1)。

通常、3Dの人物モデルに特定のポーズをとらせるには、3D空間上で関節を動かし、視点を切り替えて結果を確認する、という作業を何度も繰り返す必要がある。こうした作業は普段2Dのイラストしか扱っていないユーザにとって難しい。そのようなユーザでも手軽に3Dのポーズデータを利用可能にするため、本研究ではあらかじめ作成されたポーズデータベースからポーズ検索を行うシステムを提案する。

本研究では、検索対象を特定の視点から見たポーズに限定するのではなく、ユーザが望んだ視点から見たポーズを検索できるようにするため、様々な視点から見たポーズを

基に特徴量を計算する。具体的には、各関節を2D画面に投影した際のx,y座標に関する、他の関節との順序関係を利用する。提案手法では、関節同士の順序関係を単純なビット列として表現するので、データ量が少なく、特徴量同士の比較もビット演算を用いることで非常に高速に行うことができる。これによって1万程度のデータに対して検索を行う場合でも高速に結果を返すことができ、ユーザが入力を微修正する度に検索結果を更新するといった対話的な処理が実現できる。

### 2. 関連研究

3Dのポージングを実現するには、ユーザが自分でポーズを指定していく方法と、すでに用意されたポーズを検索する方法とが考えられる。

ユーザ自身によるポーズ指定については、前述の通り、3D空間での操作に慣れていないユーザには難しいので、それを支援する手法が提案されている。ユーザが2Dの棒人間を入力することで、システムがそれらしい3Dポーズを生成する手法がいくつか提案されている[1,2,3,4]。これらの手法では、希望する3Dポーズの細部までユーザが指定しなければならない。一方、2Dイラスト用のポーズ集が存在

<sup>†1</sup> 筑波大学  
University of Tsukuba.

<sup>†2</sup> 株式会社  
CELSYS, Inc.

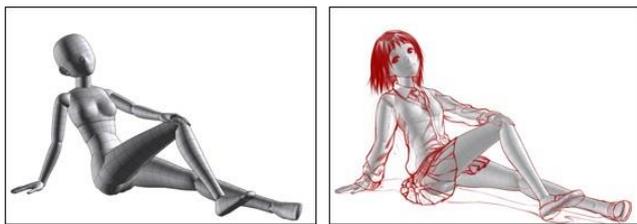


図 1 3D モデルにポーズを適用し、イラスト制作に利用するソフトウェアの例 (POSE STUDIO ©セルシス).

するように、イラストで使われる典型的なポーズはある程度決まっている。ユーザが望むポーズが予め用意されたポーズとは異なる場合でも、基本となるポーズから編集する方が手間は少なく済む。よって、少ない操作で希望するものに近いポーズを取得できれば有益であると考えられる。提案手法では、ユーザが棒人間の主要な関節をひとつ入力する度に、その場で検索結果が得られる。ポーズによっては、そのポーズが2Dでどのように見えるかわかりづらく、指定しづらいものもあるが、提案手法では関節すべてを入力しなくても所望のポーズが得られる場合がある。

その他のポーズ指定の方法として、3D ポーズを実世界の人型デバイスで指定する研究もある[5,6]。また、そのような製品も存在する[7]。しかしこれらはポージングにデバイスが必要という点が問題となる。

3D のポーズを検索するために実際に多く用いられている方法は、ポーズデータに付加されたタグを用いたテキスト検索である[8]。しかし、タグは手動で付加せねばならず、しかも適切なタグを設定するのは容易ではない。例えば座っているポーズがあったときに、単に「座る」というタグでは十分でなく、姿勢の違いで様々な座り方が考えられる。また、実際にイラストを描くことを考えると、そのポーズをどの方向から見るかも重要となる。

テキストを用いない検索に関しては、2D 上で棒人間を入力し、モーションキャプチャデータの検索と可視化を行う、というシステムが提案されている[9]。このシステムでは、ユーザによる棒人間のスケッチ入力を基に、データベース内のモーションキャプチャデータの姿勢や動きを考慮して検索を行うとともにリアルタイムで結果の可視化も行うことでモーションの検索を対話的に行うことができる。しかしこの手法では、入力は真横から見たポーズを入力しなくてはならない、という制約がある。これに対し、本研究では、様々な視点から見たポーズを検索できる。

### 3. 提案手法

特定の視点から見たポーズを表現するための特徴量として、本研究では次の2つを要件として考えた。

- (1) 大量のポーズデータから高速にポーズを検索するため、軽量かつ高速に比較可能であること

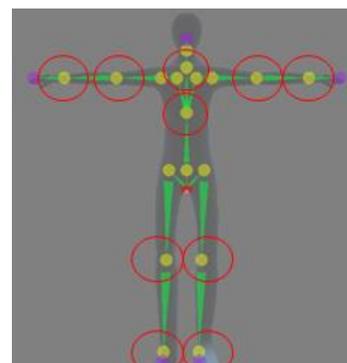


図 3 特徴量計算に用いる主要な 10 個の関節.

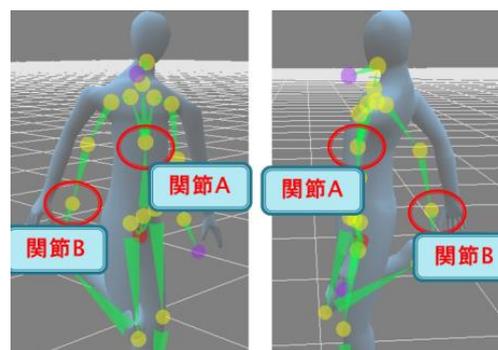


図 4 同一ポーズを異なる視点から見た図。  
左右の関節 A, B の順序関係が異なっている。

(2) ユーザ入力が 2D なので、2D で計算可能であること  
そこで本研究では、これらの条件を満たす特徴量として、関節同士の座標の順序関係に注目した。ここでいう順序関係とはすなわち、ある視点から見た 2 つの関節の内、片方がもう一方より右に位置する、といったような単純な大小関係のことである。順序関係は 3D ポーズを眺める視点を決め、関節位置を 2D 上に投影すれば一意に求めることができる。また、ひとつの関節の組み合わせに対する大小関係は 1 ビットあればよい。さらに、人物の手足の長さや体型が異なっていた場合や、ユーザが入力するポーズの手足の長さがデータベース上のモデルの手足の長さとなっていた場合でも、大小関係さえ一致していれば検索可能という利点がある。

本研究で提案するシステムの概観を図 2 に示す。大きく分けてデータベース作成と検索処理の 2 つに分類することができる。まず大量のポーズデータから検索のための特徴量を計算し、データベースを作成する。この時にひとつの 3D ポーズデータに対して複数の視点を設定し、そこから見た各関節の位置を基に特徴量を計算する。ユーザが検索を行う際には 2D による人型ポーズを入力し、この時に入力された人型の各関節の情報を基にして、データベース作成のときと同様に特徴量計算を行う。入力から計算された特徴量とデータベース内の特徴量を比較し、似ていると判断されたデータをユーザに提示する。

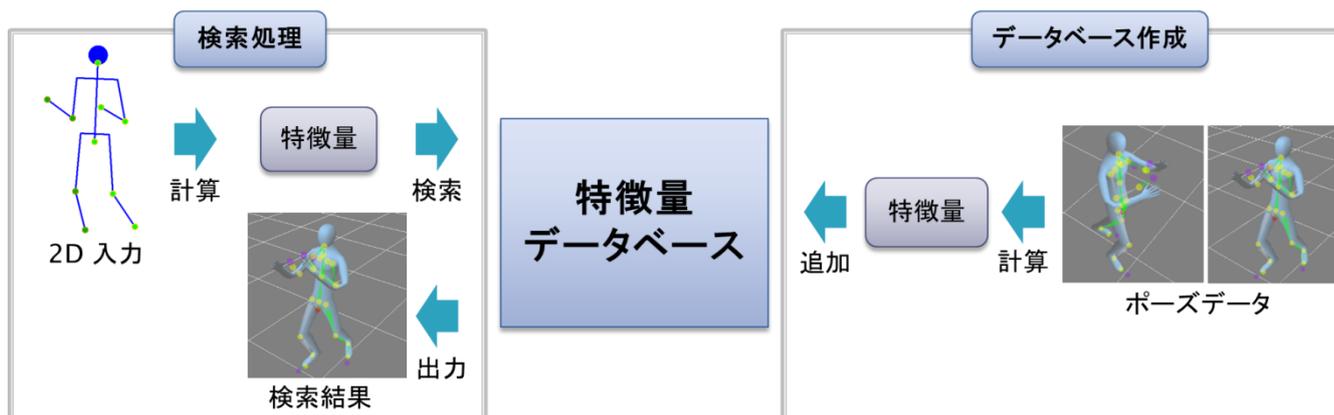


図 2 システムの流れ図.

## 4. データベース作成

### 4.1 特徴量計算

関節の順序関係を用いた特徴量の具体的な計算手順を説明する。まず、ある視点から見たポーズの各関節を 2D 平面上に投影する。次に、投影した関節の内 2 つの関節  $i, j$  の  $x$  座標を比較する。関節  $i$  の  $x$  座標の方が大きければ 1、小さければ 0 という 1 ビット分の情報を割り当てる。  $y$  座標に関しても同様に大小を比較し 1 ビットの情報を割り当て、これらの処理を全ての主要な関節に対して行うことで関節同士の順序関係を表すビット列が得られる。このビット列をこの視点から見た特徴量とする。本研究における主要な関節とは、腰、首、左右の肘、手首、膝、足首の合計 10 個の関節である(図 3)。一般に 3D のポーズデータにはこれら 10 個の関節以外にも肩や腰の付け根などの関節が含まれている。しかし、例えば肩の関節は人間がどんなポーズをとっていても背骨から大きく離れることはない。従ってこれらの関節全てを考慮した場合、結果に対して望ましくない偏りを生じると考え、本研究においてこれらの関節は除外している。10 個の主要な関節の組み合わせ( ${}_{10}C_2 = 45$ 通り)に対して  $x, y$  それぞれの順序関係に 1 ビット割り当てるため、特徴量のデータサイズは合計 90 ビットとなる。

### 4.2 多視点対応

関節の順序関係はポーズを見る視点によって変化する。図 4 は同一のポーズを異なる視点から眺めた画像であるが、左の画像では関節 A が関節 B より右側に位置しているのに対し、右の画像では関節 A の方が左に位置しており順序関係が逆転している。この問題に対応するため、ポーズの周囲に複数の視点を設定し、それぞれの視点から見た関節の順序関係に対して別個の特徴量を計算する。結果としてひとつのポーズが複数の特徴量を持つことになる。本研究では水平、俯瞰、あおりの 3 つの角度それぞれに対し 45 度刻みで周囲に 8 つの視点を設定、合計 24 視点としている。

これにより、ユーザが入力したポーズが 24 視点の内いずれかの視点に近いものであれば対応可能になる。人物のイラスト制作においてはポーズを眺める視点はある程度定番が決まっており、本研究ではポーズ集[10,11]などを参考に視点位置を設定したが、より細かく視点を配置することも可能である。

## 5. 検索処理

### 5.1 ユーザ入力

ユーザがポーズを検索する際には、2D 上で主要な 10 関節を順番にひとつひとつ入力させる(図 5)。このときの関節の入力順は固定となっているが、関節位置はドラッグ操作によって後からでも編集可能になっている。3D ポーズデータと同様に、入力された関節位置から順序関係を調べ、特徴量の計算を行う。

### 5.2 特徴量の比較

特徴量は順序関係を合計で 90 ビット分のビット列として表現しており、2 つの特徴量同士の類似度を求めるため単純に 2 つの特徴量のビットが一致している数を調べている。ビットが一致している数を調べるという操作は高速に行うことが可能なため、現在は全ての特徴量データに対して線形探索を行っている。

## 6. 実験結果

検索システムを C++、OpenGL を用いて実装した。このシステムを Intel Core i5 2.5GHz、メモリ 8GB を搭載したノート PC にて実験を行った。また、ポーズのデータベース作成に際してはカーネギーメロン大学が提供しているモーションキャプチャデータベースを利用した。モーションキャ

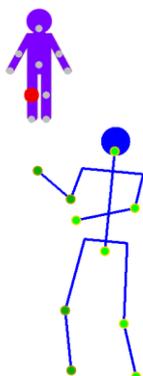


図 5 ユーザ入力. 左上の人型モデルで、次に入力すべき関節(赤丸)が表示され、ユーザは指示に従い順番に関節(緑丸)を入力していく。

ャプチャデータの持つ各フレーム情報を 6 フレームごとに取り出し、これを単一のポーズと見なし特徴量計算を行いデータベースに追加した。モーションキャプチャデータは類似したポーズを多数含んでいる傾向が強いと考えられるため、特徴量をデータベースに追加する前にその時点でのデータベース内で検索を行い、最大スコアが閾値よりも低かった場合のみ追加するようにした。今回の実験では閾値を 80 とし、結果として合計で 13,827 の特徴量データを得ることができた。

図 6 に検索結果の例を示す。検索結果の中に入力したポーズと比較的近いポーズが含まれているのがわかる。また、スコアに極端な偏りがないかを確認するため、検索ごとに全ての特徴量のスコア分布を調べたものを図 7 に示す。検索結果の上位はおおよそ 80 程度のスコアを持っていたが、スコア分布を見ると 80 前後のスコアを持つデータはごくわずかであることがわかり、データベース中の特徴量の中で順序関係がよく類似したものが検索結果として現れていることがわかる。また、分布を見ると多くのデータがスコア 40 から 60 程度になっている。これは  $x$  座標に対する順序関係が異なっても、 $y$  座標に関する順序関係は直立したポーズであればそこまで大きくは異ならないためだと考えられる。

1 回当たりの検索時間はおおよそ 20 から 50 ミリ秒程度であった。このためユーザが関節の入力・変更を行う度に検索処理を更新してもユーザを待たせることなく検索結果を表示させることが可能であった。

## 7. まとめと課題

本研究では 3D ツールに不慣れなユーザであっても手軽に 3D ポーズデータを利用可能にするためのポーズ検索システムを提案した。検索のための特徴量として 2D に投影した主要な関節同士の順序関係をビット列として表現することで軽量かつ高速な検索処理実現することができた。ま

た、複数の異なる視点から見たポーズの特徴量を計算することにより、ユーザがポーズを入力する際も様々な視点から見たポーズ入力が可能になった。

しかし本手法では関節同士の順序関係に 1 ビット分の情報量しか割り当てないため、2 つの関節の位置座標が誤差のような違いしかない場合も片方の関節の座標が大きい・小さいと判断されてしまう。そのためユーザが直立したポーズを入力した際でも、必ず左右の足の  $y$  座標のどちらかが大きいと判断される。ユーザが、片方の足がもう一方の足より上にある、という意図を持って入力したわけではなくても生成される特徴量には 1 ビット分の情報として格納されてしまう。このことはユーザが意図しない検索結果を生む可能性がある。また、ほんの些細な入力の違いが異なる検索結果を生じるというのは検索システムとしては適切ではない。そのためこうしたわずかな座標のずれを考慮する仕組みが必要になる。例えば、大きい・小さい、という 1 ビットの情報だけではなく、「ほとんど同じ」という第三の情報を持たせることが考えられる。特徴量には合計で 2 ビット分の情報量を割り当て、特徴量比較の際にはこの情報が一致した時のみスコアを +1 する。または関節位置の大小比較ではなく距離を測り、それを離散化した量の特徴量とするという方法も考えられる。ただし情報量が大きく複雑になるとデータサイズや速度が犠牲になると思われるので、現在の実装のように単純な線形探索ではリアルタイムに検索結果を返すことができない可能性もあり、その場合は別途、高速化の仕組みを導入する必要があると考えられる。

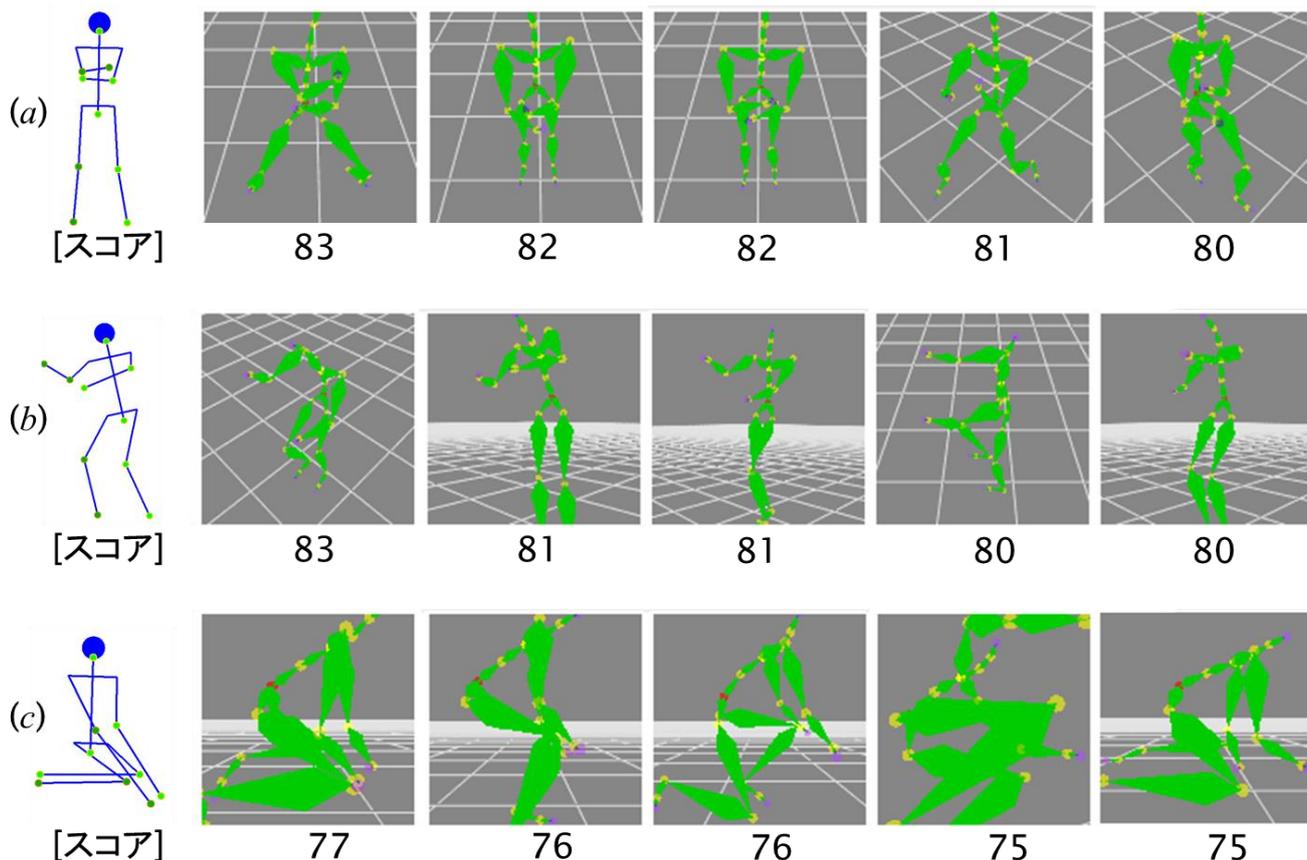


図 6 検索結果例. 一番左の列がユーザ入力であり, その他の列は検索結果. 検索結果の下の数字はスコアを表す.

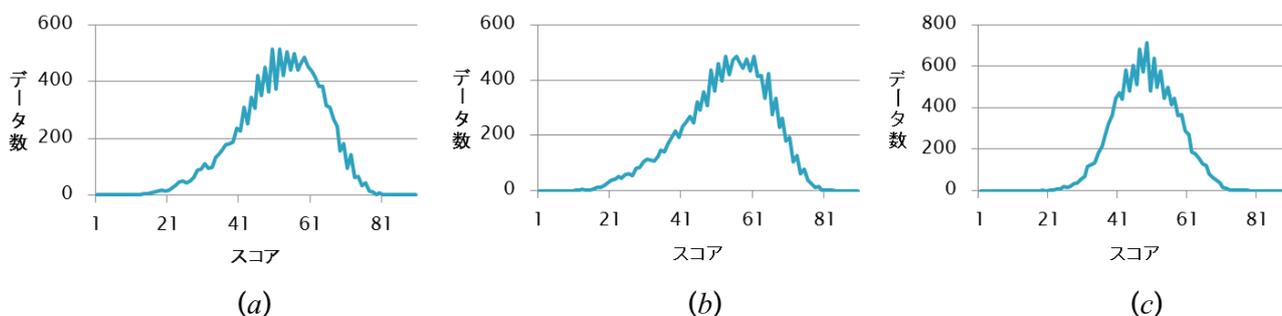


図 7 図 6 (a), (b), (c) の検索結果例に対するスコア分布.

## 参考文献

- 1) DAVIS J., AGRAWALA M., CHUANG E., POPOVIC Z., SALESIN D.: A sketching interface for articulated figure animation. In Proceedings of the ACM SIGGRAPH/Eurographics symposium on Computer animation (2003), Eurographics Association, pp. 320–328.
- 2) MAO C., QIN S. F., WRIGHT D. K.: A sketch-based gesture interface for rough 3D stick figure animation. Perspective (2008).
- 3) LIN J., IGARASHI T., MITANI J., SAUL G.: A sketching interface for sitting-pose design. EUROGRAPHICS Symposium on Sketch-Based Interfaces and Modeling (2010).
- 4) WEI X., CHAI J.: Intuitive interactive human-character posing with millions of example poses. IEEE Comput. Graph. Appl. 31, 4 (2011), 78–88.
- 5) CHIENG FENG T., GUNAWARDANE P., DAVIS J., JIANG B.:

- Motion capture data retrieval using an artist's doll. In Pattern Recognition, 2008. ICPR 2008. 19th International Conference (2008), pp. 1–4.
- 6) NUMAGUCHI N., NAKAZAWA A., SHIRATORI T., HODGINS J. K.: A puppet interface for retrieval of motion capture data. In The proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (2011).
- 7) 株式会社セルシス: QUMARION. <http://www.clip-studio.com/quma/>.
- 8) CMU motion capture database. <http://mocap.cs.cmu.edu/>.
- 9) Myung Geol Choi, Kyung-young Yang, Takeo Igarashi, Jun Mitani, Jehee Lee, Retrieval and Visualization of Human Motion Data via Stick Figures, Computer Graphics Forum (Proc. of Pacific Graphics 2012), Volume 21, Number 7, pp. 2057-2065.
- 10) POSEMANIACS. <http://www.posemaniacs.com/blog/>.
- 11) デッサン人形ポーズ集. <http://posesyu.web.fc2.com/>.