Vol.2011-CG-145 No.1
Vol.2011-CVIM-179 No.1
2011/11/17

センサ内蔵人形を用いたモーションキャプチャデータの検索手法

沼口直紀[†1]　中澤篤志[†1,†2]
白鳥貴亮[†3]　ジェシカ ホギンス[†4]

モーションキャプチャは現在の映像やゲームの制作現場において必須の技術である．特に最近では安価なモーションキャプチャシステム等の開発もあり，モーションキャプチャデータ（動作データ）の一般ユーザーへの利用も広がっている．このような背景から，大規模なモーションキャプチャデータベース（動作データベース）がインターネット上に公開され，ダウンロード・利用できるようになっている．本研究では，大規模な動作データベースからユーザーの所望のデータを検索するために，センサを内蔵した人形インタフェースを用いる手法を提案する．このセンサは 17 自由度を持ち，人と同様の形状・自由度であるため，キーワード等の方法に比べ直感的に検索できる．我々は人形動作と動作データの比較手法として，部分空間法を用いた異常データ検出の考え方に基づき，人形動作と動作データを互いの部分空間に投影することで比較を行う手法を開発した．実験ではユーザーテストを行ってインタフェースの操作性や検索精度および検索時間を評価し，提案手法の有効性を確認した．

A Instrumented Puppet Interface for Retrieval of Motion Capture Data

Naoki Numaguchi ,[†1] Atsushi Nakazawa ,[†1,†2]
Takaaki Shiratori [†3] and Jessica K. Hodgins[†4]

Intuitive and efficient retrieval of motion capture data is essential for effective use of motion capture databases. In this paper, we describe a system that allows the user to retrieve a particular sequence by performing an approximation of the motion with an instrumented puppet. This interface is intuitive because both adults and children have experience playacting with puppets and toys to express particular behaviors or to tell stories with style and emotion. The puppet has 17 degrees of freedom and can therefore represent a variety of motions. We develop a novel similarity metric between puppet and human motion by computing the reconstruction errors of the puppet motion in the latent space of the human motion and those of the human motion in the latent space of the puppet motion. This metric works even for relatively large databases. We conducted a user study of the system and subjects could find the desired motion with reasonable accuracy from a database consisting of everyday, exercise, and acrobatic behaviors.

1. はじめに

映像コンテンツ制作においてモーションキャプチャシステムやモーションキャプチャデータ（以下「動作データ」）は，プロフェッショナル用途にとどまらず一般ユーザーへも利用が拡大している．このような背景から，大規模な動作データベースがインターネット上[1)] で公開され，一般向けアニメーションツール[2)3)] 等で利用できるようになっている．

大規模動作データベースから所望の動作データを検索するには，キーワードを用いる方法や，動作のサムネイル動画等から検索する方法が採用されているが，キーワードでは表現しきれない動作や，大規模なデータベースに適用することは難しい．これは，プロフェッショナルの現場でも同様で，動作データの再利用は殆どされていないのが実情である．一方で，動作データの取得やデータの後処理は大きな手間のかかるものなので，動作データを蓄積してデータベースを構築し，検索し再利用することは非常にニーズの高い技術である．

本論文では，センサの内蔵された人形インタフェースを用いて，大規模な動作データベースから動作を検出するシステムを提案する．人形インタフェースは人間と同様の形状および自由度を持つため，ユーザは所望の動作と同じ動きを人形で表現するだけで，動作データの検索を行う事ができる．このセンサは，手足の関節に 14 個の回転角センサ（ポテンショメータ）と，体の三自由度の方向を検出する方向センサ（IMU: Interia Measurement Unit）からなり，計 17 自由度の情報を取得可能である．ユーザーが操作した人形動作は，これらのセンサにより読み取られ，人間の関節スケルトンに投影（リターゲット）され，動作データベース内のデータと比較される．比較アルゴリズムの結果，類似値の高い動作が検索結果としてユーザーに提示される．

人形は誰しも慣れ親しんだデバイスであり，大人から子供までユーザーを選ばず，操作を

†1 大阪大学サイバーメディアセンター, Cybermedia Center, Osaka University
†2 科学技術振興機構さきがけ, PRESTO, JST
†3 ディズニーリサーチ・ピッツバーグ, Disney Research, Pittsburgh
†4 カーネギーメロン大学, Carnegie Mellon University

© 2011 Information Processing Society of Japan

Vol.2011-CG-145 No.1
Vol.2011-CVIM-179 No.1
2011/11/17

直感的に行えるという点が本システムの最も大きな利点である．また，キーワードで表現しにくい動作や動作の感情，人間では再現しづらいアクロバティックな動作も容易に再現することができる．これは Microsoft Kinect[4] 等，人の動作をそのまま取得できるセンサに対して有利な点である．

人形インタフェースの問題点として，ユーザーが四肢を同時に操作することが難しく，複雑な姿勢を再現しづらい点が挙げられる．これに対して，我々が設計したインタフェースでは，関節のトルクを調節することにより，ユーザーが手を離しても姿勢を固定させることが可能である．また，ユーザーは希望する動作を完全に再現する必要はなく，似た動作を大まかに再現できれば，後述する固有空間法による動作マッチングで類似動作を検索する事ができる．

動作データと人形の操作は「似ているようで異なる」点がいくつか見られるため，両者のデータを直接比較しただけでは適切な検索を行う事は難しい．また，人形の操作はユーザーに応じて大きく異なるため問題は更に複雑である．これに対し我々は，固有空間を用いた異常検出の考え方に基づき，人形動作と動作データを相互に互いの固有空間に対してマッピングし，比較を行う双対部分空間法 (Dual subspace projection method (DSPM)) を提案する．この方法は，ユーザー間の操作方法のゆれに対してロバストであり高速に演算可能である．

以下 2 章では関連研究を述べ，3 章では開発した人形インタフェースのハードウェアについて述べる．4 章では動作データの人形操作データの比較法，5 章で実験結果を述べる．

2. 関 連 研 究

検索インタフェース 従来より，マウスやペンタブレット等で，キャラクタの動きやポーズを描く事で動作を検索・生成する研究が行われている[5),6)]．また，動作データの生成手法として広く用いられている Motion Graph アルゴリズムのバリエーションとしても，様々な動作指定インタフェースが導入されている[7)–9)]．Chai と Hodgins は，ユーザー自身の動きを用いて検索を行う手法を開発している．彼らは，通常のモーションキャプチャよりも少ないマーカーセットで，データベースから所望の動作が検索可能であることを示した[10)]．Ren らの研究では，複数視点から得られた人体のシルエットを用いて動作データを検索する手法を提案している[11)]．Slyper らの研究では，服に複数の加速度センサを取り付け検索を行うシステムを提案している[12)]．ただしこれらの研究で用いられたデータベースサイズは限られたものであり，大規模なデータベースに対する有効性は確認されていない．また人体そのものを用いる検索手法はユーザーへの負担が大きく，アクロバティックな動き等容易に再現できない動作などへの適用が難しい．

人形を用いたインタフェースの提案もいくつかなされている．Johnson らの研究では，センサを内蔵した鳥状のインタフェースを開発し仮想キャラクタとのインタラクションに用いている[13)]．Esposito らは Monkey と呼ばれるセンサ内蔵人形を提案し，CG 制作におけるキャラクタのポーズ指定に用いている[14)]．Yoshizaki らの研究では，小型の二足歩行ロボットを用いてキャラクタのポーズを指定するシステムを提案している[15)]．Feng らは，人形状のインタフェースを用いてデータベースからの動作データ検索を行うシステムを提案している．このインタフェースはセンサは内蔵しておらず，各関節につけられたカラーマーカを複数台のカメラで認識し，その三次元位置からキャラクタのポーズを復元して検索に用いている[16)]．我々の提案手法はこれらの従来手法に比して，インタフェースを動的に動かすことが可能であり，またその動的な情報を検索に用いるという点で異なっている．また，人形操作の特性に合わせた比較手法を新たに開発することで，大規模なデータベースに関しても有効な検索が行えることを確認している．

人形操作と動作データの比較アルゴリズム 人形操作と動作データの比較は時系列データ同士の比較であり，既存の時系列データのマッチングアルゴリズムを用いることが考えられるが，前述したように人形動作は人間の動作データと一致することはまれなため，それらの差を吸収するアルゴリズムが必要である．動作データの比較に良く用いられているアルゴリズムとして DTW (Dynamic Time Warping) がある[17)]．Muller らは動作データの比較のために，体部位やマーカーの幾何的な相対関係をコード化し比較する手法を提案している[18),19)]．Ishigaki らは，ユーザーの意図を反映させたキャラクタアニメーションを生成するために，主成分分析 (PCA) を用いて，あらかじめ用意された複数の動作データと入力となるユーザーの動作データの比較を行っている[20)]．Deng らは，確率的 PCA を動作データのセグメンテーションと検索のために利用している[21)]．また山本らは相互部分空間法 (Mutual Subspace Method(MSM)) を提案し，時系列顔画像の比較に適用し高い性能を発揮することを示している[22)]．

3. 人形インタフェースのハードウェア

提案する人形インタフェースのハードウェア構成を図 1 に示す．インタフェースのサイズは高さ 33cm，幅 (手を広げた状態)44cm，重量約 400g であり，ユーザーが手に持って操作しやすい大きさに設計されている．人形は 14 関節（肩 3 関節 ×2，肘 1 関節 ×2，腰 2 関節

© 2011 Information Processing Society of Japan

Vol.2011-CG-145 No.1
Vol.2011-CVIM-179 No.1
2011/11/17

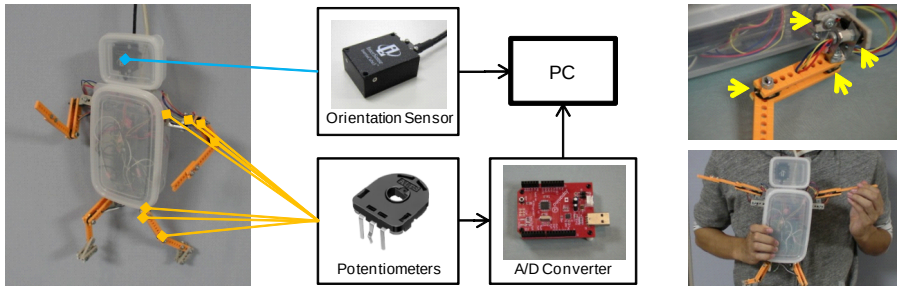


図 1 人形インタフェースのハードウェア．(左) 人形インタフェースの構成．本インタフェースは 17DOF をもち，14 個のポテンショメータと IMU から構成される．ポテンショメータは各関節に取り付けられ，その抵抗値が角度値を示す．ポテンショメータの値は内蔵された Arduino マイクロコントローラで読み取られ PC に 60Hz で送られる．また IMU の値は直接 PC に入力される．(右上) 関節の構成．各矢印はポテンショメータの位置 を示す．(右下) 実際の人形インタフェースの操作．

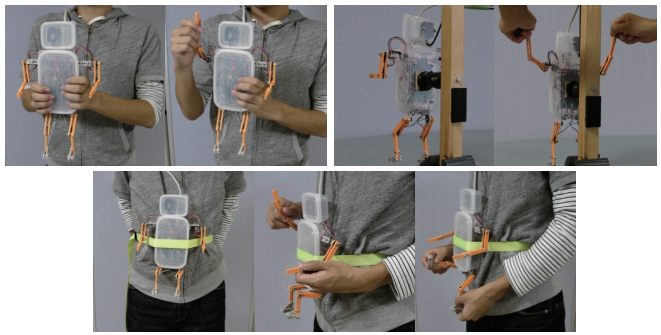


図 2 3 種類の操作方法．(左上) 胴体を片手で保持する方法．(右上) ステージに取り付ける方法．(下) ユーザーの体に取り付ける方法

×2，膝 1 関節 ×2）をもち，各関節には角度センサが取り付けられている．また体全体の姿勢 3 自由度を IMU で取得する．すなわち，計 17 自由度を持つことになる．各センサデータは Arduino マイクロコントローラで読み取られ，USB シリアルインタフェースで PC に 60Hz で送られる．腕部分の構造を図 1（右上）に示す．図中矢印はポテンショメータの位置を示している．

簡単な予備実験から，ユーザーによる人形インタフェースの操作方法として主に以下の 3 通りが見られた．最も簡易な方法は片手で胴体を持ち操作するスタイル（図 2（左上））だが，片手で体を支えるため人形の四肢を操作するためにもう一方の手しか使えない．図 2（右上）のように人形本体をステージに設置する方法もあり，この場合ユーザの両手により人形の手足を操作できるという利点がある．一方で，人形の姿勢のコントロールが不自由になる．図 2（下）のように，ユーザの体にバンド等で固定する方法を提案したユーザもあり，この場合両手を使うことができ，また人形の姿勢のコントロールもある程度可能である．

4. 固有空間を用いた人形操作と動作データの比較

人形動作と動作データの比較は時系列データ同士の比較となる．アルゴリズムを設計するために，試行実験として幾人かのユーザーに「動作データをビデオで提示し同じ動作を人形で再現させる」実験を行ったところ，人形動作には以下のような特徴があることが分かった．

(1) 人形と動作データの関節角度は大きく異なる．これには以下の 2 種類の理由が考えられる．
  (a) 人と人形の関節の動作範囲が異なるため．
  (b) ユーザーの人形操作は，動作データの関節角度ではなくエンドエフェクタ（手・足先）の軌跡をまねるように操作するため，体の大きさの異なる人と人形の間に関節角度の差が生じる（逆運動学問題）．
(2) 人形は人間の動作と異なり物理法則に左右されない．これは体全体がユーザーに支えられているためである．

これらの点から，関節角度ベースの時系列データの比較アルゴリズムをそのまま適用することは困難であり，人の動作と人形動作の違いを吸収できる比較手法が必要になる．これに対し，我々は以下のアプローチを取ることにした．

(1) 人形と人の物理的な大きさの違いに関しては，人形の動きを人のスケルトンにマッピングし比較に用いる．
(2) 特徴量として関節角度ではなく体中心座標系での関節位置を用いる．
(3) 関節動作範囲の違いに関して，部分空間法のアプローチを拡張した双対部分空間法（Dual Subspace Projection Method（DSPM））アルゴリズムを開発し用いる．

本研究では，主成分分析（Principle Component Analysis（PCA））および GP-LVM（Gaussian Process Latent Variable Models）の 2 種の部分空間表現を用いた．以下ではまず人形データおよび動作データベースに対する前処理について述べ，次に通常の部分空間法について述べる．次に提案する双対部分空間法を用いた検索手法について記述する．

© 2011 Information Processing Society of Japan

Vol.2011-CG-145 No.1
Vol.2011-CVIM-179 No.1
2011/11/17

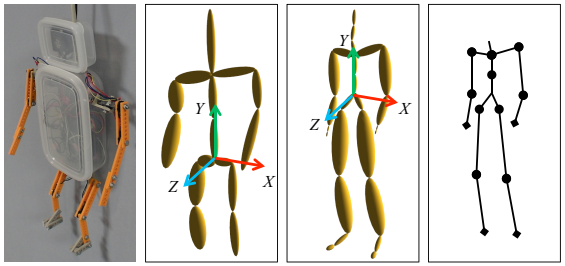


図 3 特徴ベクトル取得に用いる体中心座標系と関節位置．(a) 人形の関節角度および体方向が角度センサにより読み取られ，(b) 人形のスケルトンにマッピングされる．(c) 関節角度を人のスケルトンにマッピングした結果．腰位置に体中心座標系を設定する．(d) 特徴ベクトルに用いられる関節位置．

4.1 データの前処理

4.1.1 人形データの前処理

人（動作データ）の大きさ（リンクの長さ）と人形センサの大きさは異なるため，人形センサからの角度データを人のスケルトンにマッピングし，順運動学計算で得られた関節位置を腰を中心とした座標系（体中心座標系）に投影し，その x, y, z 座標値を特徴ベクトルとする．用いる関節は首，両肩，両肘，両手，両腰，両膝，両足の計 13 点である．体中心座標系の取り方および特徴点を図 3 に示す．これらの値および IMU から得られた体の回転角を合わせ，計 42 次元の特徴ベクトルとする．

4.1.2 動作データベースの前処理

動作データベース内の動作データファイルには，複数の異なる動作が連続して格納されているものがあるため，これらを分割し 1 動作が 1 セグメントに格納されるように分割する．我々は確率的 PCA (Probabilistic PCA) ベースの動作セグメンテーションアルゴリズム[23]を用いた．

4.2 部分空間法による時系列データの比較

部分空間法 (Subspace Method (SM)) は，入力データを比較対象のデータの部分空間に投影・復元した時の再構築誤差（再投影誤差）を類似度とする手法である．$\mathcal{Q}$ を人形インタフェースからの N フレームの特徴ベクトル $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)$ とし，$\mathcal{H}$ を動作データとする．動作データの固有ベクトルを $\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_M$ とすると，$\mathcal{H}$ の M' 次元部分空間は $\mathbf{H} = [\, \mathbf{h}_1\, \mathbf{h}_2\, \cdots \mathbf{h}_{M'}]$ と表せる．入力データ $\mathbf{x}_n$ のこの空間における再構築誤差 Err は，以下で得られる．

$$Err(\mathbf{H}, \mathbf{x}_n) = \left|\left| \mathbf{x}_n - \left(\mathbf{H}\,\mathbf{H}^{\mathsf{T}} \left(\mathbf{x}_n - \bar{\mathbf{x}}^{\mathcal{H}} \right) + \bar{\mathbf{x}}^{\mathcal{H}} \right) \right|\right|, \tag{1}$$

ただし $\bar{\mathbf{x}}^{\mathcal{H}}$ は，動作データ $\mathcal{H}$ の平均である．

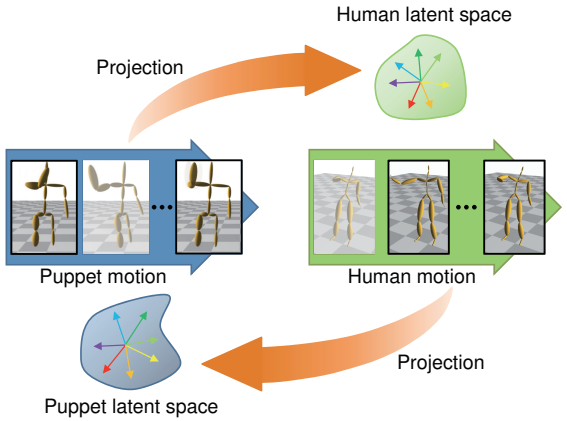


図 4 双対部分空間法 (Dual subspace projection method (DSPM))．入力データ中の 1 フレームの特徴ベクトルは動作データの部分空間に投影され，その再投影誤差が評価される．動作データの特徴ベクトルも同様に入力データの部分空間に投影され，再投影誤差が評価される．全フレーム中の最大の誤差が 2 つのデータの類似度を表す．

4.3 双対部分空間法

部分空間法では，入力データの特徴ベクトルを比較対象となる動作データに投影することで類似度を評価する．この場合，入力データの全フレームが比較対象のデータのサブセットになる場合では類似度は大きくなるという問題点がある．

双対部分空間法 (Dual Subspace Projection Method (DSPM)) ではこの問題を解決するため，順方向の評価に加え逆方向（動作データの特徴ベクトルから入力データの固有空間）への投影も類似度評価に取り入れる（図 4）．これにより，部分空間法で問題となる入力データが比較対象のサブセットになる場合であっても，逆方向の投影の類似度は小さくなるため誤対応を棄却できる．

入力の人形動作を $\mathcal{Q}$，データベースの動作データを $\mathcal{H}$ とし，それぞれの固有空間を $\mathbf{Q}$,

© 2011 Information Processing Society of Japan

Vol.2011-CG-145 No.1
Vol.2011-CVIM-179 No.1
2011/11/17

$\mathbf{H}$ とすると DSPM の評価値 D_{DSPM} は以下で得られる．

$$D_{DSPM}(\mathcal{Q},\mathcal{H}) = \max(Err(\mathbf{H},\mathbf{x}_1),\cdots,Err(\mathbf{H},\mathbf{x}_N), \\ Err(\mathbf{Q},\mathbf{x}_1^{\mathcal{H}}),\cdots,Err(\mathbf{Q},\mathbf{x}_{N^{\mathcal{H}}}^{\mathcal{H}})), \quad (2)$$

ただし $\mathbf{x}^{\mathcal{H}}$ は $\mathcal{H}$ の特徴ベクトルであり，$N^{\mathcal{H}}$ は動作データのフレーム数である．この投影は順方向，逆方向全てのフレームで行うべきだが，計算コストを削減するため，両データからランダムに K フレーム選び投影を行う．最小の D_{DSPM} を示す動作データを，類似度の高いデータとして選択する．

4.4 Gaussian Process Latent Variable Models (GP-LVM)

SM および DSPM は，人体動作の表現によく用いられる GP-LVM[24] と組み合わせることも可能である．GP-LVM は確率的 PCA とカーネル法を組み合わせた非線形マッピング法で，スタイルベースの逆運動学[25] や人体動作からキャラクター動作へのマッピング[26] 等に利用されている．PCA と同様，動作データベースや人形インタフェースから得られた時系列特徴ベクトルは GP-LVM を用いて潜在空間 (Latent space) に投影され，それを逆投影して元の値との違いを計り類似度評価に用いる．

4.5 Joint Velocity Histogram

部分空間法および双対部分空間法は，動作中の姿勢の類似度を評価する．これに加え，関節角の時系列的な類似度を評価するため，関節角の動きヒストグラム（Joint Velocity Histogram (Jvh)) を類似度評価に加える．Jvh は以下で得られる．

$$Jvh = \sum_{n=1}^{N-1} \left(|x_{n+1}^1 - x_n^1|, |x_{n+1}^2 - x_n^2|, ..., |x_{n+1}^M - x_n^M| \right), \quad (3)$$

ただし x_n^k は $\mathbf{x}$ の n フレームにおける k 次元目の要素である．$Jvh_{\mathcal{A}}$ と $Jvh_{\mathcal{B}}$ の類似度は，以下のように Bhattacharyya 係数を用いて得られる[27]．

$$Sim(Jvh_{\mathcal{A}}, Jvh_{\mathcal{B}}) = \sum_{i=1}^{M} \sqrt{\frac{Jvh_{\mathcal{A}}^i}{\sum_{j=1}^{M} Jvh_{\mathcal{A}}^j} \cdot \frac{Jvh_{\mathcal{B}}^i}{\sum_{j=1}^{M} Jvh_{\mathcal{B}}^j}}, \quad (4)$$

5. 実験結果

提案手法を実装しユーザーテストを行った．実験には，設計した人形インタフェースおよび PC(Windows XP, Intel Core2 Duo 2.66GHz, 4GB RAM) を使用した．データの取得及び画像のレンダリングには C++(Microsoft Visual Studio 2008) を使用し，検索手法の実装は Matlab R2008b で行った．実装に対して並列化処理は行っていない．



図 5 実験シーン．

5.1 動作比較アルゴリズムのの評価実験（実験 1）

提案する DSPM 法と他の比較アルゴリズムの比較実験を行った．CMU Motion Capture Database 中の様々な日常動作を示すカテゴリである Subject 14 から，37 個 (計 932 秒) の動作データを選んだ．これらのデータに節の前処理を行った結果 269 個のセグメントが得られた．5 名のアニメーション制作の経験のないユーザーで以下の 2 種類の実験を行った．実際の実験シーンを図 5 に示す．

Protocol 1: 検索対象となる動作をデータベースから選び，そのアニメーションをユーザーに見せる．ユーザーは同じ動作を人形で再現し検索する．

Protocol 2: 検索対象となる動作のキーワードを提示し，ユーザーはその動作を人形で再現して検索を行う．

図 6 に人形操作の一例（「あごに手をやる (“Hand on Chin”)」）を示す．得られた人形操作データを用いて複数の動作比較アルゴリズムによって検索を行い有効性を比較する．用いたアルゴリズムは DSPM(PCA),DSPM(GP-LVM),SM(PCA),SM(GP-LVM),MSM,DTW の 6 種である．前述したように MSM は，顔画像認識等に適用された時系列データ同士の比較アルゴリズムで，2 つのデータセットが張る固有空間の正準角 (canonical angle) を評価値とする手法である．今回の実験では PCA は 5 次元，GP-LVM では 2 次元の固有空間を用いた．また DSPM での動作データからのランダムサンプル数は 50 フレームとした．これらのパラメータを変化させ実験を行ったが，性能に特段の変化は見られなかった．

ユーザーは提示された上位（上位 5, 10）の検索結果に対し，所望の動作が存在するかを評価する．Protocol 1 および 2 の評価結果を表 1 に示す．表中の値は 5 人中何人が，所望の動作が検索結果に含まれていたかを示している．ここから SM(PCA),DSPM(PCA),MSM が DTW,SM(GP-LVM),DSPM(GP-LVM) よりも良い結果を得ていることが分かる．Protocol 1 に関しては，ground-truth となる動作データの順位も評価した．結果を表 2 に示す．ユーザーによる評価と同様，DSPM(PCA) が他の方法に比べて優位であることが確認できた．

© 2011 Information Processing Society of Japan

Vol.2011-CG-145 No.1
Vol.2011-CVIM-179 No.1
2011/11/17

検索に要した時間を表 4 に示す．MSM が最も高速であるが (0.2 秒)，他の PCA ベースの手法も現実的な実行速度であると言える．一方，GP-LVM を用いた非線形手法および DTW は低速であり，インタラクティブなアプリケーションに用いるのは現実的ではない．

5.2 大規模データベースを用いた実験（実験 2）

第 1 の実験から MSM および DSPM(PCA) が良い結果を得たため，より大規模なデータベースを用いた実験を行う．

対象となる動作データは，CMU Motion Capture Database のうち様々なカテゴリの動作が記録された Subject 13, 14, 15, 79, 80, 81, 85, 87, 88, 89, 90, 125, 126, 141, 143 からなる 479 個のファイル（計 2.3 時間）である．ここからセグメンテーションを行い 3065 個の動作セグメントを得た．5 人の新たなユーザーにより Protocol 2 で実験を行った．これは，大規模なデータベースでは類似した動作が多く含まれる可能性があるため，キーワード提示 (Protocol 2) による実験が実利用状況に適しているためである．実験 1 と同様に，ユーザーはアニメーション制作等に対するスキルはない．また検索に使用したパラメータも実験 1 と同様である．

表 3 に実験結果を示す．表中の値は実験 1 と同様，所望する動作が上位 10（上位 5）候補中に存在したユーザー数を表している．ここから，特にアクロバティックな動作やエクササイズ動作（”Russian dance”,”breast stroke”, ”back stroke”, ”fly syle”) 等に対して DSPM(PCA) は MSM よりも良い結果が得られている．また，Joint velocity historgram を用いることで，検索性能が向上していることも確認できる．

表 5 に計算時間を示す．この実験では Matlab R2008b に加え C++(MS-Visual Studio 2008, 並列処理無し) による実装も行い評価した．MSM は DSPM による計算時間を上回っているが，DSPM であっても C++による実装により大規模データベースを対象にした場合でも十分な速度で検索が可能であることを確認した．

6. 結論・考察

本論文では大規模な動作データベースから所望の動作データを検索するため，人形型のインタフェースを用いる手法を提案した．我々は，関節角度と体姿勢が取得可能なセンサ内蔵の人形インタフェースを設計し，人形動作と動作データとの差や個人間の人形操作の違いを吸収可能な検索法を新たに開発し，大規模なデータベースに対しても実用的な検索が可能であることを示した．本システムは人形インタフェースによる直感的な操作によりアニメーション制作等の経験が無いユーザーでも容易に学習・使用が可能であることを確認した．

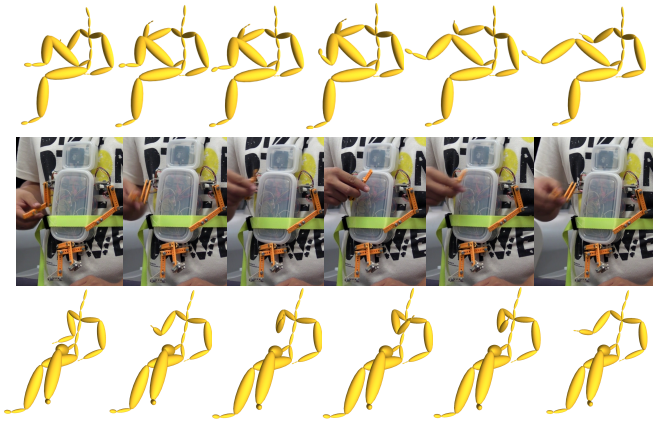


図 6 「あごに手をやる」動作の例.（上）動作データ,（中）人形操作,（下）人形操作を人のスケルトンにマッピングした結果.

表 1 Protocol 1 (P1) および Protocol 2 (P2) の実験結果（実験 1）．各数字は「所望の動作が検索結果候補の中に存在した人数」を表す（最大 5）．括弧内の値は上位 5 位の検索候補に所望の動作が含まれていた人数．括弧外の値は上位 10 位の中に含まれていた人数を表す.

	DTW		**SM (PCA)**		SM (GP-LVM)		**MSM**		**DSPM (PCA)**		DSPM (GP-LVM)	
	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
right punch	5 (4)	4 (4)	5 (5)	4 (3)	5 (5)	5 (5)	4 (3)	3 (3)	5 (3)	1 (1)	3 (3)	3 (2)
squats	0 (0)	0 (0)	4 (3)	3 (2)	0 (0)	0 (0)	2 (2)	4 (2)	4 (4)	3 (1)	0 (0)	2 (0)
wash window 1	4 (3)	3 (3)	5 (3)	5 (5)	2 (0)	3 (2)	5 (5)	5 (5)	5 (5)	5 (5)	3 (2)	3 (3)
wash window 2	3 (1)	4 (4)	5 (5)	5 (5)	4 (1)	5 (1)	5 (5)	5 (5)	5 (5)	5 (5)	5 (4)	3 (3)
side twists	4 (3)	3 (3)	4 (4)	5 (5)	0 (0)	0 (0)	4 (4)	5 (5)	5 (5)	5 (4)	1 (1)	1 (1)
climb 2 steps	3 (3)	3 (2)	4 (3)	5 (3)	0 (0)	0 (0)	5 (4)	5 (5)	5 (5)	4 (3)	4 (3)	1 (1)
point by right hand	3 (1)	2 (0)	5 (4)	5 (3)	2 (0)	3 (2)	5 (5)	5 (5)	5 (5)	5 (5)	4 (2)	3 (1)
sit on high stool	3 (1)	3 (3)	5 (5)	5 (5)	0 (0)	0 (0)	5 (4)	4 (3)	4 (3)	5 (4)	4 (4)	3 (1)
sit and swing legs	5 (5)	5 (5)	4 (4)	5 (5)	0 (0)	0 (0)	5 (5)	5 (5)	5 (5)	5 (5)	2 (2)	3 (2)
hand on chin	5 (2)	2 (2)	3 (3)	2 (2)	0 (0)	0 (0)	1 (1)	3 (1)	5 (4)	5 (4)	4 (4)	4 (4)
drink soda	3 (3)	2 (2)	4 (2)	3 (2)	0 (0)	0 (0)	5 (5)	5 (4)	5 (4)	5 (5)	1 (0)	0 (0)
Average	3.5 (2.4)	2.8 (2.5)	**4.4 (3.7)**	**4.3 (3.6)**	1.2 (0.5)	1.5 (0.9)	**4.2 (3.9)**	**4.5 (3.9)**	**4.8 (4.4)**	**4.4 (3.8)**	3.4 (2.3)	2.4 (1.6)

実験を行った全ユーザーは，人形インタフェース自体の直感的な操作性に関しては好意的に受け止めていた．また，現在のインタフェースの自由度は，様々な動作を再現するのに適切であるという指摘があった．ユーザーは平均 5 分程度でインタフェースに慣れる事ができ，検索を開始することができるようになった．ユーザーから，特にキーワードで表現しに

© 2011 Information Processing Society of Japan

Vol.2011-CG-145 No.1
Vol.2011-CVIM-179 No.1
2011/11/17

表 2 Protocol 1 における提示動作の順位（実験 1）．提示動作が検索結果のうち何位に位置したかを示す．

	DTW	SM (PCA)	SM (GP-LVM)	MSM	**DSPM (PCA)**	DSPM (GP-LVM)
right punch	15.8	8.6	75.8	84.8	5.6	25.2
squats	186.8	8.4	66.8	27	7.6	45.8
wash window 1	37.4	61.2	134.4	17.6	8.8	117.0
wash window 2	43.2	30.6	126.4	50.4	10.2	48.2
side twists	31.2	22.0	83.0	10.0	1.8	62.4
climb 2 steps	47.8	32.6	80.2	9.0	9.0	13.6
point by right hand	16.2	6.6	135.8	2.4	4.8	47.0
sit on high stool	67.4	12.4	143.8	19.0	23.6	70.4
sit and swing legs	1.8	1.2	106.6	1.2	1.0	33.4
hand on chin	5.6	12.8	86.2	33.8	4.2	6.0
drink soda	15.6	12.8	179.4	1.0	5.2	106.4
Average	42.6	19.0	110.8	23.3	**7.4**	52.3

くい動作の検索には有効であるとの指摘があった．一方，ユーザーは同時に 2 つの手足しか動かすことができないため，その範囲で表現できる動作なら良いが，より複雑な動作は再現しにくいという指摘があった．

検索アルゴリズムの評価に関しては，提案した DSPM(PCA) が他の手法より良い成績だった．また，MSM も計算が単純かつ高速に処理でき，検索精度も 2 番目に良かった．GP-LVM による手法は計算コストが多くかかる割に，良い結果を残さなかった．これは，GP-LVM では学習データに対して非常に狭い学習空間を構築するため，今回のマッチングのように人形と人体動作の違いや個人の揺れを吸収できなかったのではないかと推測しているが，さらなる検証が必要である．

本手法の限界として，手の関節がないため grasping のような動作に対しては適用が難しい点が挙げられる．また，インタフェースの平行移動成分を取ることができないため，一部の動作（歩行動作と階段を上る動作等）は区別できないという問題点もある．手の関節や体の移動成分を取るセンサを導入することは可能であるが，ハードウェアや設備の大型化につながってしまうため，簡便さやコストとの兼ね合いとなる．

人形インタフェースのハードウェアは比較的安価であり，ポテンショメータ，Arduino マイクロコントローラを含め 100USD 以下で制作することができる．また今回使用した IMU は比較的高価なものを用いたが，これも安価なユニットに交換することで，センサ全体で 150USD 程度で制作可能である．

謝辞 本研究の遂行に当たり，初期システムの開発及び実験に協力いただいた二宮啓氏（現任天堂株式会社），カーネギーメロン大学の Moshe Mahler 氏，Justin Macey 氏に感謝します．本研究は，JST 戦略的創造研究推進事業さきがけおよび総務省 SCOPE 若手 ICT 研

表 3 大規模データベースに対する検索結果（実験 2）．各数字は「所望の動作が検索結果候補の中に存在した人数」を表す（最大 5）．括弧内の値は上位 5 位の検索候補に所望の動作が含まれていた人数．括弧外の値は上位 10 位の中に含まれていた人数を表す．

	MSM	MSM +Jvh	DSPM (PCA)	DSPM (PCA)+Jvh
right punch	0 (0)	1 (0)	2 (2)	3 (3)
squats	1 (1)	1 (0)	4 (2)	4 (3)
wash window 1	5 (5)	5 (5)	5 (5)	5 (5)
wash window 2	5 (5)	5 (5)	5 (5)	5 (5)
side twists	3 (2)	2 (2)	4 (4)	5 (4)
climb 2 steps	4 (4)	4 (3)	4 (2)	4 (2)
point by right hand	5 (2)	5 (3)	5 (4)	5 (4)
sit on high stool	3 (3)	3 (3)	2 (1)	3 (1)
sit and swing legs	5 (3)	4 (3)	5 (4)	5 (4)
hand on chin	0 (0)	1 (0)	4 (3)	5 (3)
drink soda	4 (1)	4 (1)	3 (3)	4 (3)
hand stand kicks	3 (3)	3 (2)	5 (5)	5 (5)
backflip	3 (2)	4 (1)	4 (4)	4 (4)
cartwheel	3 (2)	3 (1)	4 (4)	5 (4)
spins	0 (0)	0 (0)	0 (0)	1 (0)
handstands	4 (3)	3 (1)	3 (3)	4 (3)
somersault	4 (2)	4 (1)	3 (2)	3 (2)
jump kick	2 (1)	2 (0)	0 (0)	2 (2)
front hand flip	4 (4)	4 (2)	5 (5)	5 (5)
rug pull fall	5 (2)	4 (3)	5 (5)	5 (4)
Russian dance	1 (0)	1 (0)	5 (5)	5 (4)
breast stroke	0 (0)	0 (0)	5 (5)	5 (5)
free style	0 (0)	0 (0)	5 (5)	5 (5)
back stroke	1 (0)	1 (0)	5 (5)	5 (5)
fly style	1 (1)	1 (0)	5 (5)	5 (5)
Average	2.2 (1.4)	2.6 (1.4)	3.9 (3.5)	4.3 (3.6)

表 4 計算時間（実験 1）(sec)．

DTW	SM(PCA)	SM (GP-LVM)	**MSM**	**DSPM (PCA)**	DSPM (GP-LVM)
57.1	8.6	189.7	0.2	2.9	157.7

表 5 計算時間（実験 2）(sec)．

	MSM	MSM + Jvh	DSPM(PCA)	DSPM(PCA) + Jvh
Matlab R2008b	1.7	7.1	37.1	43.0
C++ (Visual Studio 2008)	0.0023	0.0093	0.54	0.76

© 2011 Information Processing Society of Japan

Vol.2011-CG-145 No.1
Vol.2011-CVIM-179 No.1
2011/11/17

究者育成型研究開発（ICT 安心・安全技術）の支援を受けて行われました.

参 考 文 献

1) CMU Graphics Lab Motion Capture Database: http://mocap.cs.cmu.edu.
2) Animate Me, Inc.: Animeeple. http://www.animeeple.com.
3) Smith Micro Software: Poser. http://poser.smithmicro.com.
4) Microsoft Kinect: http://www.xbox.com/Kinect.
5) Thorne, M., Burke, D. and vande Panne, M.: Motion Doodles: An Interface for Sketching Character Animation, *ACM Transactions on Graphics*, Vol.23, No.3, pp. 424–431 (2004).
6) Li, Q. L., Geng, W. D., Yu, T., Shen, X. J., Lau, N. and Yu, G.: MotionMaster: authoring and choreographing Kung-fu Motions by Sketch Drawings, *Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp.233–241 (2006).
7) Kovar, L., Gleicher, M. and Pighin, F.: Motion Graphs, *ACM Transactions on Graphics*, Vol.21, No.3, pp.473–482 (2002).
8) Lee, J., Chai, J., Reitsma, P. S. A., Hodgins, J.K. and Pollard, N.S.: Interactive Control of Avatars Animated with Human Motion Data, *ACM Transactions on Graphics*, Vol.21, No.3, pp.491–500 (2002).
9) Arikan, O. and Forsyth, D. A.: Interactive Motion Generation from Examples, *ACM Transactions on Graphics*, Vol.21, No.3, pp.483–490 (2002).
10) Chai, J. and Hodgins, J.K.: Performance Animation from Low-dimensional Control Signals, *ACM Transactions on Graphics*, Vol.24, No.3, pp.686–696 (2005).
11) Ren, L., Shakhnarovich, G., Hodgins, J. K., Pfister, H. and Viola, P.: Learning Silhouette Features for Control of Human Motion, *ACM Transactions on Graphics*, Vol.24, No.4, pp.1303–1331 (2005).
12) Slyper, R.Y. and Hodgins, J.K.: Action Capture with Accelerometers, *Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (2008).
13) Johnson, M.P., Wilson, A., Blumberg, B., Kline, C. and Bobick, A.: Sympathetic Interfaces: Using a Plush Toy to Direct Synthetic Characters, *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems*, pp.152–158 (1999).
14) Esposito, C., Paley, W.B. and Ong, J.: Of mice and monkeys: a specialized input device for virtual body animation, *Proc. ACM SIGGRAPH Symposium on Interactive 3D Graphics*, pp.109–114 (1995).
15) Yoshizaki, W., Sugiura, Y., Chiou, A.C., Hashimoto, S., Inami, M., Igarashi, T., Akazawa, Y., Kawachi, K., Kagami, S. and Mochimaru, M.: An actuated physical puppet as an input device for controlling a digital manikin, *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems*, pp.637–646 (2011).
16) Feng, T.-C., Gunawardane, P., Davis, J. and Jiang, B.: Motion capture data retrieval using an artist's doll, *Proc. International Conference on Pattern Recognition* (2008).
17) Bruderlin, A. and Williams, L.: Motion Signal Processing, *Proc. ACM SIGGRAPH 95*, pp.97–104 (1995).
18) Müller, M., Röder, T. and Clausen, M.: Efficient Content-based Retrieval of Motion Capture Data, *ACM Transactions on Graphics*, Vol. 24, No. 3, pp. 677–685 (2005).
19) Müller, M. and Röder, T.: Motion Templates for Automatic Classification and Retrieval of Motion Capture Data, *Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp.137–146 (2006).
20) Ishigaki, S., White, T., Zordan, V.B. and Liu, C.K.: Performance-based control interface for character animation, *ACM Transactions Graphics*, Vol.28, No.3, pp. 1–8 (online), DOI:http://doi.acm.org/10.1145/1531326.1531367 (2009).
21) Deng, Z., Gu, Q. and Li, Q.: Perceptually Consistent Example-based Human Motion Retrieval, *Proc. ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pp.191–198 (2009).
22) Yamaguchi, O., Fukui, K. and Maeda, K.: Face Recognition Using Temporal Image Sequence, *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, pp.318–323 (1998).
23) Barbič, J., Safonova, A., Pan, J., Faloutsos, C., Hodgins, J.K. and Pollard, N.S.: Segmenting Motion Capture Data into Distinct Behaviors, *Proc. Graphics Interface*, pp.185–194 (2004).
24) Lawrence, N.D.: Gaussian process latent variable models for visualisation of high dimensional data, *Proc. Neural Information Processing Systems*, p.2004 (2004).
25) Grochow, K., Martin, S.L., Hertzmann, A. and Popović, Z.: Style-Based Inverse Kinematics, *ACM Transactions on Graphics*, Vol.23, No.3, pp.522–531 (2004).
26) Yamane, K., Ariki, Y. and Hodgins, J.: Animating Non-Humanoid Characters with Human Motion Data, *Proc. ACM SIGGRAPH / Eurographics Symposium on Computer Animation* (2010).
27) Kailath, T.: The Divergence and Bhattacharyya Distance Measures in Signal Selection, *IEEE Transactions on Communication Technology*, Vol.COM-15, pp.52–60 (1967).

© 2011 Information Processing Society of Japan