

関節速度の特徴識別による動作データの自動分割

吉 川 裕 規[†] 栗 山 繁[†] 金 子 豊 久[†]

動作データの再利用性を高めるために、基本となる動作単位に分割してデータベースを構築する必要がある。本研究では、動作データの各関節の速度の局所最小点が多くの場合に動作の分割点となることに着目し、ユーザにより指定した関節部位ごとの速度情報を教師信号とするサポートベクターマシンを用いて、動作を自動分割する手法を提案する。計算された分割点の、人間が手で与えたものに対する再現率と適合率を求めて、手法の有効性を評価した。

Automatic Segmentation of Motion Data with Feature Recognition of Joint Velocity

HIROKI YOSHIKAWA,[†] SHIGERU KURIYAMA[†] and TOYOHISA KANEKO[†]

Recycling of motion data requires dividing it to fundamental motion clips for efficiently constructing a database. This paper proposes a method for automatically clipping motion data with a support vector machine. Supervised signals are given by the velocity of each joint whose local minima often correspond to the frame partition of the motion. Usability of our method is evaluated by computing the reproduction and adaptation ratios of the detected frame partitions against manually selected ones.

1. はじめに

近年、CG アニメーションの分野で動作データが広く用いられているが、動作データの制作には多くの時間と費用を要する。そのため、過去に取得した動作データ資源の再利用が盛んに試みられている。中でも、長い動作系列を基本となる動作単位に自動的に分割する方法は、動作データの再利用性を高める上では重要な技術と考えられる。

本研究では、モーションキャプチャ装置により取得した動作データを、各関節の速度の特徴を用いて自動的に人間の動作の基本単位に分割することを試みる。動作データにおいて各関節の速度の局所最小点は、基本動作の分割点となることが多い。ただし、その局所最小点のすべてが分割点となるわけではないので、適応的に選別する必要がある。本手法では、関節速度を特徴ベクトルとするサポートベクターマシン (SVM) を関節部位ごとに構成することにより、分割点を含むフレーム区間を識別して、自動的に動作を基本単位に分割する手法を提案する。

2. 関連研究

動作分割における関連研究として、動的計画法を用いて類似動作を検出する手法¹⁾ や SVM により動作を注釈付けする手法²⁾ がある。これらのパターン認識技術を用いた手法は比較や学習に用いる動作データが必要となる。また、関節速度や短時間スペクトル分析を用いた動作分割³⁾⁴⁾ や舞踊動作プリミティブの検出⁵⁾、および DOF セグメンテーション⁶⁾ 等のパターン認識技術を用いない手法は、人体関節から得られる特徴に閾値を設定することにより動作を分割する。これらの手法においては学習のための動作データは必要ないが、特徴の閾値がすべての動作に対して有効であるとは限らないので、汎用性に欠ける。

以上の考察に基づき、本手法ではパターン認識技術を用いた後、人体関節による速度の局所最小点に着目する分割方法を導入する。

3. 動作分割

動作データは再利用の観点からは基本的な動作毎に保存するのが望ましい。しかしながら、多数の基本動作を含む系列を少数回の計測でまとめて取得し、後処理として基本単位に分割する方が計測作業の観点からは効率的である。一方、動作データの処理においては、

[†] 豊橋技術科学大学 情報工学系
Toyohashi University of Technology

言語処理の分野の形態素解析で用いられるような辞書がないので、基本となる動作単位を自動的に検出するのは困難である。そこで、分割の際に重要となる関節部位を特定することにより、関節速度の局所最小点から分割点を推定する手法を提案する。

3.1 重要関節部位の特定

動作データを基本単位に分割する際に、ユーザーにおけるポリシー、すなわち動作の分割点の決定法は一般に異なる。本手法ではポリシーはユーザーによって与えられ、それに基づいて学習することによりユーザーに適した分割点が検出される。ユーザー自身が判断する動作における停止点と基本動作間の遷移点を分割点の候補とする。なぜなら、動作の停止点は次の行動に移ることが明確に理解でき、基本動作間の遷移点は前後の動作における違いが、意識的に判断できると考えたからである。

動作データにおける分割点周辺の動作の特徴を利用して、ユーザーが見本として与えた各分割点に対して個別の識別器を構成することも考えられるが、大規模な動作データに対しては与える分割点の個数も膨大なものとなるので現実的ではない。また、与えられたすべての分割点を発見する1つの識別器を作成できればよいが、分割点の前後に含まれる動作には様々な特徴のものが考えられるので、分割点の推定が困難となる。したがって、本研究では、人間は動作を分割するとき特定の関節部位に着目しており、各部位で識別される動作分割点の特徴にはある程度の共通点があると仮定し、部位ごとに識別器を構成した。すなわち、分割点周辺の動作は多くの場合、機能的に類似した動作を行うため（例えば、腕なら伸縮や振り等）、部位ごとの動作特徴はばらつきが小さくなり、分割点の識別が容易になると考えられる。

3.2 学習方法

動作データの分割は、未知の動作データに対しても幅広く対処できる必要がある。そこで本手法ではパターン認識技術の1つであるSVM⁷⁾を用いた。SVMは学習パターン（特徴ベクトル）を使って最適な識別境界を決定することにより未知パターンを判定する。識別境界は式(1)の最適化問題を解くことにより決定される。

$$\begin{aligned} \min_{\mathbf{w}, b, \xi} G(\mathbf{w}, \xi) &= \frac{1}{2} \|\mathbf{w}\|^2 + c \sum_{i=1}^n \xi_i \quad (1) \\ \text{subject to} \quad &\forall i, \quad y_i \cdot (\mathbf{w}^t \mathbf{x}_i + b) \geq (1 - \xi_i) \\ &\forall i, \quad \xi_i \geq 0 \end{aligned}$$

ここで、 $\mathbf{x}_i (i = 1, \dots, n)$ は n 個の学習データであり、

$y_i (i = 1, \dots, n)$ は $\mathbf{x}_i$ の属するクラスを表し、1と0の値のみを取る。 $\mathbf{w}^t, b$ は境界識別のパラメータであり、 $\xi_i (i = 1, \dots, n)$ は線形分離不可能な場合に条件を緩める正変数である。SVMは学習パターン中の識別境界候補からマージンがより大きなものを選ぶ（マージン最大化）という自然な考え方に基づいているため、汎化能力の高いパターン認識手法といえる。また、複雑な識別境界を持つ場合でも、カーネルトリックと呼ばれる方法を用いて、非線形の識別関数を構成することができる。本手法ではSVM library⁸⁾を用いて、Radial Basis Function(RBF) カーネルを使用した。ただし、学習誤りを最小化するか、マージンを最大化するかのトレードオフを決定する定数 C は $C = 1000$ と設定した。

ここでの目的は、関節部位ごとに識別器を構成し、各部位における分割点周辺の区間を検出することである。区間を検出する理由は、分割点とその周辺の分割点の候補の特徴が類似しており、SVMでは ξ_i により制約条件を緩めていることから正確に分離することが困難なためである。そこでまず分割点が存在する区間を認識して、その区間の中から別の方法で最適な分割点を検出する。

SVMは、各特徴ベクトルに対して、分割点の候補集合に属する($y_i = 1$)か、属さない($y_i = 0$)かを指定することにより学習する。ただし、ユーザーが指定した分割点から前後0.5秒のフレーム区間に属するすべてのデータを分割点候補に属するものとする。本手法の動作データは腰部の関節を基点とし体躯の前向きを一致させた場合の座標により表される。図1に人体の関節位置イメージを示す。

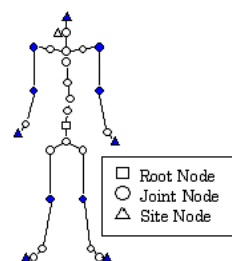


図1 姿勢データの各関節位置

特徴ベクトルは体中心座標系の両手先、両足先、両肩、両肘、両膝、頭の11部分それぞれの位置の変位量（速度）と、そして体の向きを変えるときに発生する度の変位量から成る値を考慮する。さらに現在のフレームから前後1.0秒を0.2秒刻みで同様に特徴抽出

したものを含めることにより、各フレームにおける特徴ベクトルの次元数は計 132 次元となる。図 2 に特徴抽出のイメージを示す。既存の手法²⁾では、クラスタリングされた動作データに対して注釈付けしており、1 秒間にわたる関節位置を動作の特徴としている。各関節の位置では類似した動作しか検出できないが、本手法では速度の特徴を用いることにより意味的に類似した動作（例えば、前に手を伸ばす、横に手を伸ばす等）を柔軟に検出できると考えられる。

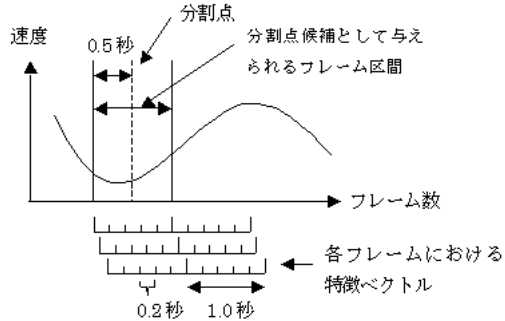


図 2 特徴抽出

学習システムの概要を図 3 に示す。学習用動作データから特徴を検出し、分割点候補の集合で構成されるフレーム区間（以後、分割フレーム区間）を部位ごとに SVM で学習することにより個別の識別器を構成す

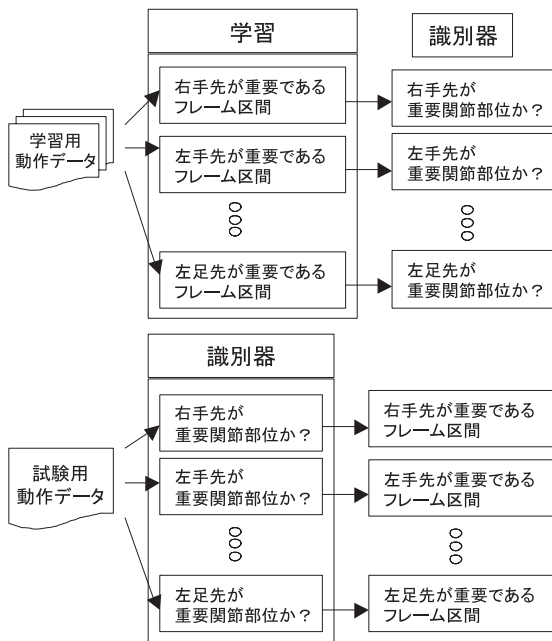


図 3 学習システムの概要

る。作成された個別の識別器を未知の動作データから得られた特徴に用いることにより、部位ごとの分割フレーム区間を特定できる。

3.3 ベル型関数による速度曲線の当てはめ

次に、以上の処理により特定した分割フレーム区間から分割点を決定し、全部位に対する分割点から最適なものを選択する。分割フレーム区間で分割点となり得るのは、その区間の関節速度の局所最小点と考えられる。一方、動作として最適な分割点は、基本動作の開始と終了の部位が異なる（左足で始まり右手で終わる等）場合もあり、部位ごとの分割点を時間軸に沿って並列に考慮する必要がある。この節では部位ごとの分割フレーム区間から求めたすべての分割点の候補からユーザの意図する分割点のみを選択する手法について述べる。

全部位に対する分割点は多数あるが、フレーム数の近い分割点は姿勢が良く似ているので、同じ分割点としてまとめることができる。ゆえにここでは、隣接する分割点の特徴ベクトルのユークリッド距離が閾値以下なら候補から削除する。

一方、一般に最適な分割点は、エンドエフェクタによる到達運動の開始や終了が考えられる。目標に手を伸ばす到達運動の際の手先の軌跡は、始点と終点を結ぶ直線に近い緩やかな曲線で、速度曲線はいわゆるベル型である（図 4）。ゆえに、全ての基本動作はベル型の速度曲線を描くという仮定を設ける。これにより速度の速い局所最小点である分割点の候補を削除することができ、もっともらしい分割点が選別される。

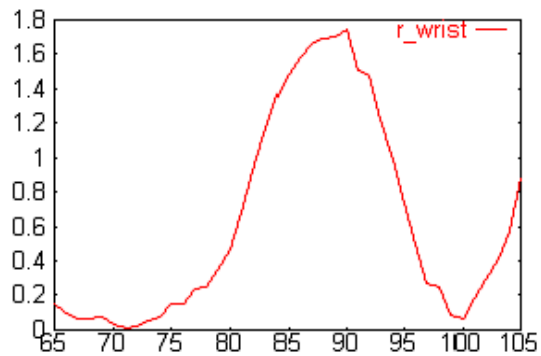


図 4 到達運動における右手先速度

この速度曲線を式 (2) の関数で近似する。

$$f(x) = h \cdot \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) + v_{min} \quad (2)$$

ここで $h = v_{max} - v_{min}$ を表し、 v_{max} は局所最大点、 v_{min} は局所最小点を表す。 μ は基本動作の開始を示

す局所最小点と、次の局所最小点の中間に位置するフレーム番号の値を表す。また、 σ はベル型関数における $f(0)$ の値から

$$\sigma = \sqrt{-\frac{\mu^2}{2\log_e\left(\frac{f(0)}{h}\right)}} \quad (3)$$

で計算される。このベル型関数の形状が速度曲線を最も精度良く近似するように v_{max} と v_{min} を設定する。ただし、ベル型関数の形状が速度曲線と合うように経験的に $f(0) = 0.05$ と設定した。また、各関節の軌道とこの曲線との差異 $diff(i, j)$ は以下の式

$$diff(i, j) = \frac{\sum_{k=i}^j |v(k) - f(k-i)|}{|i-j| \cdot h} \quad (4)$$

で与えられ、この値を最小化する分割点を選択する。ただし、ここで $i, j (i < j)$ は局所最小点のフレーム番号を示しており、 $v(k)$ は k フレームにおける速度を示している。また、基本動作に費やすフレームの数ができるだけ小さくなるようにパラメータを設定する。なぜなら、基本動作の開始と終了が大きく離れていた場合であっても、ベル型曲線との差異が小さくなる可能性があるからである。そこで、指数関数を用いてベル型曲線との差異 $diff(i, j)$ に以下のペナルティ項を付加する。

$$\exp\left(\text{weight} \cdot \left(\frac{|i-j|}{\text{maxFrame}}\right)\right) \quad (5)$$

ただし、 maxFrame は動作データの最大フレーム数を表す。 weight を大きくすることにより、近くの局所最小点が優先され、分割点の数が増える。すなわち、この weight は分割点数の調整に利用でき、基本的な動作に分割する際の粒度に影響する。ベル型の速度曲線を当てはめる処理手順としては、基本動作の開始を示す局所最小点 i と残りすべての局所最小点 j を逐次比較していき、それぞれのベル型曲線との差異 $diff(i, j)$ を求め、最も小さい値となったときの j を基本動作の終了とする。この j を次の基本動作の開始を示す局所最小点 i とし、同様な処理を繰り返す。

基本動作の開始と終了の部位が異なることも考慮すると、ベル型の速度曲線の前半は動作開始の部位の速度曲線と比較し、後半は動作終了の部位の速度曲線と比較する。前半の差異と後半の差異の和を総合的な差異とする。これにより異なる部位の速度を考慮し、基本動作の分割点を選別する。

4. 分割点の自動判別実験

予備実験として3つの動作データを2つは学習用、1つは試験用として実験した。今回我々が用いた動作

データは作業的な動作である。作業動作は多くの場合、決まった動きを繰り返し行うものである。動作データは主に締め付け作業であり、いろいろな向きのボルト締めを行っている。用いた動作データの総フレーム数と時間を表1に示す。動作データの数字(1,2,3)は同じ種類の動作であることを意味し、1は工具なしで大まかな動きであり、2は工具ありで腕による細かい動き、および3は1と2を混ぜたような動きである。アルファベット(a,b,c)は数字で示された種類の動作を同様に複数回計測したものである。ただし、演技者は同一人物である。実験手順は、はじめに3.2節で述べた学習方法で構成された識別器を用いることにより、分割フレーム区間から関節速度の局所最小点を求めて分割点とした検出精度を求め、次に、その結果に対してベル型関数による速度曲線を当てはめ分割点を検出した。最後に学習データを増やし、他の動作データでも実験した。

表1 動作資源

動作データ	総フレーム数	時間(秒)
m1a	2390	80
m1b	1940	65
m2a	3220	107
m2b	2720	91
m3a	4050	135
m3b	1630	54
m3c	1040	35

4.1 認識結果

表1のm1aとm2aを学習した後、m1bを試験した結果を図5に示す。波形として示されているのが関節の速度の波形であり、帯状に網掛けて示されているのがSVMによって認識された分割フレーム区間である。すなわち、その関節部位において分割点を含むと判断された区間である。破線は被験者が正解と判断した分割点のフレーム位置を示す。

次に、検出精度を示す指標として、情報検索技術でよく用いられる再現率R、適合率P、およびF値を用いた。適合率とは、検索結果中の正解の割合(正解率)を表し、再現率とは、検索結果として現れた正解がすべての正解に占める割合(カバー率)を表す。F値は適合率と再現率の双方が大きい値であれば大きい値をとり、 $F = 2PR/(P+R)$ で表され、その値の範囲は0~1である。

本手法の各段階の検出精度を評価した結果を表2に示す。正解分割点数は被験者が主観的に選んだ分割点の数であり、検出分割点数は計算によって自動的に検

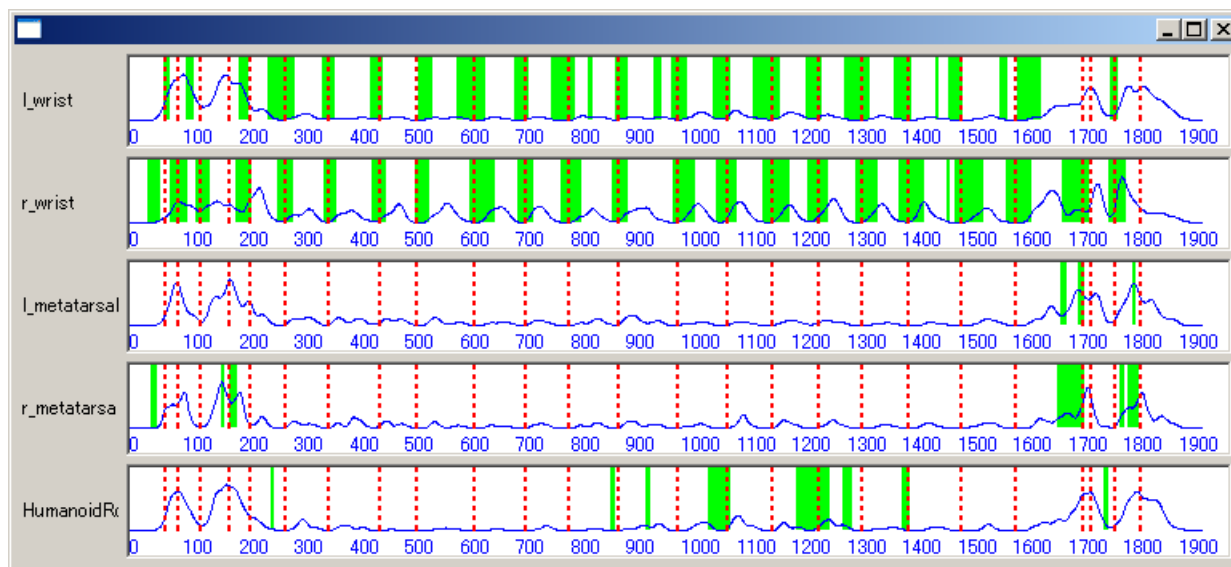


図 5 タイムライン

出した分割点の数である。ただし、人間が選んだ分割点と計算で選ばれた分割点は、双方の時間差が 0.5 秒以内であれば一致していると判定した。表 2 の手法

表 2 検出精度

手法	(ア)	(イ)	(ウ)
総正解分割点数	25	25	25
検出分割点数	64	37	22
検出正解分割点数	25	25	21
再現率 R	1	1	0.84
適合率 P	0.39	0.68	0.95
F 値	0.56	0.81	0.89

(ア) は、分割フレーム区間から関節速度の局所最小点を分割点とした結果であり、手法 (イ) は、手法 (ア) により求めた分割点と隣接する分割点で、その特徴ベクトルのユークリッド距離が一定の閾値 (0.1) 以下のものを削除した結果である。手法 (ア)、手法 (イ) において、再現率はよい精度が得られているが、計算により検出した分割点の数が多いため適合率が悪い。手法 (ウ) は、手法 (イ) により求めた分割点に、ベル型関数による速度曲線を当てはめて、分割点の候補を削除したものである。正解分割点がわずかに削除され再現率が低下したが、適合率が上がり約 0.9 の F 値を得ることができた。この結果より本手法で分割点候補から適当な個数のものを絞り込むことができるとわかる。

次に、ベル型関数の *weight* の値が分割点の検出精度にどのように影響するのか調べた結果を表 3 に示す。

動作データ m1b のすべての局所最小点の数が 485 であり、正解分割点の数は 25 である。*weight* に対してベル型関数による分割点候補の削除後の局所最小点の数と正解分割点の数について調べた。ベル型関数を用いることで正解分割点を保ちつつ分割点の候補を削除することができた。わずかではあるが、正解分割点も削除されており、これはポリシーに反した動作の速い瞬間における分割点だと考えられる。*weight* の値を大きくすると近くの分割点が優先され分割点の数が多くなる効果があるので、すべての正解分割点をカバーすることができるが、局所最小点を多く取り過ぎる結果となる。この結果では、F 値を見ると *weight* は 8 程度が良いとわかる。

表 3 ベル型関数の有効性

weight	8	16	32	64	128
局所最小点の数	34	51	81	113	137
正解分割点の数	22	23	24	24	25
再現率 R	0.88	0.92	0.96	0.96	1
適合率 P	0.74	0.49	0.31	0.22	0.18
F 値	0.80	0.64	0.47	0.36	0.31

同様に残りの動作データに対して処理を行った結果を以下に述べる。図 6 は、動作データ m1a, m2a, m3a を学習した後の検出精度である。ただし、試験データには m1b, m2b, m3b, m3c を用いた。m1b 以外の動作データは腕による細かい動作が含まれているため、検出精度がやや悪くなっている。平均して約 0.7 の F 値が得られた。*weight* を大きくすることにより分割

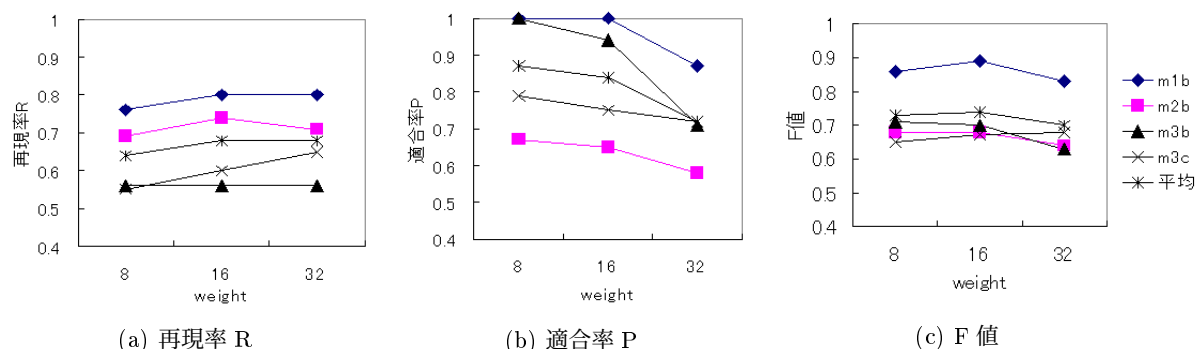


図 6 検出精度 (m1a, m2a, m3a を学習)

点を多く取るため、再現率 R は増加する傾向が見られるが適合率 P は減少する。この場合、F 値を見ると *weight* は 16 程度が良いと考えられる。したがって、*weight* の値は大まかな動作では小さくし、細かな動作では大きくする必要があるものと思われる。

5. ま と め

本研究では SVM を用いて人間の動作を基本的な動作に自動的に分割する手法を提案した。動作における速度の局所最小点が多くの場合に基本動作の分割点になることを考慮し、その周辺の特徴を認識することにより、動作の分割点を求めた。またベル型関数により分割点を選別することにより、基本動作の検出精度を高めた。

動作データを切り出す際には、動作分割のポリシーに個人差があるので、別のユーザが同じ識別器を用いて満足のいく結果が得られるとは限らない。ゆえに本手法のように、ユーザの指定した分割点を用いて学習して適応的な識別器を構成することは重要であると考えられる。また、本手法ではユーザは動作データを分割するときに特定の部位に着目しているという仮定を設けたが、分割点によっては複数の部位に着目することや、全身に着目する場合がある。複数の部位に着目することにより多数の局所最小点が分割点の候補となるが、ベル型関数による速度曲線を当てはめることにより最適な分割点を選別することができる。

実際の動作の分割点は、局所最小点以外の特徴を有することも考えられるので、その特徴を発見する方法を開発する必要がある。また、細かい腕の振りなどを含む動作データでは切り出しの検出精度が低いので、今後はその改善を検討する。

動作データを切り出した後、コード化することにより動作データを容易に検索できる。また、動作の接続性に関する特徴を得ることができれば、再利用性をよ

り向上できると考えられる。一方、本研究では作業的な動作データ⁹⁾を用いたが、一般的な動作データ¹⁰⁾でも実験する必要がある。今後の課題として、基本的な動作のコード化や動作の接続性に関する特徴の検出、および一般的な動作データを用いた実験等が挙げられる。

謝 辞

本研究は、日本学振興会・科学研究費・基盤研究 (C) の援助を受けて実施された。また、その一部は文部科学省 21 世紀 COE プログラム「インテリジェントヒューマンセンシング」と「東海産業技術振興財団」の援助を受けている。

参 考 文 献

- 1) Kovar.L, Gleicher.M, Automated Extraction and Parameterization of Motions in Large Data Sets, SIGGRAPH 2004
- 2) Arikan.O, Forsyth.D, O'Brien.J, Motion Synthesis from Annotations, SIGGRAPH2003, 402-408
- 3) Uehara.K, Shimada.M, Extraction of Primitive Motion and Discovery of Association Rules from Human Motion Data, Progress in Discovery Science 2001, LNAI 2281, pp. 338-348
- 4) 森隆規, 上原邦昭, 動作データからの基本動作の検出と相関ルールに基づく動作認識, 電子情報通信学会, 人工知能と知識処理研究会, 2002
- 5) 中澤, 中岡, 池内, モーションキャプチャデータからの舞踊動作プリミティブの抽出, 第 19 回日本ロボット学会学術講演会予稿集, 2001
- 6) Ashraf.G, Wong.K, Semantic Representation and Correspondence for State-Based Motion Transition, IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL.9, NO.4, 2003
- 7) 津田 宏治, 電子情報通信学会誌 Vol.83 No.6 pp.460-466 2000.6
- 8) <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- 9) <http://tomolow.vcl.ics.tut.ac.jp/>
- 10) <http://mocap.cs.cmu.edu/>