統計的相関に基づく動作パターンの リアルタイム教師なし分節化と 原始シンボルの自律的獲得

高 野 渉* 中 村 仁 彦*

Realtime Unsupervised Selftuning Segmentation of Behavioral Motion Patterns Based on Probabilistic Correlation and Its Application to Automatic Acquisition of Proto-Symbols

Wataru Takano* and Yoshihiko Nakamura*

Mimesis is the hypothesis that human intelligence originated in the interactive communication of motion recognition and generation through imitation. This is attractive for artificial intelligence. We have developed a mimesis system using Hidden Markov Models (HMMs) and their parameter sets are defined as the proto symbols. In conventional systems, designers have to segment a motion pattern in sequential motion data to embed the motion pattern in an HMM. However it is necessary to have the ability of motion pattern segmentation in order to autonomously learn and develop through imitation. In this paper, we propose a motion segmentation method that consists of three phases. In the first phase short sequences of motions are encoded. In the second phase the correlation matrix of the encoded sequences are computed. In the third phase motion patterns are segmented based on error between the encoded sequences observed and predicted from the correlation matrix. Moreover we show that it is possible to acquire the proto symbols by providing the mimesis system with the segmented motion patterns.

Key Words: Behavior Segmentation, Competitive Learning, Correlation Learning, Hidden Markov Model

1. は じ め に

ロボットの知能研究において、動作を模倣して新しい行動を習得する見まね学習モデルが提案されている [1]~[4]。それらの1つとして、身体運動時系列データをモデル化し、そのモデルを用いた行動の認識、行動の生成を行う双方向の情報処理の枠組みであるミメシスモデルが提案されている [5]~[10]。このモデルは、認知心理学において提唱されているミメシス理論 [11] および神経科学におけるミラーニューロンの発見 [12] [13] などの近年の脳科学のパラダイムの発展にヒントを得て、身体運動情報の記号化とそれを操作して環境にはたらきかけることを機能としてもつ人工知能を構築しようとする双方向情報処理モデルである。ミメシス理論とは、ヒトは音声、言語によるコミュニケーション以前に身振りによるコミュニケーションを行っており、他者の行動の見まねを通じた行動認識、行動生成の相互的な循環により、シンボル操作を基礎とする高度な知能を獲得したとする理論である。ミラーニューロンは、他者の特定の行動

を観察するときに発火し、自己が同じ行動をするときにも発火するとされるニューロン群のことであり、これが類人猿の脳の言語野に存在するという発見がなされた。人間の脳にもミラーニューロンが存在するという説も報告され、言語やシンボル操作と見まね学習との関係が注目されている [14]。

筆者らはミメシスモデルとして、多自由度の時系列信号として表現される人間の全身運動パターンを記号化可能であることが示されている隠れマルコフモデル (HMM) を採用する。各 HMM は対応する運動パターンを学習し、HMM のパラメータ集合として記号化・記憶する。学習した HMM を運動パターンの時系列信号をモデルパラメータとして離散化・低次元化した抽象的な表現としてみなすことによって原始シンボル (*proto-symbol*) と定義する。運動学習において各運動パターンに対して HMM の尤度が最大となるように HMM パラメータを推定することによって人間の運動パターンの獲得を行っているが、ここで用いる運動パターンはあらかじめ設計者が運動の境界を見つけ、分節化した時系列データであった。このような設計者により分節化された運動パターンを記号化するアプローチは人間が直感的に理解しやすい原始シンボルをロボットに付与することができるが、膨大な運動パターンを学習する場合において設計者への作業負担が増大することや未知の運動を自律的に学習することが不可能な問題がある。ロボットが他者行動の観察を通じて自

原稿受付 2008 年 3 月 28 日

*東京大学大学院情報理工学系研究科

*Graduate School of Information Science and Technology, The University of Tokyo

■ 本論文は有用性で評価されました。

律的に運動パターンを分節化・記号化するアプローチは、人間とは異なる理解不能な原始シンボルを運動知識として蓄積していくことがあるが、膨大な運動パターンの学習や未知な運動の学習には必要不可欠な機能である。

本論文では人間の運動に関して、頻繁に観察される運動パターンの教師なし学習を行い、リアルタイムで運動パターンの境界を検出する分節化手法 (RUSS: Realtime Unsupervised Self-tuning Segmentation) を提案する。本手法は、一連の運動データを一定の時間幅をもつ時系列データに区切り、切り出された運動時系列データを HMM の特徴量に変換、圧縮することによって運動パターンを特徴量の時系列として表現する。変換された特徴量時系列から相関行列を獲得し、特徴量の出現規則を学習する。獲得した規則が不明瞭な時点が運動パターンの境界と定義することによって行動の分節化を行う。さらに、ヒューマノイドロボットが分節化された運動データを抽象化することによって自律的に原始シンボルを獲得する知能の枠組みを提案し、その有効性を示す。

2. 関連研究

動作の分節化の研究として、Mori ら [15] は、教師あり分節化手法を提案している。提案モデルは、日常の動作においてまとまりのある単位としての動きの切れ目は曖昧であるという考えに基づき、区切れ目の広がりや動作の分節可能度の分布として表現している。分節可能度の計算は、人間の身体の階層性にも着目し、木構造を用いることによって体の動かし方を定性的に表現する。推定モデルの学習では、いくつかの運動データに対して、各フレームごとに分節可能度を 4 段階で被験者に評価してもらい、この評価値を教師データとして採用することによってより人に近い分節化基準をもつ推定モデルを獲得できる。しかし、運動パターンごとに被験者が教師データを入力する作業が必要であり、分節化したい運動パターンの数に比例して被験者に掛かる負荷が増大するため、膨大な運動データの分節化のモデル作成は容易ではない。

Wang ら [16] は、教師なし分節化法として compressive 手法を提案している。compressive 手法は、入力の時系列データに関して圧縮率をデータ長およびその時系列の出現回数、頻度から算出し、圧縮率の大きい時系列を一つのまとまりのある単位とし、この処理を繰り返すことによって時系列データを分節化する方法である。compressive 手法を用いることによって指揮者の周期的な手の運動から運動パターンの抽出が可能となる。この方法は、入力データを記憶しておき、ある特定の時系列パターンの繰り返し回数を数える必要があるために、長時間の運動動作に対してはデータを保存するメモリへの負荷が増大になり、リアルタイムで圧縮率の高い動作単位を逐次的に獲得することは困難である。

Shiratori ら [17] は、一連の動作はプリミティブな動作単位のシーケンスであるという仮定 [18] に基づき、ダンスデータからプリミティブな動作の抽出を行っている。動作単位の抽出には、手先、足先およびボディ重心の速度に着目し、各部位の速度がゼロ付近になる時点をキーポーズの候補とし、音楽のリズム情報を用いてキーポーズの選定を行う。これは、キーポーズや音

楽といったダンスの特徴とそれに付随する情報を用いた教師なし分節化の方法であるが、日常動作において、このような特徴や情報を前提とすることは適切ではない。

Kim ら [19] も、同様に教師なし分節化手法を用いてリズムカルな運動からプリミティブな運動動作の抽出を行っている。各関節ごとに対して、速度がゼロになる時点を求め、検出した時刻系列を正弦関数によって近似する。各関節ごとの正弦関数を重ね合わせて得られた動作の区切れ点の参照関数に基づき、一連の運動からプリミティブな動作の抽出を行っている。この研究では、参照関数が、周期関数で表されているために運動プリミティブの周期が同一であるとして扱われている。

また、自然言語処理や音声認識などの分野においても文章や会話から単語の抽出は欠かすことのできない情報処理過程であり古くから分節化の研究がなされている [20]。これらの分節化手法は、大きく分けて文章や発言の終端と単語の終端の類似性に着目した手法 (Utterance-boundary strategy) [21] [22]、先行する音素や文字から次に現れるものの予測に基づく手法 (Predictability strategy) [23] [24]、逐次獲得した単語との一致、不一致判定を通じて新規の単語を蓄えていく手法 (Word-Recognition strategy) [25] が挙げられる。Utterance-boundary strategy や Predictability strategy はニューラルネットワークなどのコネクションリストモデルを適用しており、入力データに対してリアルタイムでの分節化は可能であるが単語の認識を行うことはできない。モデルの学習に関しても時間が掛かりリアルタイムでの学習は困難であると思われる。一方、Word-Recognition strategy は、音素または文字の各組み合わせに対して既存の単語との一致、不一致判定を行うため、分節化と同時に単語の認識が可能であるが、リアルタイムでの分節化を行うモデルは提案されていない。

自然言語の分野で処理される文章や会話は単語の記号列から構成されており、人間の手により単語間の区切りを明確に与えることができるため、その区切り位置を分節化モデルの教師信号として用いることが可能である。また、使われる単語をあらかじめデータベースとして蓄えておくことによって、データベース中の単語との照合を通じて文章中の単語間の区切り位置を検出することができる。しかし、言語と運動の情報処理に多くの類似点はあるものの、運動は連続であり運動パターン間の境界は明白ではないことから、運動パターンの分節化は一層困難な問題である。

3. 運動パターンのリアルタイム教師なし分節化法

3.1 特徴 HMM の獲得

人間の連続した運動時系列データを少数フレーム数 w_{span} 毎の観測データに区切り、その i 番目の観測データを $\mathbf{o}(i)$ とする。

$$\mathbf{o}(i) = [\hat{\mathbf{o}}((i-1)w_{span} + 1) \cdots \hat{\mathbf{o}}(iw_{span})] \quad (1)$$

ここで、 $\hat{\mathbf{o}}(t)$ は時刻 t における運動データの列ベクトルとする。この少数フレームの観測データ $\mathbf{o}(i)$ を分類する状態数の少ない HMM を総数 N_D 個用意する。HMM は時系列データを入力とし、その時系列データがどのカテゴリに含まれるかを決定するための確率モデルであり、音声認識等に活用されているモ

Fig. 1 A procedure for optimization of a feature HMM

デルである [26]. HMM は、状態の集合 $\mathbf{Q} = \{q_1, \dots, q_n\}$ 、状態 q_i から状態 q_j への遷移確率 a_{ij} の集合である状態遷移確率行列 $\mathbf{A} = \{a_{ij}\}$ 、状態 q_i でのベクトル $\mathbf{x}$ を出力する確率分布の集合 $\mathbf{B}$ 、および初期状態確率の集合 $\mathbf{\Pi} = \{\pi_i\}$ の 4 組の集合 $\lambda = \{\mathbf{Q}, \mathbf{A}, \mathbf{B}, \mathbf{\Pi}\}$ によって表される。

N_D 個の HMM ($\lambda_k : k = 1, 2, 3, \dots, N_D$) の競合学習手順を示す (Fig. 1)。

- step1 HMM λ_k の初期パラメータをランダムに設定する。
- step2 観測データ $\mathbf{o}(i)$ に対して、HMM λ_k の尤度 $P(\mathbf{o}(i)|\lambda_k)$ を算出する。
- step3 観測データ $\mathbf{o}(i)$ に対して、最大尤度を出力する HMM(λ_{max}) を求める。

$$\lambda_{max} = \arg \max_{\lambda_k} P(\mathbf{o}(i)|\lambda_k) \quad (2)$$

- step4 HMM(λ_{max}) のパラメータを最適化するための観測データとして、観測データ $\mathbf{o}(i)$ を保存する。
- step5 HMM(λ_{max}) パラメータの最適化のために保存した観測データ数がある閾値に達すれば、保存した複数の観測データを用いて HMM(λ_{max}) パラメータを最適化し、保存されていた観測データを消去する。保存した観測データ数がある閾値未満の場合は、HMM(λ_{max}) パラメータの最適化を行わない。HMM パラメータの最適化は、EM アルゴリズムの一種である Baum-Welch アルゴリズム [27] を用いる。
- step6 step2~step5 までの処理回数が閾値回数以上に達した時点で、十分に学習が行われたとみなして HMM λ_k の競合学習を終了する。step2~step5 までの処理回数が閾値回数未満の場合は、次の観測データ $\mathbf{o}(i+1)$ を入力として step2 に戻る。

step1~step6 の手順により各 HMM(λ_k) は少数フレーム数の観測データの分類および特徴の学習を行っているため、HMM(λ_k) を特徴 HMM と定義する。

3.2 特徴 HMM 時系列を用いた相関学習

観測データ $\mathbf{o}(i)$ を出力する尤度が最大の特徴 HMM の時系列は、運動のまとまりの単位ごとにおいては、統計的規則があると考えられる。そこで、特徴 HMM の時系列の関係性に着目し、その関係性を相関学習によって獲得する。

まず、観測データ $\mathbf{o}(i)$ が特徴 HMM から生成される尤度を計算し、算出した尤度の大きい順に特徴 HMM を N_S 個選択する。選択された特徴 HMM の集合を $S(i) = \{\lambda^1, \lambda^2, \dots, \lambda^{N_S}\}$ と置く。

k 番目の特徴 HMM(λ_k) が集合 $S(i)$ に含まれる場合、 k 次元目の要素を 1、集合に含まれない場合、その要素を 0 とするような N_D 次元の列ベクトル $\hat{\mathbf{x}}(i)$ を求める。

$$\hat{\mathbf{x}}(i) = [\hat{x}_{i1} \hat{x}_{i2} \dots \hat{x}_{iN_D}]^T \quad (3)$$

$$\hat{x}_{ik} = \begin{cases} 1 & \lambda_k \in S(i) \\ 0 & \lambda_k \notin S(i) \end{cases} \quad (4)$$

ここで、 T は行列の転置を示す。過去の履歴情報を含めたベクトル $\mathbf{x}(i)$ を新たに定義する。式 (6) に示すように、 $\mathbf{x}(i)$ は $\hat{\mathbf{x}}(i-j)$ ($j = 0, 1, \dots, M-1$) を列方向に並べ、規格化された列ベクトルである。

$$\mathbf{x}^*(i) = [\hat{\mathbf{x}}(i-M+1)^T \dots \hat{\mathbf{x}}(i)^T]^T \quad (5)$$

$$\mathbf{x}(i) = \frac{\mathbf{x}^*(i)}{\|\mathbf{x}^*(i)\|} \quad (6)$$

ここで、 M は相関学習に用いる履歴情報の時間幅を表す。このように過去の履歴情報を用いることによって直前の特徴量との相関が強くなり過ぎることにより生じる運動パターンを細かく分節化することを防ぐ。また、以後 $\mathbf{x}(i)$ を特徴ベクトルと呼ぶことにする。特徴ベクトルを用いた相関学習によって相関行列を求める。

相関学習は K 組の入出力ベクトル $\{\mathbf{u}_l, \mathbf{y}_l \mid l = 1, 2, \dots, K\}$ が与えられたとき、入出力関係を相関行列 $\mathbf{W}_0 = \sum_{l=1}^K \mathbf{y}_l \mathbf{u}_l^T$ によって表現する手法である。この相関行列によって、与えられた入力ベクトル $\mathbf{u}_l$ ($l = 1, 2, \dots, K$) が互いに直交するとき、 $\mathbf{y}_l = \mathbf{W}_0 \mathbf{u}_l$ の関係を満たすことから入力ベクトルから出力ベクトルが正しく想起できる。また、入力ベクトルが互いに直交関係を満たさない場合においても、想起される出力ベクトルは実出力ベクトルに近い値をとる。

提案する分節化手法において相関学習は、現時刻の特徴ベクトル $\mathbf{x}(i)$ を出力ベクトル、直前の特徴ベクトル $\mathbf{x}(i-1)$ を入力ベクトルとみなし、式 (7) に示すように逐次的に相関行列を更新することによって行う。

$$\mathbf{W}(i) = \alpha \mathbf{W}(i-1) + (1-\alpha) \mathbf{x}(i) \mathbf{x}(i-1)^T \quad (7)$$

ここで、 α は安定化係数を表す。 $\mathbf{x}(i) \mathbf{x}(i-1)^T$ は直前の特徴ベクトル $\mathbf{x}(i-1)$ から現時刻の特徴ベクトル $\mathbf{x}(i)$ への写像を表す。式 (7) は、過去の特徴ベクトルのダイナミクス $\mathbf{W}(i-1)$ に現時刻での特徴ベクトルのダイナミクスを逐次足し合わせていくことによって、特徴ベクトルのダイナミクス $\mathbf{W}(i)$ を求めている。この特徴ベクトルのダイナミクスを表現する $\mathbf{W}(i)$ を用いることによって特徴ベクトルの履歴情報から次の特徴ベクトルを予測することが可能となる。

3.3 運動パターンの境界判定

まとまりのある運動単位内においては、現在の動きは過去の

Fig. 2 An illustration for the segmentation method

運動履歴から受ける影響が大きく、履歴情報から現在の動きを容易に予測できると思われる。一方、ある運動パターンから他の運動パターンに移移する時点において、過去の情報から次の動きを予測することは困難である (Fig. 2)。このような仮定に基づき現在の特徴ベクトル $\mathbf{x}(i)$ と履歴情報による予測特徴ベクトル $\mathbf{W}(i-1)\mathbf{x}(i-1)$ との誤差 $E(i)$ を式 (8) のとおりに算出する。

$$E(i) = \|\mathbf{x}(i) - \mathbf{W}(i-1)\mathbf{x}(i-1)\| \quad (8)$$

この誤差を運動の予測不確実性と捉え、予測不確実性の大きい時点の運動は未知の運動、予測不確実性の小さい時点の運動は既知の運動パターンの一部として判断できる。したがって、式 (8) が閾値 E_{th} を越える時点運動パターンの境界 b として運動パターンを抽出する。すなわち運動パターンの境界 b は下式を満足する時点として検出される。

$$\begin{aligned} E(b-1) &> E_{th} \\ E(b) &< E_{th} \end{aligned}$$

身体部位の速度がゼロとなる時点を検出し、その時点運動境界と見なすのではなく、多数出現する運動パターンを一つの運動のまとまりであると判定することによって運動データを分節化することが、提案する分節化手法の特徴である。

4. 原始シンボルの競合学習法

これまで、運動パターンを HMM により抽象化する学習フェーズ、抽象化して得られた HMM のパラメータセットを原始シンボルと位置づけ観測された運動データに対して各 HMM の尤度に基づく運動パターンの認識フェーズおよび原始シンボルから確率的に運動データの時系列を再生する運動生成フェーズからなる学習、認識、生成の統合的な数理モデルが提案されている [7] [8]。提案されているモデルにおいて、設計者があらかじめ分節化し、運動パターンごとに分類した運動データを教師信号として原始シンボルを獲得する必要がある。本章では、提案した分節化手法を用いて自動的に分節化された運動データから自律的に原始シンボルが獲得される競合学習の枠組みを提案する。

原始シンボルの自律的に獲得する枠組みとして、前節の特徴

HMM の学習と同様に競合学習を用いる。原始シンボルの獲得過程は以下のとおりである。

- step1 N_U 個の $HMM(\Lambda_k, k = 1, 2, 3, \dots, N_U)$ を用意する。各 $HMM(\Lambda_k)$ の初期パラメータをランダムに設定する。
 step2 i 番目に分節化されたパターンデータ $\mathbf{o}_{segment}(i)$ に対して、最大尤度を出力する $HMM(\Lambda_{max})$ を求める。

$$\Lambda_{max} = \arg \max_{\Lambda_k} P(\mathbf{o}_{segment}(i) | \Lambda_k) \quad (9)$$

- step3 $HMM(\Lambda_{max})$ のパラメータを最適化するための訓練データとして、パターンデータ $\mathbf{o}_{segment}(i)$ を保存する。
 step5 $HMM(\Lambda_{max})$ パラメータの最適化のために保存したパターンデータ数がある閾値に達すれば、保存した複数のパターンデータを用いて $HMM(\Lambda_{max})$ パラメータを最適化し、保存されていたパターンデータを消去する。保存したパターンデータ数がある閾値未満の場合は、 $HMM(\Lambda_{max})$ パラメータの最適化を行わない。
 step6 step2～step5 までの処理回数が閾値回数以上に達した時点で、十分に学習が行われたとみなして $HMM(\Lambda_k)$ の競合学習を終了する。step2～step5 までの処理回数が閾値回数未満の場合は、次のパターンデータ $\mathbf{o}_{segment}(i+1)$ を入力として step2 に戻る。

step1～step6 に示すように、分節化されたパターンデータを $HMM(\Lambda_{max})$ に分類し、その $HMM(\Lambda_{max})$ のパラメータを最適化することによって $HMM(\Lambda_{max})$ は類似したパターンデータを表現した原始シンボル HMM へと発展していく。

運動パターンの認識は、観測された運動データに対する各原始シンボル HMM の尤度 $P(\mathbf{o}_{segment} | \Lambda_k)$ の最大となる原始シンボル HMM を認識結果とする。

$$\Lambda_{KMAX} = \arg \max_{\Lambda_k} P(\mathbf{o}_{segment} | \Lambda_k) \quad (k = 1, 2, 3, \dots, N_U) \quad (10)$$

原始シンボルからの運動生成について、初期状態を q_1 とし、状態遷移確率 $\mathbf{A}$ に従って状態遷移系列を算出する。ここで、終端状態に到達した時点運動パターンの終わりとして定義する。状態遷移系列の長さは、確率的に決定されるため数回の状態遷移系列を計算し、平均化することによって状態遷移を求める。算出された状態遷移に従って、状態ごとの確率密度関数 $\mathbf{B}$ によって運動データ時系列を求める。これらの状態ごとに生成されるデータも確率的に決定されるために数回の試行を重ねて平均化することによって運動パターンの生成を行っている。

5. モーションキャプチャを用いたリアルタイム分節化の実験結果

実際の人間の運動に適用することによって提案した運動パターンの分節化手法の有効性を検証した。34 個のマークを取り付けた人の運動を MotionAnalysis 社製の光学式モーションキャプチャにて計測する。キャプチャのサンプリング時間は 50 [ms] である。各マークの計測位置から 20 自由度のヒューマンフィギュアモデル (各腕 4 自由度、各足 6 自由度) のボディの位置、姿勢および各関節角をリアルタイムで逆運動学を解くことによ

て算出する [28]. ボディーの位置・姿勢および全関節角の時系列によってヒューマンフィギュアモデルの運動は一意に決定されるが, 絶対座標系におけるボディの水平位置およびボディのヨー角の差異によって異なる運動パターンとして運動を蓄積していくのは好ましくない. さらに, 関節角度はヒューマノイドロボットの運動生成において運動指令値として直接利用可能である. 本稿では, 運動はボディーの高さ, ロール, ピッチ角および関節角の時系列によって表現されるとした. そこで, 5 フレーム, 46 次元のデータ (ボディの高さ, ロール, ピッチ, 各関節角およびこれらの速度) $\mathbf{o}(i)$ から特徴 HMM の最適化を行う. この HMM は状態数 3, 各状態に 46 次元確率密度関数を持つ Left-to-Right 型 HMM である. 特徴 HMM の総数を $N_D = 50$ とした. 相関学習における特徴ベクトルは, 履歴情報の時間幅 $M = 4$ の $\mathbf{x}(i) = [\hat{\mathbf{x}}(i-3)^T, \hat{\mathbf{x}}(i-2)^T, \hat{\mathbf{x}}(i-1)^T, \hat{\mathbf{x}}(i)^T]^T$ と表されるような 200 次元のベクトルとした. また, 安定化定数は $\alpha = 0.99$ とした.

実験では被験者の行う運動は, 「右キック」「左キック」「右パンチ」「左パンチ」「右足一步後退」「左足一步後退」「かがむ」の 7 種類の運動パターンであり, 順不同に被験者はこれらの運動パターンを行うものとした. Fig. 3 は, モーションキャプチャによって被験者の一連の運動を計測し, リアルタイムで運動パターンの境界を検出した実験結果, および (左右) 股関節角・肩関節角・全関節の平均角速度の時系列を表している. Fig. 3 は, 提案した手法が一連の運動を「左パンチ」「かがむ」「右キック」「左パンチ」「右パンチ」「左パンチ」「かがむ」と運動パターンに分節化しており, 設定した 7 種類の運動パターン間の遷移時刻を運動パターンの境界として判定していることを示している. また, 全関節の平均角速度の時系列からも身体部位の速度がゼロになる時刻を運動パターンの境界として検出する手法により運動が細かく分節化されることなく, 多数出現する運動時系列を運動パターンとして抽出できていることが確認できる.

提案する分節化手法の計算時間について調べた. モーションキャプチャシステムによって計測されるマーカー位置から 20 自由度ヒューマンフィギュアモデルを用いた逆運動学計算に要する平均時間は 7.3 [ms], 計算された運動時系列データから全特徴 HMM の尤度の平均計算時間は 0.3 [ms] [2.1 節 step2, step3], 特徴 HMM の尤度計算を通じて得られる特徴ベクトルから運動パターンの境界判定の計算時間は 0.6 [ms] [式 (8)] である. さらに, 学習段階では特徴 HMM の学習計算 [2.1 節 step5] や相関行列の更新 [式 (7)] が必要であるが, 各計算時間は 6.3 [ms], 0.8 [ms] である. 学習を含めた全工程に要する時間は 15.3 [ms] であり, モーションキャプチャのサンプリング時間の 50 [ms] 以内に計算が完了する. このことから提案する分節化手法がリアルタイムでの運動の分節化に有効であることを確認した. なお, 実験で用いた計算機の CPU は 3.6 GHz Xeon である.

また, 分節化された運動パターンに関してクラスタ分析を用いて提案した分節化手法の有用性を検証した. 実験で得られた 333 個の切り出された運動パターンに対して互いの非類似度 (距離) を求める. 分節化された運動パターンはフレーム数が異なるため各運動パターンを HMM (運動パターン HMM と呼ぶこ

Fig. 3 Snapshots of a human figure shows segmentation result for a sequence of captured motion data. Dashed lines denote boundaries of motion patterns determined by the proposed segmentation method. The capture data is divided into “left punch”, “bend”, “right kick”, “left punch”, “right punch”, “left punch” and “bend” motion patterns in real-time. These segmented motion patterns are recognized as 10th, 12th, 1st, 10th, 2nd, 10th and 12th HMM respectively. Three graphs at the bottom shows the time profiles of “right hip joint angle”, “left hip joint angle”, “right shoulder joint angle”, “left shoulder joint angle” and the average angular speed of all the 20 joints

とにする) に変換し, その HMM 間の非類似を表す Kullback Leibler 情報量を各運動パターンの距離として用いる. すなわち, 運動パターン HMM は一つの運動パターンデータのみを学習データとして最適化されている. また, 運動パターン HMM 間の Kullback Leibler 情報量は式 (11) として計算することができる.

$$D^*(\lambda_i^p, \lambda_j^p) = \frac{1}{T_{Gi}} \ln P(\mathbf{o}_{Gi} | \lambda_i^p) - \ln P(\mathbf{o}_{Gi} | \lambda_j^p) \quad (11)$$

ここで, k 番目の運動パターン HMM を λ_k^p ($k = 1, 2, 3, \dots, 333$) として記述し, $\mathbf{o}_{Gk}$ は k 番目の運動パターン HMM から生成されるフレーム数 T_{Gk} の運動データである. ここで, Kullback-Leibler 情報量 $D^*(\lambda_i^p, \lambda_j^p)$ は対称性 ($D^*(\lambda_i^p, \lambda_j^p) = D^*(\lambda_j^p, \lambda_i^p)$) を満足しないため, 式 (12) に示す非類似度を定義する.

$$D(\lambda_i^p, \lambda_j^p) = \frac{D^*(\lambda_i^p, \lambda_j^p) + D^*(\lambda_j^p, \lambda_i^p)}{2} \quad (12)$$

この非類似度 $D(\lambda_i^p, \lambda_j^p)$ を各運動パターン HMM の距離とみなし、この距離をできる限る満たすように運動パターン HMM を多次元空間上に配置する。空間の配置には多次元尺度法を用いた。本稿での多次元尺度法では、式 (13) で表される対象 i と対象 j 間の実距離 f_{ij} の二乗と空間上での距離 d_{ij} の二乗の重み付け誤差最小二乗法により各対象を多次元空間上に布置する。

$$E = \frac{1}{N C_2} \sum_{i=1}^N \sum_{j=i+1}^N \frac{(f_{ij}^2 - d_{ij}^2)^2}{4 f_{ij}^2} \quad (13)$$

式 (13) で表される対象の距離二乗の誤差の二乗和は、各対象の多次元空間上での座標値の四次多項式として表現できるため、最適化計算が容易になる。

Fig. 4 は、多次元空間の空間次元数と式 (13) で表される距離誤差の関係を表している。 d 次元空間での距離誤差 E_d と $d+1$ 次元空間での距離誤差 E_{d+1} の変化率 ($\gamma = (E_d - E_{d+1})/E_d$) が $d=4$ で 0 となり、四次元空間で多次元尺度法の計算は収束していることが確認できる。すなわち、本実験で得られた運動パターンを多次元空間上の点として布置することによって、運動の低次元化を行う場合は空間の次元数は 4 で十分であると判断できる。

Fig. 5 は、このように分節化された運動パターンを運動パターン HMM へ変換し、各 HMM を多次元尺度法によって四次元空間上に布置して得られる運動空間を表す。Fig. 5 に示される運動空間上の点の色および形は、その運動パターンが属するクラスターを表している。クラスター分析は、自律的に獲得されたシンボル (次章に詳細を述べる) を用いた認識手順に従う。すなわち、分節化された運動パターンを生成する尤度が最大の原

始シンボルを求め、その原始シンボルにあらかじめ人手により与えられたシンボルラベル (「右キック」「左キック」「右パンチ」「左パンチ」「右足一歩後退」「左足一歩後退」「かがむ」) ごとに運動パターンをクラスターに分類した。Fig. 5 から同一のクラスターに分類される運動パターンが四次元空間上において密に集まったクラスター構造を形成していることが分かる。この運動パターンが密に集まるクラスター構造は、そのクラスターに属する類似した運動パターンが多く存在することを意味し、提案した手法が頻繁に現れる運動時系列データをパターンとして抽出することに有効であることを示している。

このようにモーションキャプチャを用いた実験より、人間の運動に対して提案した分節化手法がリアルタイムで適切に運動境界を検出できることを確認した。

6. 運動パターンの分節化を通じた運動の記号化・認識・生成

6.1 原始シンボルの自律的獲得

リアルタイム教師なし分節化および原始シンボルの競合学習法を用いてヒューマノイドロボットが自律的に原始シンボルを獲得する実験を行った。実験は前章で述べたモーションキャプチャシステムにて計測した人の運動から自動的に切り取られた運動パターンを学習データとしてロボットがリアルタイムに原始シンボルを獲得する。ここで原始シンボル HMM は、状態数 10、各状態に 23 次元確率密度関数 (ボディの高さ、ロール、ピッチおよび 20 個の関節角の 23 次元) を持つ Left-to-Right 型 HMM である。原始シンボル HMM の総数を $N_U = 20$ とした。長時間の人の動きの観察からヒューマノイドロボットは 15 個の原始シンボルを獲得した。競合学習法により獲得した原始シンボルの有効性を原始シンボル空間を用いて検証した。原始シンボル空間を原始シンボルと前節の運動パターン HMM の合計 438 個の HMM より構成した。**Fig. 6** は構築した原始シンボル空間を表す。原始シンボルが運動パターンが形成する各クラスター内に配置されていることから、原始シンボルが各クラスターの運動パターンの特徴的なダイナミクスの汎化能力を獲得していることが分かる。このように、リアルタイム教師なし

Fig. 4 The relationship between the number of spatial dimension and distance error

Fig. 5 Space of segmented motion patterns. Segmented motion patterns are converted into HMMs respectively. The HMMs is optimized by using one segmented motion pattern as a training data. The HMMs are located in 4 dimensional space based on dissimilarities among themselves. The shape and the color of each mark indicate the cluster which each segmented motion pattern belong to. The upper figure projects the 4 dimensional space onto the 3 dimensional space for visualization. The lower figures project the 4 dimensional space onto 2 dimensional spaces

Fig. 6 The proto symbol space consists of proto symbols and segmented motion patterns. The fill color mark and blank mark correspond to the proto symbol and the motion pattern respectively. The proto symbols are located close to the cluster area which includes the motion patterns that are recognized as the proto symbols

Fig. 7 Generated motion patterns by each proto symbol. The motion patterns of (A)–(G) can be subjectively classified as “right kick”, “left kick”, “right punch”, “left punch”, “retract the right leg”, “retract the left leg” and “bend”

分節化法と原始シンボルの競合学習法を併用することによって原始シンボル HMM が運動パターンを逐次学習し、運動パターンの抽象的表現へと発展していくことが確認できた。しかし、ロボットが観察した運動パターンは主観的に「右キック」、「左キック」、「右パンチ」、「左パンチ」、「右足一歩後退」、「左足一歩後退」、「かがむ」の七つの運動パターンに分類できる。しかし、上記の実験結果では 15 個の原始シンボルが獲得されており、七つの運動パターンとして分類されていない。Fig. 6 に示すように類似した運動パターンを複数の原始シンボルとして分類している。例えば、「かがむ」運動パターンを三つの原始シンボルとして分類した結果が得られている。このように、クラスタリングでよく問題になるいくつかのクラスタで分類するのと同様に、与えられた運動パターンデータから適切な数の原始シンボルによって運動パターン进行分类する課題が残されている。

6.2 原始シンボルを用いた運動認識

分節化を通じて獲得した原始シンボルを用いて運動の認識実験を行った。Fig. 3 の“HMM”ラベルに続く数字が、分節化されて抽出された運動パターンに対する認識結果を示している。Fig. 3 の 3 種類の「左パンチ」運動パターンはすべて 10 番目の原始シンボル HMM として認識され、同様に、2 種類の「かがむ」運動パターンはともに 12 番目の原始シンボル HMM と認識されている。このように、同じ運動パターンに含まれる運動データを同一の原始シンボルと認識していることから、獲得されたシンボルが運動認識に有効であることが確認された。

Fig. 8 The humanoid robot can perform each motion pattern generated by the proto symbol. (A)–(G) represents the generated motion patterns of “right kick”, “left kick”, “right punch”, “left punch”, “retract the right leg”, “retract the left leg” and “bend” respectively

6.3 原始シンボルからの運動生成

獲得した原始シンボルを用いて運動の生成実験を行った。Fig. 7 は原始シンボル空間の各クラスターに属する原始シンボルの一つから生成される運動パターンを表している。生成された運動パターンは、「右キック」「左キック」「右パンチ」「左パンチ」「右足一歩後退」「左足一歩後退」「かがむ」とラベル付け可能であることが分かる。さらに、Fig. 8 は小型ヒューマノイドロボットが各原始シンボルから生成された運動パターンを振る舞った結果を表している。獲得した原始シンボルがロボットの運動生成にも有効であることを確認した。

以上のように、運動パターンの記号化・認識・生成の観点か

らも提案した運動の分節化手法がヒューマノイドロボットの知能の枠組みの基盤技術として有用であることを確認した。

7. 大規模運動データを用いた運動パターンの分節化・記号化

ヒューマノイドロボットが観察した行動を原始シンボルとして記号化し、運動の知識を蓄えていく知能構築アプローチにおいて、大規模な運動パターンにも適用可能であることが求められている。しかし、ロボットが原始シンボルとして学習・記憶する運動が多様になるにつれて、従来のように人手により運動パターンを切り出して教え込む方法は非現実的である。ロボットが膨大な運動から自律的に運動パターンを抽出し、その運動パターンを原始シンボルとして記憶するアプローチは、ロボットが膨大な運動パターンから知識を構築するうえで必要不可欠な基盤技術である。本節では、膨大な運動データに対する教師なしリアルタイム分節化手法の有効性を長時間の運動データを用いて実験した。先述と同様に運動データは人間の動きをモーションキャプチャシステムにて計測して得られるボディの高さ、ロール、ピッチおよび関節角の時系列データを用いた。本実験では約2時間48分の野球にかかわる運動を採用した。なお、特徴HMMは状態数3、各状態に46次元確率密度関数を持つLeft-to-Right型、総数を $N_D = 300$ とした。相関学習の特徴ベクトルは、4ステップ前までの履歴情報を保持している($M = 4$)。また、安定化定数は $\alpha = 0.999$ とした。

このような特徴HMMおよび相関学習の条件によって、分節化された運動パターンから競合学習を通じて $N_U = 50$ の原始シンボルを獲得した。Fig. 9は獲得した各原始シンボルから生成される運動パターンを表している。Fig. 10は獲得した原始シンボルを六次元の空間上に配置してできる原始シンボル空間の平面投影図である。なお、Fig. 10では、多次元空間上に配置された原始シンボルを六つのクラスターに分類し、各クラスターに属する原始シンボルごとの空間配置図を示している。また、分節化されて抽出した運動パターンを多次元空間上に配置したときの運

Fig. 9 Each snapshot shows a motion pattern generated by a proto symbol

動パターン間距離と実距離との誤差と多次元空間の次元数との関係を Fig. 11 に示す。距離誤差変化率 ($\gamma = (E_d - E_{d+1})/E_d$) が $d = 6$ で $\gamma = 0.003$ と十分に小さく収束している。Fig. 10 に Group1 には、「走る (右足を出す)」($\lambda_{13}, \lambda_{15}$) や「歩く (右足を出す)」(λ_{16}) の類似した原始シンボルが空間上で近接した位置に布置されている。同様に、Group4 には「バットを振る」($\lambda_{11}, \lambda_{14}, \lambda_{25}$) や「手を挙げる」($\lambda_{10}, \lambda_{28}, \lambda_{29}, \lambda_{41}$)、Group5 には「座っている」($\lambda_3, \lambda_5, \lambda_7$)、Group6 には「ストレッチ (足の伸展)」($\lambda_4, \lambda_{36}, \lambda_{38}$) の類似した原始シンボル同士が空間上においても類似の関係を保持していることが確認できる。

このように運動の類似関係を保存する多次元空間は、分節化によって抽出された運動パターンの妥当性を評価するうえで有効であると考えられる。すなわち、抽出された各運動パターンを運動パターン HMM に変換し、運動パターン HMM 間の非類似度に基づき運動パターン HMM を多次元空間上の点として布置する。空間上に点として表現される運動パターンをクラスタ分析によってクラスに分類する。各クラス内分散の和 σ_{inner} とクラス間分散 σ_{intra} の比率によって分離度 ($S = \sigma_{intra}/\sigma_{inner}$) を求める。この分離度は、同じクラス内では運動パターンが類似し、異なるクラスでの運動パターンが異なるほど値が大きくなる性質を有する。したがって、分節化されて抽出された運動パターンの分離度が大きいほど、頻繁に現れる類似した運動パターンが抽出されていることを表している。

Fig. 12 は提案する分節化手法によって切り出された運動パターンとランダムに運動パターンの境界を決定して得られた運動データの分離度を表している。ここで、分離度を計算する場合にあらかじめクラスの数を決める必要があり、本稿ではクラス数が 10, 20, 30, 40, 50 の 5 パターンの設定に対する分離度を比較した。なお、提案する分節化手法によって得られた運動パターンの時間長の平均および分散を求め、ランダムに分節化される運動パターンの時間長はこの平均・分散を有する正規分布に従うものとした。Fig. 12 からすべてのクラス数に対して、提案する分節化手法で得られた運動パターンの分離度が高いことが読み取れる。したがって、提案する分節化手法は大規模な運動データに対しても有効であることが確認できる。

8. リアルタイム教師なし分節化法および原始シンボルの競合学習法の設計変数に関する考察

8.1 原始シンボル HMM の状態数

特定の分節化された運動パターンデータを HMM に学習させる。学習した HMM がその運動パターンデータを生成する尤度はその HMM が学習したデータを記号化した正確度・詳細度を示す。そこで、HMM の状態数と HMM から学習データが生成される尤度の関係を調べた。ここではあらかじめ分節化された「かがむ」の運動データを用いた。この運動データは時間長さ 3 [s]、サンプリング時間 10 [ms] の 23 次元時系列データである。Fig. 13 は HMM の状態数と学習データを生成する尤度の関係を表している。状態数を増やすと学習データを生成する尤度は増加する傾向にあるが、状態数が 5 以上になると尤度の増加が収束する。このような人間の運動パターンデータと HMM のモデルの詳細度の関係から、本稿では原始シンボルの HMM の状

Fig. 10 50 proto symbols are located in 6 dimensional space. The proto symbols are classified into 6 groups. Each figure shows symbol space with the proto symbols in the same group

態数を 10 と固定して原始シンボルの獲得実験を行った。

8.2 運動データのサンプリング時間

ロボットが人間の運動を見まねするときの視覚情報の時間分解能に相当するモーションキャプチャのサンプリング時間について検証した。サンプリング時間が 10 [ms], 20 [ms], ..., 300 [ms] の運動データを各 HMM が学習する。用いる運動データは上記と同じ「かがむ」の運動データであり、サンプリング時間 10 [ms] の運動データを元データとし、その元データを間引くことによ

て異なるサンプリング時間の運動データを作成している。各サンプリング時間に対応する HMM から生成される運動データと元データとの誤差によりロボットの見まね学習に十分な視覚時間分解能を調べた。ただし、サンプリング時間 χ [ms] の運動データを学習した $HMM(\lambda^\chi)$ から生成される運動データのサンプリング時間は χ [ms] であり、サンプリング時間 10 [ms] の元データとの誤差を求めるには時間の対応をとる必要がある。 $HMM(\lambda^\chi)$ の第 i 番目の状態における再帰遷移確率 a_{ii}^χ からそ

Fig. 11 The relationship between the number of spatial dimension and distance error

Fig. 12 This graph shows the comparison between proposed segmentation and random one in terms of degree of separation. The score can be computed as the ration of intra-class variation divided by inner-class variation. Appropriate clustering provides high score

Fig. 13 The relationship between the number of HMM nodes and the likelihood that the HMM generates the human motion data. As the number of the HMM nodes increases, the likelihood becomes larger. The increment of the likelihood converges at the HMM with 5 nodes

Fig. 14 The relationship between sampling time and the average error of joint angles between original training data and generated data by HMM. The average errors are maintained below 0.05 rad

の状態の期待滞在時間 τ_i^x

$$\tau_i^x = \frac{1}{1 - a_{ii}^x} \quad (14)$$

が求められる。サンプリング時間 10 [ms] のモデルに変換するために、HMM(λ^x) の全状態の期待滞在時間が $\chi/10$ 倍になるように状態遷移確率を変更する。このようにして求められた HMM から学習データと同じ時間長さの運動データを生成し、各時間フレームごとの生成データと元データとの誤差から平均値を算出する。Fig. 14 は、関節角に対応する 20 次元の時系列に関

する生成データと元データとの平均誤差と運動データのサンプリング時間との関係を示している。ここで状態数 10 の HMM を用いた。時間分解能が 10 [ms]~300 [ms] の範囲内では関節角の平均誤差はほぼ一定の 0.05 [rad] 以内である。この結果から本稿でのモデル構築に用いた運動データのサンプリング時間 50 [ms] は、「かがむ」や「パンチする」といった簡単な人間の運動パターンを表現するには十分な時間分解能であることを確認した。

8.3 特徴 HMM の総数

運動の分節化能力は本稿で提案したモデルの最基盤となる技術であり、特徴 HMM の総数に依存する。そこで、特徴 HMM の総数と分節化性能の関係を調べた。分節化技術の性能を表す指標として「正確度」(accuracy)と「完全度」(completeness)を採用した。Brent らはこれら指標を英語文章の分節化に適用している [25]。英語文章では、単語の切れ目が明らかであるため正確な単語の境界と分節化処理により検出された境界とを比較することが可能である。しかし、人間の運動は運動パターンが時間的に連続に繋がった構造をしており、明確な運動パターンの境界は分からない。そこで、設計者が主観的な評価基準に基づいて運動パターンの境界を検出し、検出された時点を中心とした一定時間範囲内(本稿では ± 1 [s] と設定した)を主観的運動パターン境界とする。この主観的運動パターン境界と提案した分節化手法により検出された運動パターン境界との比較を行うことによって「正確度」および「完全度」を求める。

「正確度」および「完全度」を計算するために、以下に説明する運動パターンの境界点を数え上げる。

(1) True Positive: 主観的運動パターン境界と一致した分節化手法に基づく運動パターン境界

(2) False Positive: 主観的運動パターン境界と一致しない分節化手法に基づく運動パターン境界

(3) False Negative: 分節化手法に基づく運動パターン境界と一致しない主観的運動パターン境界

以上の運動パターンの境界点の個数から「正確度」($Z_{accuracy}$)および「完全度」($Z_{completeness}$)が求められる。

$$Z_{accuracy} = \frac{N_{tp}}{N_{tp} + N_{fp}} \quad (15)$$

$$Z_{completeness} = \frac{N_{tp}}{N_{tp} + N_{fn}} \quad (16)$$

ここで、 N_{tp} , N_{fp} , N_{fn} はおのこの True Positive, False Positive, False Negative の個数である。

Fig. 15 は、モーションキャプチャにて計測した約 4 [min] の運動データを用いて分節化実験を行った結果を表している。ここで、特徴 HMM の総数は 10~110 であり、間隔幅は 10 である。特徴 HMM の総数が 10 の場合は、「完全度」が低い。これは運動データを時間長の大きい運動パターンとして区切る傾向があることを示している。特徴 HMM の総数を 10~20 に増やすことによって、より詳細に運動パターンを抽出することができ「完全度」が改善される。特徴 HMM 数が 20 以上では、「正確度」、「完全度」とともに劇的に改善されることなく、分節化性能の向上は収束する。また、特徴 HMM 数が 20 以上の領域では、「正確度」が 90%、「完全度」が 70% と高い精度で運動パ

Fig. 15 Evaluation score of accuracy and completeness. The number of feature HMMs is set from 10 to 110

Fig. 16 This graph shows the relationship between degree of separation and duration of feature HMMs

ターンが分節化されている。

8.4 特徴 HMM の時間幅 (w_{span})

運動の特徴を表現する特徴 HMM は、少数フレームの運動データ $o(i)$ の時間幅 w_{span} によって変化する。したがってこの時間幅 w_{span} と分節化の性能の関係を調べることは、本稿提案の分節化手法を設計するうえで重要である。分節化の性能指標として先述の分離度を用いることにする。**Fig. 16** は時間幅 w_{span} と分離度の関係を示している。時間幅を変更することによって分離度は変化するが、ランダムに運動を切り出した場合と比較するとすべての時間幅において分離度は高くなっていることが確認できる。このことは、時間幅がある一定の時間分解能を保持する領域内に設定されていれば、本提案手法によって適切に運動パターンを分節化できることを示している。したがって、時間幅 w_{step} について特別の設定・チューニングを要する手法ではないと考えられる。

8.5 相関行列の履歴情報 (M)

本稿で提案する分節化手法は運動データ時系列のダイナミクスを相関行列によって学習し、ダイナミクスの予測に基づき頻繁に現れる運動パターンを抽出するアプローチである。すなわち、運動の予測精度が分節化性能に大きく影響を与える。運動の予測計算において直前の運動時系列の情報が豊かな予測精度が向上する、一方モデルの計算量や複雑度が増大、学習計算が困難になると考えられる。このように運動時系列のダイナミクスを埋め込む相関行列が扱う過去の履歴情報の大きさ (M) と分節化性能には密接なかわりがある。そこで、本節では履歴情報の大きさと分節化性能の関係性について調べた。

Fig. 17 This graph shows the relationship between degree of separation and temporal size of correlation

Fig. 17 は履歴情報の大きさと分節化の性能の関係を表している。ここでも、分節化の性能指標として分離度を用いた。履歴情報の大きさが $M = 2$ から $M = 10$ の範囲内においては、ランダムに運動パターンを分節化した場合と比較して、提案する分節化手法が頻繁に現れる類似した運動パターンを抽出することに有効であることが確認できる。すなわち、前節の時間幅 w_{span} と同様に相関行列の履歴情報の時間幅の変数に関しても特別な設定知識を要する手法ではないと考えられる。

9. おわりに

本研究で得られた成果は以下の3点にまとめられる。

- (1) 人間の一連の動作を特徴 HMM により特徴量の時系列に変換し、特徴量の時間的相関関係に着目した運動パターンのリアルタイム教師なし分節化手法を開発した。
- (2) モーションキャプチャシステムと提案した分節化手法を統合し、リアルタイムで運動データの特徴量量化、特徴量間の相関性の算出、相関性に基づく運動パターンの分節化計算が可能であることを確認した。提案した分節化手法を用いて検出される運動パターン境界が主観的な運動分節点と一致すること、および分節化された運動パターンのクラスター分析から本稿で提案した分節化手法の有効性を示した。
- (3) 運動の分節化を通じてヒューマノイドロボットが自律的に原始シンボルを獲得する競合学習法を提案し、獲得される原始シンボルのクラスター構造から適切に原始シンボルが獲得できることを確認した。さらに、獲得した原始シンボルを用いることによる運動認識・運動生成の実験を行い、提案した分節化手法および原始シンボルの競合学習法がヒューマノイドロボットの知能化の有効な基盤技術であることを実証した。

謝辞 本研究は、科学研究費補助金基盤研究 (S) (研究課題名: 知能の力学的情報処理モデルの展開, 課題番号: 15100002, 研究代表者: 中村仁彦) の支援を受けた。また、相関学習について門根秀樹氏から貴重な助言をいただいた。

参考文献

- [1] Y. Kuniyoshi, M. Inaba and H. Inoue: "Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance," IEEE Transaction on Robotics and Automation, vol.10, no.6, pp.799-822, 1994.
- [2] J. Morimoto and K. Doya: "Hierarchical reinforcement learning for motion learning: learning "stand-up" trajectories," Advanced Robotics, vol.13, no.3, pp.267-268, 1999.
- [3] M.J. Mataric: "Getting humanoids to move and imitate," IEEE Intelligent Systems, vol.15, no.4, pp.18-24, 2000.
- [4] 鮫島, 片桐, 銅谷, 川人: "モジュール競合による運動パターンのシ

- ボル化と見まね学習”, 電子情報通信学会論文誌, vol.J85D-II, no.1, pp.90-100, 2002.
- [5] M. Haruno, D. Wolpert and M. Kawato: “MOSAIC Model for Sensorimotor Learning and Control,” Neural Computation, vol.13, pp.2201-2220, 2001.
- [6] J. Tani and M. Ito: “Self-Organization of Behavioral Primitives as Multiple Attractor Dynamics: A Robot Experiment,” IEEE Transactions on Systems, Man and Cybernetics Part A: Systems and Humans, vol.4, no.4, pp.481-488, 2003.
- [7] T. Inamura, I. Toshima, H. Tanie and Y. Nakamura: “Embodied Symbol Emergence based on Mimesis Theory,” International Journal of Robotics Research, vol.23, no.4, pp.363-377, 2004.
- [8] 江崎: “自己行動要素による他者の行動理解”, 東京大学大学院工学研究科機械情報工学専攻修士論文, 1999.
- [9] A. Billard, S. Calinon and F. Guenter: “Discriminative and Adaptive Imitation in Uni-manual and Bi-manual Tasks,” Robotics and Autonomous Systems, vol.54, pp.370-384, 2006.
- [10] S. Chiappa, J. Kober and J. Peters: “Using Bayesian Dynamical Systems for Motion Template Libraries,” Advances in Neural Information Processing Systems 21, pp.297-304, 2008.
- [11] M. Donald: Origin of the Modern Mind. Harvard University Press, 1991.
- [12] V. Gallese and A. Goldman: “Mirror neuron and the simulation theory of mind-reading,” Trends in Cognitive Science, vol.2, no.12, pp.493-501, 1998.
- [13] G. Rizzolatti, L. Fogassi and V. Gallese: “Neurophysiological mechanisms underlying the understanding and imitation of action,” Nature Reviews, pp.661-670, 2001.
- [14] G. Rizzolatti and L. Craighero: “The mirror-neuron system,” Annual Review of Neuroscience, vol.27, pp.169-192, 2004.
- [15] T. Mori, Y. Segawa, M. Shimosaka and T. Sato: “Segmentation of Sequential Daily-Actions Based on Hidden Markov Models and Human Body Hierarchy,” Proc. of The Second International Workshop on Man-Machine Symbiotic Systems, pp.207-218, 2004.
- [16] T.S. Wang, H.Y. Shum, Y.Q. Xu and N.N. Zheng: “Unsupervised Analysis of Human Gestures,” Proc. of the Second IEEE Pacific-Rim Conference on Multimedia, no.10, pp.174-181, 2001.
- [17] T. Shiratori, A. Nakazawa and K. Ikeuchi: “Detecting Dance Motion Structure through Music Analysis,” Proc. of the sixth IEEE International Conference on Automatic Face and Gesture Recognition, no.5, pp.857-862, 2004.
- [18] T. Flash and N. Hogan: “The Coordination of Arm Movements: An Experimentally Confirmed Mathematical Model,” The Journal of Neuroscience, vol.5, no.7, pp.1668-1703, 1985.
- [19] T. Kim, S.I. Park and S.Y. Shin: “Rhythmic-Motion Synthesis Based on Motion-Beat Analysis,” Proc. of ACM SIGGRAPH2003, 2003.
- [20] M.R. Brent: “Speech segmentation and word discovery: a computational perspective,” Trends in Cognitive Science, vol.3, no.8, pp.294-301, 1999.
- [21] M. Redington and N. Charter: “Probabilistic and distributional approaches to language acquisition,” Trends in Cognitive Sciences, vol.1, no.7, pp.273-281, 1997.
- [22] M.H. Christiansen and S. Curtin: “Transfer of learning: rule acquisition or statistical learning,” Trends in Cognitive Sciences, vol.3, no.8, pp.289-290, 1999.
- [23] J.L. Elman: “Finding Structure in time,” Cognitive Science, no.14, pp.179-211, 1990.
- [24] P. Cairns, R. Shillcock, N. Charter and J. Levy: “Bootstrapping Word Boundaries: A Bottom-up Corpus-Based Approach to Speech Segmentation,” Cognitive Psychology, vol.33, no.2, pp.111-153, 1997.
- [25] M.R. Brent: “An efficient, probabilistically sound algorithm for segmentation and word discovery,” Machine Learning, vol.34, pp.71-106, 1999.
- [26] 中川: 確率モデルによる音声認識. 電子情報通信学会, 1988.
- [27] L. Rabiner and B.H. Juang: Fundamentals of Speech Recognition. Prentice Hall Signal Processing Series, 1993.
- [28] K. Yamane and Y. Nakamura: “Natural Motion Animation through Constraining and Deconstraining at Will,” IEEE Transactions on Visualization and Computer Graphics, vol.9, no.3, pp.352-360, 2003.

高野 渉 (Wataru Takano)

1999 年京都大学工学部物理工学科卒業。2001 年同大学大学院工学研究科精密工学専攻修士課程修了。同年三菱自動車工業株式会社入社。2006 年東京大学大学院情報理工学系研究科知能機械情報学専攻博士課程修了。2006 年同大学特任助教, 2007 年同大学助教, 2009 年同大学講師, 現在に至る。博士 (情報理工学)。ヒューマノイドロボットの統計的脳型情報処理, 自動車の知能に関する研究に従事。(日本ロボット学会正会員)

中村仁彦 (Yoshihiko Nakamura)

1977 年京都大学工学部精密工学科卒業。1982 年同大学大学院博士課程退学, 1982 年同大学助手。1987 年カリフォルニア大学サンタバーバラ校助教授, 準教授。1991 年東京大学工学部機械情報工学科助教授。現在同大学大学院情報理工学系研究科教授。工学博士。非ホロノミックロボット, 脳型情報処理, ヒューマノイドや CG の力学計算, 外科手術用ロボットなどの研究に従事。システム制御情報学会, 計測自動制御学会, 日本コンピュータ外科学会, IEEE, ASME などの会員。日本機械学会フェロー, 日本 IFToMM 会議委員長, IEEE Robotics and Automation Society, Distinguished Lecturer (2006~2007)。(日本ロボット学会正会員)