

統計的相関に基づく動作パターンの分節化

Segmentation of human behavior patterns based on the probabilistic correlation

高野 渉*¹ 中村 仁彦*¹
Wataru Takano Yoshihiko Nakamura*¹ 東京大学
Univ. of Tokyo

Mimesis theory is attractive that human intelligence originated in the interactive communication of motion recognition and generation through imitation for artificial intelligence. We have proposed the mimesis system using hidden Markov models(HMMs) and their parameter sets represent the proto symbols. In the conventional system, when a specific motion pattern is embedded in correspondent HMM, the designers have to segment the motion pattern such as periodic motion pattern. However it is necessary to have the ability of motion pattern segmentation in order to learn through imitation and develop autonomously. Therefore in this paper, we propose the motion segmentation method by the correlation among the movements for a brief period. Moreover we show that it is possible to acquire the proto symbols by providing the segmented motion patterns with the mimesis system.

1. はじめに

ヒューマノイドの知能獲得のアプローチとして、隠れマルコフモデル(HMM)を用いた運動パターンの認識、生成の双方向情報処理の枠組みを提案されている[Inamura04]。このモデルは、認知心理学の分野において提唱されているミメシス理論[Donald91]および生物学におけるミラーニューロンの発見[Gallese98][Rizzolatti01]に基づいている。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、シンボル等の高度な知能を獲得したとする理論である。他人の特定の行動を観察する時に発火し、自分が同じ行動をする時にも発火するミラーニューロンが類人猿、人間の脳の言語野に存在するとする興味深い事実が発見された。この事実は、ミラーニューロンが他者の行動をシンボルとして抽象化し、さらに抽象化したシンボルから自身の行動を生成することに関わることを示唆している。

従来モデルでは、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を学習・抽象化を行う。この抽象化に用いる運動パターンは予め設計者が運動の境界を見つけ、分節化された時系列データである。したがって、従来の枠組みでは自ら観察した結果を分節化し、新たな原始シンボルを獲得することが不可能であった。しかし、ヒトは他者の行動を観察する際に、自ら行動パターン毎に区切って認識および学習を行っている。このような分節化能力は、自律的にシンボルを獲得するなどの知能に必要不可欠である。

特に単語を基本単位とする自然言語処理において、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文章や会話の分節化についての研究がなされている。近年、計算機処理の高速化にともない、言語コーパスにリカレントニューラルネットワークなどの学習データの統計的性質を抽出するのに適しているコネクショニストモデルを適用することによる言語分節化のアプローチが提案されている[Redington97][Christiansen99]。リカレントニューラルネットワークは、文脈依存のモデルとし

て優れているものの、計算コストが大きいなどの問題が考えられる。

そこで、本稿では人間の運動パターンの分節化法を述べる。一定時間幅の運動時系列データを下位のHMMの記号に変換、圧縮することによって運動パターンを記号列として表現する。変換された記号列から相関行列を獲得し、記号列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づけることによって行動の分節化を行う。さらに、分節化された運動データから上位のHMMを学習することによって獲得される原始シンボルから提案した分節化手法の有効性を検証する。この分節化は、リカレントニューラルネットワークを用いた言語分節化法とは異なり計算量が少ないことが特徴である。

2. 人間動作パターンの分節化

2.1 HMMを用いた運動パターンの記号列化

人間の連続した運動時系列データ O を一定の時間幅の観測データ時系列 o 毎に区切り $O = \{o_1, o_2, \dots, o_m\}$ と表現する。この観測データ o_i に対して下位の k 番目 HMM の尤度 $P(o_i|\lambda_k)$ を計算する。ここで、HMM は、状態の集合 $Q = \{q_1, \dots, q_n\}$ 、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 $A = \{a_{ij}\}$ 、状態 q_i でのベクトル x を出力する確率分布の集合 B 、および初期状態確率の集合 $\pi = \{\pi_i\}$ の 4 組の集合 $\lambda = \{Q, A, B, \pi\}$ によって表される。観測データ o_i に対して最大尤度を出力する HMM を選び出すことによって、観測データ時系列を HMM の記号列表現に変換する。Fig.1 は、観測データを HMM 記号列に変換する概略図を表す。なお、最大尤度を出力する HMM は、入力となる観測データ o_i に対して EM アルゴリズムによって逐次最適化される。

2.2 相関学習

運動データから変換された HMM 記号列は、ある運動パターン内において統計的規則があると考えられる。本稿では、前後の記号列の関係に着目し、その関係性を相関学習によって獲得することを目指す。

$x(t)'$ は、下位 HMM と同数の要素数を持つ列ベクトルであり、観測データ o_t に対して第 k 番目の HMM が選択された場合、第 k 番目の要素が 1、他の要素を 0 とする。しかし、ベクトルは時刻 t の運動データ情報のみしか持たないため、過去

連絡先: 高野 渉, 東京大学大学院, 東京都文京区本郷 7-3-1, 03-5841-6381, takano@ynl.t.u-tokyo.ac.jp

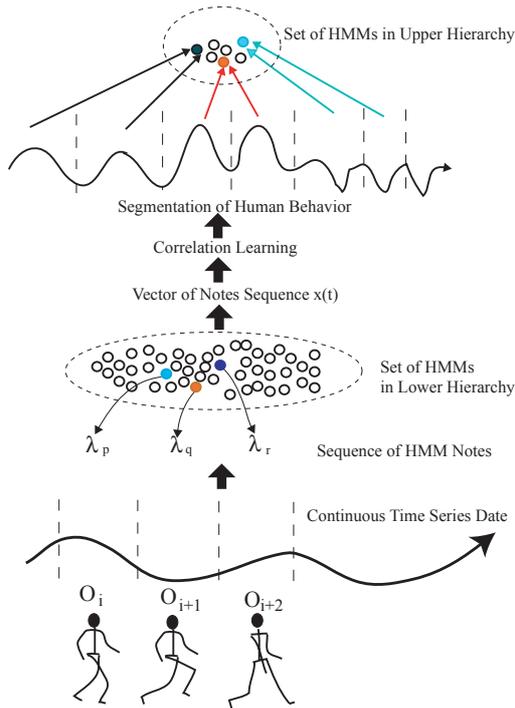


図 1: An illustration for the method of the sequence of HMM notes

の履歴情報も含めた記号列ベクトル $x(t)$ を以下のように定義する.

$$x_t^T = [x(t-N+1)^T, \dots, x(t-1)^T, x(t)^T]. \quad (1)$$

ここで, T は行列の転置を表す. $x(t-1)$ を入力パターン, $x(t)$ を出力パターンと見なすことによって, 相関行列を $W_0 = x(t)x(t-1)^T$ と表すことができる. 入力パターン $x(t-1)$ と出力パターン $x(t)$ が直交関係をみたすならば, 入力パターンから予測した出力パターン $W_0x(t-1)$ が実測の出力パターン $x(t)$ と等しくなり, 相関行列により正確に次の記号列を想起できることを示唆する. また, 入力パターンと出力パターンが互いに直交関係を満たさない場合においても, 想起される出力パターンは実出力パターンと近い値をとることが期待される. ここでは, 逐次学習を行うために以下のように相関行列 $W(t)$ を算出する.

$$W(t) = \alpha W(t) + \eta x(t)x(t-1)^T \quad (2)$$

ただし, α, η は, 各々安定化係数, 学習係数を表す.

2.3 運動パターンの境界判定

ある運動パターン内において, 現時刻の動きと次の動きの間にはある程度の相関があると考えられるので, その特定の運動パターン内において次の動きが, 現在の動きおよび過去の履歴から予測できると思われる. しかし, ある運動パターンから他の運動パターンに遷移する場合は, 過去の履歴情報から次の動きを予測することは困難である. この観点から現在のパターンベクトル $x(t)$ と予測パターンベクトル $W(t-1)x(t-1)$ との誤差 $E(t)$ を計算する.

$$E(t) = \| x(t) - W(t-1)x(t-1) \| \quad (3)$$

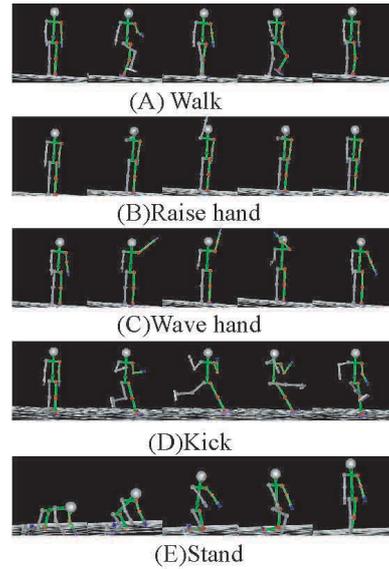


図 2: Motion Patterns by human figure

(A) Walk (B) Raise hand (C) Wave hand (D) Kick (E) Stand

この誤差を不確実性とみなし, 不確実性が大きくなることを運動パターンの切り替わる境界と判定することによって運動パターンの分節化を行う.

3. 分節化の計算結果

前節で述べた人間の運動パターン分節化手法の有効性を確認するために様々な運動パターンをモーションフィギュアを用いて作成した [Yamane03]. 作成した運動パターンは「足踏み」「挙手」「手を振る」「キック」「立ち上がる」の5種類の運動であり, 各種の運動に対してわずかに動きの異なる運動パターンを10個, 合計50の運動パターンを用意した. Fig.2は, 作成した運動パターンを示している. 尚, このモーションフィギュアモデルは, 肘2, 肩2, 腰3, 膝1, 足首2の自由度, すなわち合計20自由度のモデルであり, 各運動パターンのサンプリング時間は33ms, 1つの運動時間は3.3sの100フレームから構成されるようにした.

観測データ o_t の時間幅は330ms(10フレーム)とし, 下位のHMMを50個用意した. 各HMMは, Left-to-Right型, 状態数3, 20次元の確率密度関数とした. 順不同に運動パターンを提示しながら, 逐次HMMの最適化および相関学習を行った. 十分に学習した後「足踏み」「挙手」「手を振る」「キック」「立ち上がる」の運動パターンを10パターンずつ提示した場合の選択されるHMMの記号列をFig.3が示す. ここで, 点線は実際の運動パターンの境界を表す. Fig.3から, 必ずしも同じ種類の運動において同じ記号列が出現するとは限らないことがわかる.

相関学習において本稿では, 過去40フレームの履歴情報を利用するために記号列パターンベクトルを $x(t) = \{x(t-3)^T, x(t-2)^T, x(t-1)^T, x(t)^T\}$ と表されるような200次元のベクトルとした. このパターンベクトルを用いて十分に相関学習を行った後, Fig.3と同時刻において獲得した相関行列から想起されるパターンベクトルと実パターンベクトルの誤差を求めた. Fig.4は, 誤差の時系列変化を表し, 点線は

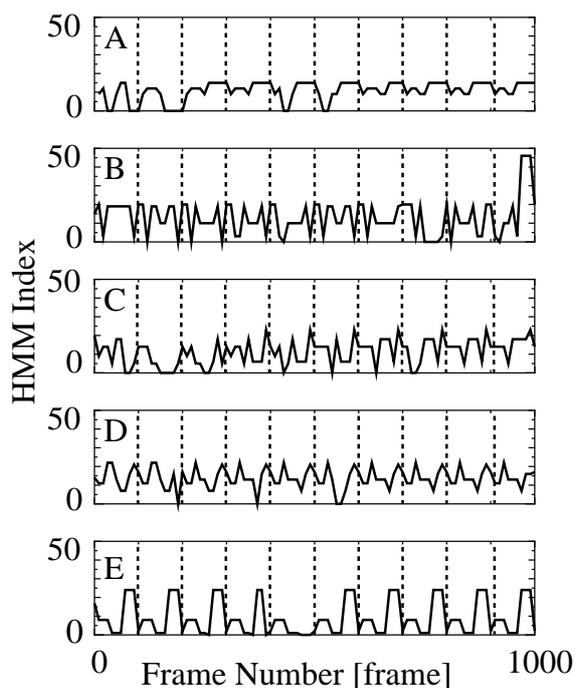


図 3: HMM index patterns correspondent to the observation interval in (A) Walk (B) Raise hand (C) Wave hand (D) Kick (E) Stand and the dot lines indicating the boundaries of motion patterns

誤差から判定した運動パターンの境界を表している。Fig.4 からはほぼ 100 フレーム毎に運動パターンの境界を判定しており、適当に運動時系列を分節化していることが確認できる。

先述の分節化される運動パターンは相関学習に利用された運動データである。そこで、学習に用いない未学習運動データを適用した場合の分節化の精度を検証した。Fig.5 は「足踏み」、「拳手」、「手を振る」、「キック」、「立ち上がる」の順に運動パターンを観察した時の分節化結果を表している。運動パターンが切り替わる付近で誤差が大きくなり始め、それに伴い運動パターンの境界判定を行っていることが分かる。この結果から、未学習運動パターンにおいても、提案した分節化手法の有効性が確認できる。

4. 分節化した運動パターンを用いた運動の認識・生成

前節までは、運動パターンの分節化手法および分節化結果について述べた。本節では、分節化して切り取られた運動時系列データが意味のあるまとまりであるかを運動の認識および生成の観点から検証した。

4.1 運動パターンの認識

同一の運動種類において、分節化された異なる運動パターンを同一の運動であると認識されるかどうかを検証する。具体的には「足踏み」に属する異なる運動パターンを提示した時に分節化して切り取られた運動パターンを共に「足踏み」と認識できるかどうかを確認する。この検証を行うために運動種と同数の 5 個の上位の HMM を用意する。各 HMM は、状態数 10、20 次元の確率密度関数から構成される Left-to-Right

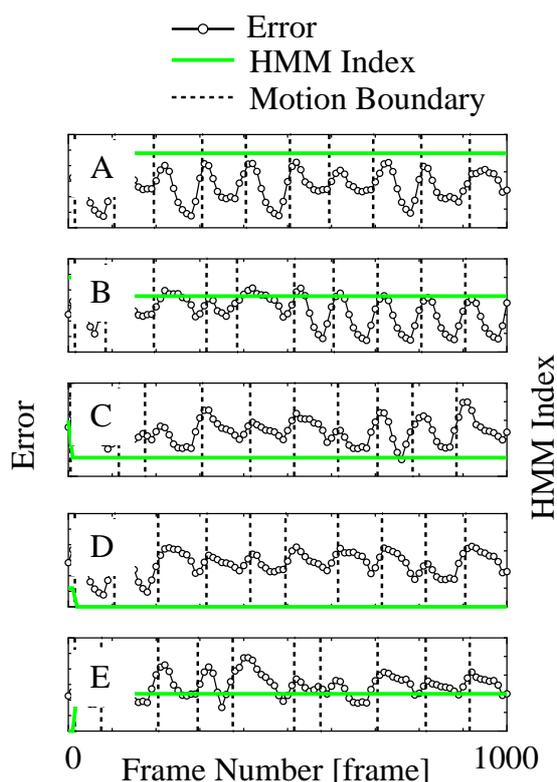


図 4: Segmentation simulation result. The round mark, the dot line and the solid line signify the error between the observed movement and predicted movement, the segment of the motion patterns and the HMM index corresponding to the recognition result of motion patterns. The given motion patterns are (A) Walk (B) Raise hand (C) Wave hand (D) Kick and (E) Stand beginning at the top.

型モデルである。各 HMM は、競合学習によって最適化される。最適化された HMM を用いた認識結果を Fig.4 の実線が示す「足踏み」、「拳手」、「手を振る」、「キック」、「立ち上がる」の運動パターンの認識結果が HMM のインデックス 4,3,1,0,2 であり、同種の運動パターンを同じ運動クラスと認識し、異なる運動パターンは異種と認識していることが確認できる。

4.2 運動パターンの生成

先述の運動のパターンの認識で用いた上位の HMM は、運動の認識・生成の双方向モデルとして提案し、各 HMM パラメータの集合を原始シンボルとして位置づけてきた。そこで、獲得した原始シンボルからの運動生成を行うことによって分節化された運動パターンについて検証した。Fig.6 は、学習に用いた実際の関節角の時系列データと原始シンボルから運動生成した場合の関節角時系列の比較図、運動生成した場合のモーションフィギュアモデルのアニメーション結果を表している。尚、Fig.6 の比較に用いた関節角は「足踏み」、「拳手」、「手を振る」、「キック」、「立ち上がる」の各々の運動動作で重要であると思われる左腰、右肩、左肩、右腰、右腰部の時系列データである。学習の実関節角データは、運動パターン毎にばらつきがあるが、原始シンボルから生成した運動パターンにおける関節角はそれらの平均値に近い値をとっていることがわかる。さらに、モーションフィギュアのアニメーション結果から、そ

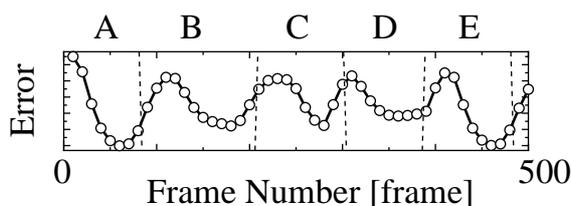


図 5: Segmentation simulation result. The round mark and the dot line signify the error between the observed movement and predicted movement and the segment of the motion patterns. The given motion patterns are (A) Walk (B) Raise hand (C) Wave hand (D) Kick and (E) Stand motion patterns that had been never learned in the learning phase

それぞれの運動パターンを主観的に「足踏み」「拳手」「手を振る」「キック」「立ち上がる」とラベル付けすることが可能であり、適切に原始シンボルから運動生成を行っていることが確認できる。

このように、分節化した運動パターンから競合学習を通して獲得した原始シンボルを用いた運動パターンの認識・生成結果からも適切に運動時系列データから運動パターンを分節化して切り取ることが可能であることが確認できる。

5. おわりに

本稿では、人間の連続の運動を記号列化に変換することによって情報を圧縮し、記号列の相関に着目した運動パターンの分節化手法を提案した。シミュレーション結果から、主観的な意味での運動の境界と提案した分節化手法を用いた場合の運動パターンの境界がほぼ一致することを示した。

さらに、分節化した運動パターンを用いることにより獲得した原始シンボルに基づいて行動の認識・生成を検証した結果、適切に機能することを確認し、本手法の有効性を示した。

今後は、本手法を実際のキャプチャシステム、ミメシスモデル等に統合することによってリアルタイム自立学習型ミメシスモデルへと発展させていく予定である。

なお、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、相関学習について博士課程の門根秀樹氏から御助言を頂き、大変感謝しております。

参考文献

- [Inamura04] 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, 2004
- [Donald91] M. Donald, "Origin of the Modern Mind", Harvard University Press, 1991
- [Gallese98] V. Gallese, A. Goldman "Mirror neuron and the simulation theory of mind-reading", Trends in Cognitive Science, Vol.2 No.12, pp.493-501, 1998.
- [Rizzolatti01] G. Rizzolatti, L. Fogassi, V. Gallese: "Neurophysiological mechanisms underlying the understand-

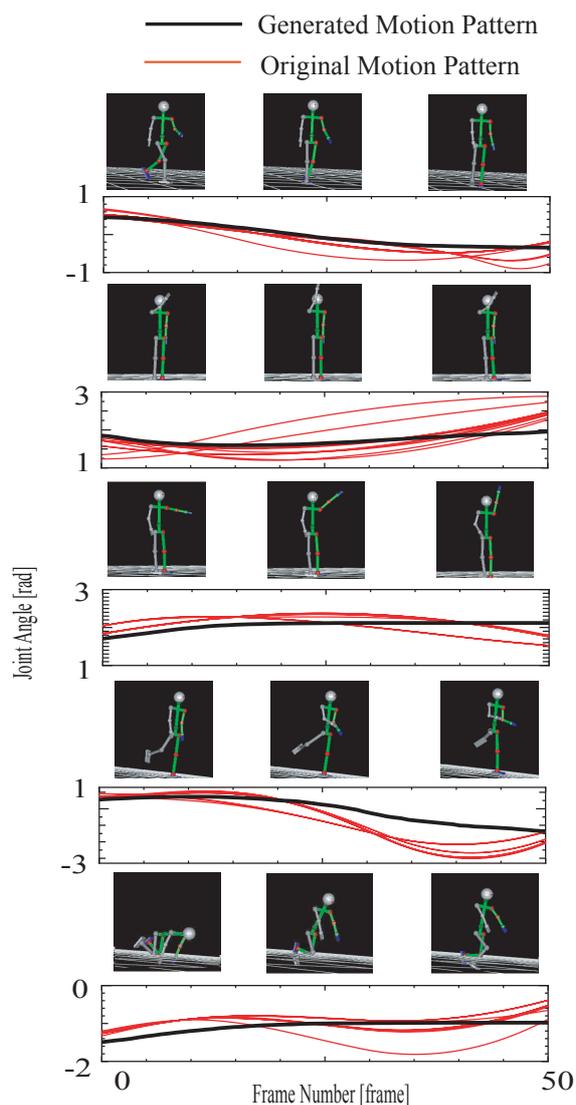


図 6: Comparison of generated motion patterns with the original motion patterns and the animation of the human figure driven by the HMM motion generators

ing and imitation of action", Nature Reviews, pp.661-670, 2001,

- [Redington97] M. Redington, N. Charter "Probabilistic and distributional approaches to language acquisition", Trends in Cognitive Sciences, Vol.1, No.7, pp.273-281, 1997
- [Christiansen99] M.H. Christiansen, S. Curtin "Transfer of learning: rule acquisition or statistical learning", Trends in Cognitive Sciences, Vol.3, No.8, pp.289-290, 1999
- [Yamane03] K. Yamane, 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