複数予測モデル遷移のN-gram統計に基づく非分節運動系列からの模倣学習手法

Imitation learning from unsegmented human motion based on N-gram statistics of linear prediction models

谷口 忠大[1,2] 岩橋 直人[2]
Tadahiro Taniguchi Naoto Iwahashi
立命館大学[1] 情報通信研究機構[2]
Ritsumeikan University NICT

Abstract: This paper presents an imitation learning method, which enables an autonomous robot to extract demonstrator's characteristic motions by observing unsegmented human motions. To imitate another's motions through unsegmented interaction, the robot has to find what he learns from the continuous time series. The learning architecture is developed mainly based on a switching autoregressive model (SARM), a keyword extraction method based on minimum description length principle, and singular vector decomposition to reduce dimensionality of high dimensional human bodily motion. In most previous research on methods of robotic imitation learning, target motions that were given to robots were segmented into several meaningful parts by the experimenters in advance. However, to imitate certain behaviors from the continuous motion of a person, the robot needs to find segments that should be learned. To achieve this goal, the learning architecture converts the continuous time series into a discrete time series of letters by using SARM after reducing its dimensionality by using SVD. After the conversion, the proposed method finds characteristic motions by utilizing n-gram statistics referring to description length. In our experiment, a demonstrator displayed several unsegmented motions to a robot. The results revealed that the framework enabled the robot to obtain several prepared characteristic human motions.

1 はじめに

ヒューマノイドロボットが開発され，自律的なロボットを人間社会へ普及させようとした際には，ロボット自身が人間社会に対して適応していく事が求められる．その適応対象には物理的な動作の獲得も含まれれば，私達人間が日常的に行っている記号的・社会的ルールもある．記号論の研究は私達の世界がいかに非自然性という意味での恣意性に満ちた存在であるかを明らかにしてきたが，ヒューマノイドロボットが人間機械系という人間を含んだ系を相手にする以上，その系への適応は不可避である．

自律ロボットの環境適応を議論する際に，主には強化学習と模倣学習が多く問題にされるが，人間が社会的なルールや他者の社会での行いを学ぶ際に主に用いられるのが模倣学習である．他者の観察に基づく模倣学習はオペラント条件付けに基づく強化学習と異なり，社会的な学習であるといえる．模倣学習は機械学習の枠組みでは，しばしば単純化され教師有り学習と同一視されるが，入力に対して出力を与えることで写像関係を学習するという教師有り学習には含まれない様々な問題を模倣学習は含んでいる．模倣学習が人間の知能を特徴づける高度な学習能力である証拠に，人間にとっては自然と達成可能な模倣学習もサルやチンパンジーにとっては非常に困難であることが知られている [1].

例えば，模倣者が被模倣者の行動を教師データとしようにも，何に着目して被模倣者の運動を自らの身体に関連づけるかという問題がその前段階として存在し Nehaniv らはこれを Correspondence problem と呼んで問題視している [2]．また，表面上は同じ行動であったとしても，それがどのような意図によって駆動されたものであるかによってしばしば模倣対象は変わってくる [3, 4]．例えば，赤いボールを持ち上げた動作を真似るには「赤いものを上げた」「右手を上げた」「ボールを上げた」など様々な解釈がありえる [4].「バイバイ」という動作にしても，アクション側のバイバイとリアクション側のバイバイでは，役割が異なるために意味が異なる [5]．さらには，連続的で非分節な時系列として提示される被模倣者の動作から，模倣者が「どの部分を真似るか？」という問題も，人間は易々と行なうにも拘わらず，ロボットが行おうとすると困難な問題の一つである．これらの問題を解決することは，社会適応可能な自律ロボットを実現する上でも，人間の知能を理解する上でも重要であると言える．

本論文では，上記問題の中で特に，連続的で非分節な動作時系列から模倣者が如何に特徴的な系列を抽出し模倣するかという点に焦点を当て，人間の動作模倣と音

声模倣を司る脳機能の計算過程についての仮説に基づき，複数予測モデル遷移の N-gram 統計を用いた発見的模倣学習手法を提案する．また，その有効性を実際の人間の上半身動作データを用いて検証する．

2 研究背景

人間の幼児は 1 歳半頃から親の動作を能動的に模倣し始め，バイバイやお辞儀などを表出するようになる．人間の幼児のように様々な動作を観察やインタラクションを通して自律ロボットが学習する事が出来れば，人間とのインタラクションを通じて自らが用いる「動作自体」を適応的に獲得するロボットを構築する事が出来，人間社会適応的なロボット構築の一歩となる．また，エンターテインメントロボットの視点からすれば，ユーザとの相互作用を通じてそのユーザ独自の学習経路を辿るペットロボットは人間ロボットインタラクションに新たな価値を生み出す可能性もある．本章では，ロボットが提示された動作系列を自動的に分節化し動作要素を学習する手法についての先行研究を概観して本稿のアプローチについて述べる．

2.1 動作の分節化と特徴的動作の抽出

非分節な動作系列から特徴的動作を抽出するという試みには二段階の問題がある．一つは，連続的な時系列を適切に区切るという問題であり，もう一つはまとまった特徴的動作を抽出するという問題である．多くの研究では複数の特徴的動作[1]が次々と切り替わる時系列データを準備し，それを適切に分割する事を問題としている [6]．しかし，人間の日常動作から特徴的な動作を切り出し，ロボットが模倣するという課題を考えると，人間の日常動作系列では特に意味のない動作が特徴的な動作の間を埋めており，背景的な動作列の中から有意味な特徴的動作を抽出するという問題も重要となる．

連続的な動作を時系列情報な局所的特徴からプリミティブに分解する研究はこれまで多くなされて来ている．古くには Rubin らが「視覚的分節境界」のプリミティブ集合を定義している [7]．ここでは，stop や impluse などの時系列の局所的な量的特徴からイベントを切り出し分節化する事が考えられている．また，Fod らは関節角の速度が 0 を跨いだ時点を分節境界としており [8]，櫻井らは特異スペクトル分解を用いて分節を導く方法を提案している [9]．一方で，境界を直接的に求めるのではなく，分節化された各要素が線形ダイナミクスを持つという視点から分節を求めるアプローチもなされている [10,11]．また，Marphy や Kawashima らは時系列データを線形ダイナミクスの確率的切り替わりとしてモデル化する手法を提案し，これにより時系列の分節化を行っている [12,13]．しかし，これらの手法で分節化された各動作単位は，多くの場合人間にとって意味ある動作というよりは，非常に物理的で断片的なプリミティブとなる．これは分節化はされているが人間観察者の意味解釈を導くようなものではない．Barbic らはこのような分節を「低次の動作要素 (low-level behavior components)」と呼んでおり，walk や sit down といったような人間の認識にとっての単位動作である「高次の動作要素 (high-level behavior components)」と区別している [6]．幼児が自律的な模倣学習により獲得しており，また我々がロボットに学習させたいのはこの高次の動作要素である．よって，観測される状態量の線形的・局所的特性に基づき時系列を区分するだけでは不十分である．これに対し，本稿では線形性の基準により断片化された低次の動作要素の連接情報（N-gram 統計量）を用いて結合する事で高次の動作要素に組織化するというアプローチをとる．

2.2 連続的動作からの自己組織化型学習

Ito や Tani らは適応的なバイアス項を有するリカレントニューラルネットワーク (RNNPB) を用い，一つのニューラルネットワーク内に予測誤差に基づき，複数の動作を自律的に獲得させる手法を提案している．この方式では一つの行動が RNN 内に分散的に表現されたアトラクタとして自己組織化される [14,15]．また，岡田や南野らは RNN や多項式で表現された非線形予測器を自己組織化マップ上に並べることで，予測誤差に基づきこれらを選択・学習させ，SOM 上に複数の動作を学習させる手法を提案している [16,17]．こらの手法により kick や squat といった高次の動作要素をロボットに獲得させる事が出来る．

このような問題を扱う際に，根本的に問題になることは「高次の動作要素が何によって規定されているか」という問題である．上記 RNN 等の非線形予測器に基づくアプローチでは，一つの動作単位を非線形予測器において安定的に形成されうる非線形アトラクタであるという考え方が根底にある．確かに，人間の動作の多くは周期性を有するリミットサイクルであるか，終端点を有する点アトラクタとして表現しうると考える事が出来，それを単位として時系列を区分する事が出来れば高次の動作要素をロボットに獲得させる事が出来る．例えば，Nakanishi らは運動のプリミティブとして明示的に周期運動や終端点を有する運動生成器のパラメータを座標変換することでヒューマノイドロボットに動作獲得を行わせている [18]．

しかし，これらの手法では時系列の分節化と学習が

[1] punch，kick，jump など

なされるが，特徴的動作とそれ以外の動作を区別し抽出する事はなされていない[2]．また，分節は基本に置く学習器の性質に強く依存し，どれだけの隠れ層を持ったRNN にするのか，何次の多項式にするのかにより分節の結果は異なってしまう．

このような競合学習器によって時系情報の分節化を行うアプローチに対し，提示された全時系列データの分布情報や部分時系列の特徴量に基づいて特徴的動作を抽出する手法も研究されている．特に，繰り返される動作を特徴的動作とみなす事により高次の動作要素を抽出する手法が考えられる．門根らは与えられた学習データから，繰り返し現れる運動パターンを自己相関に基づいて切り出す手法を提案している [19]．しかし，実験用に統制された動作ではなく人間の自然なモーションを対象にする場合にはそこには無視しがたい時間的な揺らぎが存在するが，直接的に時系列情報を時間の関数として比較する方法では時間方向の伸縮に対応できない．これに対し音声認識の世界では HMM（隠れマルコフモデル）を用いる事で直接的な特徴量空間での波形情報ではなく，離散的な隠れ状態の遷移により時系列情報の同一性を捉える事で時間方向の揺らぎに対応している．本稿の提案手法ではこれと同様の考え方により，隠れ状態の遷移において生まれる繰り返しパターンから繰り返し現れる動作パターンを抽出する．

2.3 音声言語模倣と動作模倣の類似性

本稿では音声言語模倣と動作模倣の類似性に着目し非分節運動系列からの模倣学習手法を構築する．主に音声認識に用いられてきた HMM が動作模倣に多く利用されて来ており，音声認識と動作模倣の数理的類似性が注目されている [20]．

Rizzolatti によるミラーニューロンの発見以降脳科学においても模倣学習を実現する脳内機構が注目されてきた [21]．ミラーニューロンは運動の生成と認識双方に関わっていることから，生成モデルとしてのパターン認識器である HMM との類似性が注目され，HMM を用いた模倣学習手法がミメシス理論として提案されている [20, 22]．また，音声知覚の研究においても，音声の知覚を運動系を用いて行っていると考える運動理論が再評価されてきており [23]，ミラーニューロンとの関わりが研究されている．また，ミラーニューロンは文法を司るブローカ領域の近傍に存在すると考えられているため，動作模倣と言語学習に何らかの共通の計算機構が用いられているのではないかと考える事も出来る [22]．これを本稿の主題に結びつけて考えると，幼児が連続的に発話される音声言語から単語を抽出し学習する脳内計算機構と共通の計算過程により，我々は非分節な動作時系列から高次の動作要素を抽出し学習する事が可能になるのではないかという仮説が導かれる．これは言い換えれば，音声言語や書き言葉に見られる「二重分節」の構造を人間の日常的な動作にも仮定するという事を意味する．

二重分節とは記号論において言語の特性としてのの指摘されている構造である [24]．二重分節とは音声が音素という形で分節化され[3]，それらが連なる事で単語という分節が形成されるという二段階の分節構造を音声言語がもつ事を指す．ここで重要なのは言葉は二段階目の単語という分節になって初めて「意味」を持ちうる点である．これは，動作の分節において，低次の動作要素が意味を持たず，もう少し大きな分節である高次の動作要素で初めて意味を持つ事と相似している．ここで，この二重分節の考え方に基づき音声言語模倣と動作模倣の類似性について以下の作業仮説を立てる．「音声言語において単語が形成されるのと同様に，動作模倣においても低次の動作要素が連なる事で高次の動作要素が構成され，その動作学習には音素列（文字列）から単語が抽出されるのと同様の計算機構が用いられるのではないか」という仮説である．しかし，どうやって低次の動作要素を連なりから高次の動作要素を抽出するかという問題がある．

近年，Web の急速な普及に伴い，自然言語処理の分野でも，辞書を使わず N-gram 統計量のみを用いた教師無し学習により，自動的に未知のテキスト上から単語・キーワードを抽出する手法が開発されている [25, 26]．N-gram 統計量とは文字の連接の頻度情報を指すが，近年の統計的な言語処理において，言語モデルを構築する際には標準的に N-gram 統計が用いられている．以前は自然言語の文法と言えばルールベースの手法が基本であったが，現在の音声認識では N-gram 統計により文法を確率的に学習するのが一般的である．故に計算論的な理解から言えば文法の学習とは N-gram 統計量の操作に他ならない．よって本稿では，非分節な動作系列から特徴的な動作を抽出する事にも，連続的な音声から単語を抽出し学習する事にも N-gram 統計量操作がブローカ野により行われ，ミラーニューロンによる認識と生成の双方向的な処理と協同的な学習により模倣学習が達成されているという想定に基づき学習機構を構築する．ここで，ミラーニューロンの処理に対応する学習機構が HMM のように離散的な隠れ状態を持ち認識と生成

[2] これらの実験では動作抽出も行っているように見える事があるが，そのように見えるのは学習データが特徴的な動作の遷移列で与えられている為に，分節化さえ行えば，分節化された後の動作は特徴的な動作であるという理由による．

[3] 書き言葉では文字がこの分節にあたる．

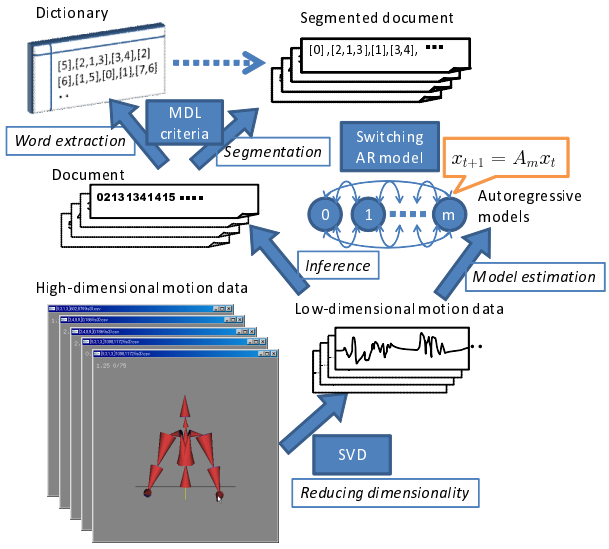

図 1: 提案する学習手法の全体像

を双方向的に行う確率モデルであり，ブローカ野の行う処理に対応するのが N-gram 統計量に基づきキーワード抽出を行う言語処理システムとなる．より具体的には本稿では，Switching AR model と最小記述長原理に基づく単語抽出手法によりそれらを表現する．

3 提案手法

提案手法の全体像を図 1 に示す．以下順次説明する．

3.1 特異値分解による低次元化

人間が熟練した動作や，繰り返し行なう動作には協応構造[4]が発達すると考えられている [27]．協応構造が組織化された動作では各関節の動作が強く関係し合い作動する為に，全体としての時系列情報が低次元化すると考えられる．一方で，時系列情報を低次元化する際には特異値分解 (SVD: Singular Value Decomposition) がよく用いられる．いま与えられた時刻 t に対して D 次元のベクトル $v_t (t \in \boldsymbol{T})$ の時系列データが存在するとするとし，これらを並べたものを行列 $A = (v_1^T, v_2^T, \cdots)^T$ とする．これを特異値分解すると

$$A = U \Sigma V^T \tag{1}$$

となる．ただし $\Sigma = \mathrm{diag}_{(T,D)}(\sigma_1, \sigma_2, \cdots, \sigma_D)$[5] このとき，$\Sigma$ を $\bar{\Sigma}_{(K)} = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_K)$ で置き換え，U,V^T をそれぞれはじめの K 列，K 行のみに制限した行列 $U^{(K)}, V^{(K)T}$ として作られる行列

$$\bar{A}_{(K)} = U^{(K)} \bar{\Sigma}_{(K)} V^{(K)T} \tag{2}$$

は $rank$ が K の行列による A のフロベニウスノルムの意味での最良近似となることが知られている．よって，時系列としては自乗誤差を最小とする低次元時系列での近似となっている．

本稿で対象とする人間の上半身動作は高次元であり，そのまま学習器に学習させるのは計算コスト的に非効率である．また，人間の動作の中でも注目されるのは，その一部の部分空間であると考えられるため，学習前に適切に時系列情報を低次元化する事で本質的でない次元の情報を削除する事が出来，次節以降の特徴抽出が効率化すると考えられる．

3.2 Switching AR model

Switching AR model (SARM) は多次元の時系列を複数の自己回帰 (AR) モデル (autoregressive model) の切り替わりによりモデル化する手法である [12]．本提案手法では SARM に基づいて被模倣者の動作系列を AR モデル（線形予測器）の切り替わりとしてモデル化する．各 AR モデルを表す隠れ状態はマルコフ過程に基づいて遷移すると仮定する．SARM は M 個の隠れ状態を持つ．我々は時刻 t における隠れ状態を s_t とする．隠れ状態 j は以下の AR モデルをもつ．

$$x_t = A_j x_{t-1} + v_t, \tag{3}$$

ここで x_t は時刻 t における時系列情報の状態ベクトル[6]，$v_t \sim N(0, Q_t)$ は分散共分散行列 Q_t の白色ガウスノイズを表す．まず，ロボットが時系列情報 $x_{1:T}$（$t = 1$ から $t = T$ までの時系列情報）を観測した際に，どの隠れ状態に相当する AR モデルから生成されたものとして認識するかについての計算機構について説明する．隠れ状態の推定には事後確率 $\Pr(s_t = j | x_{1:T})$ が計算される必要がある．隠れ状態遷移が遷移確率 Z のマルコフ過程に従うとすると，

$$\begin{aligned}
& \Pr(s_t = j | x_t, x_{1:t-1}) \\
= \; & \frac{1}{c} \Pr(x_t | s_t = j, x_{1:t-1}) \Pr(s_t = j | x_{1:t-1}) \\
= \; & \frac{1}{c} \Pr(x_t | s_t = j, x_{1:t-1}) \sum_i \Pr(s_t = j | s_{t-1} = i, x_{1:t-1}) \\
& \times \Pr(s_{t-1} = i | x_{1:t-1}) \\
= \; & \frac{1}{c} L_t(j) \sum_i Z(i,j) \Pr(s_{t-1} = i | x_{1:t-1}),
\end{aligned}$$

ここで c は正規化定数であり，$Z(i,j)$ が隠れ状態 i から j の遷移確率を表す．

$$L_t(j) = N(x_t; A_j x_{t-1}, Q_j) \tag{4}$$

[4] 文献によっては協調構造と呼ばれることもある．

[5] ここで $\mathrm{diag}_{(T,D)}(v)$ は $T \times D$ 行列で対角成分が v のものを指す．

[6] つまり，ロボットが獲得したユーザの関節角などの情報．

は j 番目の AR モデルによる時刻 t における予測誤差に基づく尤度を表わす．ここで N は中心を $A_j x_{t-1}$，そして分散共分散行列を Q_j とした多次元正規分布を表す．事後確率は backward の計算により以下のように算出される．

$$\Pr(s_t = j|x_{1:T}) = \sum_k \frac{\Pr(s_t = j|x_{1:t}) \Pr(s_{t+1} = k|s_t = j)}{\Pr(s_{t+1} = k|x_{1:t})} \times \Pr(s_{t+1} = k|x_{1:T}). \quad (5)$$

パラメータ A_j, Q_j も以下の EM アルゴリズムを用いる事で推定することができる．これは HMM のパラメータ再推定における Baum-Welch アルゴリズムに相当する [12].

$$\begin{aligned} A_j &= (\sum_l \sum_{t=2}^{T} W_t^j P_{t,t-1})(\sum_l \sum_{t=2}^{T} W_t^j P_{t,t-1})^{-1} \\ Q_j &= (\frac{1}{(\sum_l \sum_{t=2}^{T} W_t^j)} \times \\ &\quad (\sum_l \sum_{t=2}^{T} W_t^j P_t - A_j \sum_l \sum_{t=2}^{T} W_t^j P'_{t,t-1}) \end{aligned} \quad (6)$$

ここで $W_t^j \equiv \Pr(s_t = j|x_{1:T})$, $P_t \equiv x_t x_t^T$, かつ $P_t \equiv x_t x_{t-1}^T$． 元の論文 [12] では初期の分布 $P(s_1 = j) = \pi_j$ と $Z(i, j)$ も推定されているが，本稿では簡単の為，この推定は行わない．隠れ状態の遷移確率は自らへの滞在確率を ρ とし，全ての他の状態への遷移確率は一定のエルゴディック HMM とする．

次に, 計算された事後確率から最大のものを選ぶことで時系列情報を離散文字列へと変換する．これは音声認識における音素認識に相当する．

$$s_t^* = \underset{j}{\operatorname{argmax}} \Pr(s_t = j|x_{1:T}) = \underset{j}{\operatorname{argmax}} W_t^j, \quad (7)$$

ここで s_t^* は時刻 t における最も尤もらしい隠れ状態を指す．隣接する同じ隠れ状態を無視することで，隣接する文字が必ず異なる文字列へと圧縮する．以降，この文字列を *document* と呼ぶことにする．例えば，あるセッションが 10step の長さで，隠れ状態が $[1, 1, 1, 0, 0, 3, 3, 3, 3, 2]$ と遷移した場合，*document* は $[1, 0, 3, 2]$ となる．*document* は具体的なダイナミクスと時間方向の長さを無視した，抽象的なイベント列である．ここで連続的な時間の情報を捨てる事で情報の損失が起きているが，この抽象化により操作可能性が増すために N-gram 統計量の利用が可能になる点が重要である．また，SARM の遷移行列の滞在確率を ρ に固定し，他の全ての状態に当確率に遷移するエルゴディックモデルとしている為に，離散文字一文字に対する滞在時間の期待値は遷移確率の視点からすると一定となる為，具体的な滞在時間を捨象しても過剰な情報の損失は小さいと考えられる．

これにより，セッション数と同じだけの数の *document* を獲得する事ができる．

この後に，ロボットはこの document から分布情報に基づいて単語抽出を行う．一般的に，隠れ状態の列により形成される N-gram（たとえば, $[1, 3, 4, 5]$, や $[2, 1, 2, 1]$ など）はある動作系列を表象するが，そのほとんどの動作はユーザにとって意味がない．その中から繰り返し生起する意味ある N-gram を抽出する事が重要となる．

3.3 最小記述長原理に基づく単語抽出

Web の急速な拡大に伴い自然言語処理のニーズは高まっているが，その中でもブログをはじめとする文法が崩れることが多い話し言葉や辞書を適用できない新語が頻出する文章の解析が新たな課題となっており，教師無しの手法による単語抽出が研究されている [25, 26]．このような教師無し学習による単語抽出の手法は，隠れ状態の添字の並びのような未知の記号により形成される文章にも適用する事が可能である．

人間の動作の内，模倣対象となるような，高次の動作要素は線形予測モデルに対応する隠れ状態の連なりとして音声言語における単語のようになっていると考える．Taniguchi [28] らは同様の仮定に基づき，梅村らの未踏テキストからのキーワード抽出手法 [25] を用いる事でキーワードに相当する特徴的動作を抽出した．しかし，発見的手法に基づくため設計すべきパラメータが多い上，理論的なパラメータ設計手法がなく，抽出結果がセッションの区切りに強く依存するなどの問題点があった．これに対し，本稿では，セッションの区切りに依存せず，情報理論的に根拠の強い単語抽出手法として，最小記述長原理に基づく方法を採用する．松原らは同様の手法を用い日本語話し言葉の単語分割を行なっている [26].

今，最少記述長に基づく単語抽出では，文章を二段階に符号化する事を考える．二段階符号化とは，幾つもの単語を含んだ辞書を用いて対象の文書を符号化し，さらにその辞書を符号化する事を指す．ここで，分かち書きされた文章 d_i は辞書 *dict.* に含まれる単語群 $\{w_j\}$ の連なりで記述されているとし，単語 w_j は文字の連なり $a_1^j a_2^j \cdots a_{m_j}^j$ で表わされているとする．このとき文章 d_i を全て連結した全 *document*，すなわちコーパス

$\boldsymbol{d} = \{d_1 d_2 \cdots d_N\}$ の記述長は

$$L(\boldsymbol{d}) = -\sum_{w_j \in dict.} \#(w_j|\boldsymbol{d}) \log(p(w_j|\boldsymbol{d})) \quad (8)$$

$$p(w_j|\boldsymbol{d}) = \#(w_j|\boldsymbol{d}) / \sum_k \#(w_k|\boldsymbol{d}) \quad (9)$$

となる．ここで $\#(w_j|\boldsymbol{d})$ は分かち書きされたコーパス $\boldsymbol{d}$ 内に単語 w_j が出てくる回数である．例えば $d = ([ab][ab][a][ab][ca][b])$ の場合 $\#([ab]|d) = 3$ となる．また，単語を記録する辞書 $dict. = \{w_i\}$ の符号長 $L(dict.)$ は，用いた全ての単語を繋げて作った分かち書きされていない[7]文章についての符号長で与えられるとする．

$$L(dict.) = -\sum_j \#(a_j|dict.) \log(p(a_j|dict.)) \quad (10)$$

$$p(a_j|dict.) = \#(a_j|dict.) / \sum_k \#(a_k|dict.) \quad (11)$$

これらの和 $L^{double}(\boldsymbol{d}) = L(\boldsymbol{d}) + L(dict.)$ が単語抽出を行なった後の全体としての二段階符号化による符号長となる．これを最小化するように辞書の生成を行なう．

3.3.1 初期辞書の作成

辞書 $dict.$ が与えられた時に，コーパス符号長 $L(\boldsymbol{d})$ を最小化させるためには，単語 w_j 一つあたり符号長を

$$Score = -\log(p(w_j|\boldsymbol{d})) \quad (12)$$

としてこの総和が最小となるように分節化する必要がある．$Score$ が決まったときには，この分節化は動的計画法 (Viterbi サーチ) を用いる事で計算量を抑えて実行することが出来る [25]．よって，分節化は先に仮に $Score$ を決め，それに基づき分節化を行い，その結果から $Score$ を上式に従い更新するという手順をとる．必ずしも最適な分節化が出来るとは限らないが，十分良い解が得られると考える．

しかし，$L(\boldsymbol{d})$ を近似的にであれ最適化したとしても，辞書 $dict.$ が変化すると分節化の最適な経路自体が変化するために，$L^{double}(\boldsymbol{d})$ を一度に最適化する事は困難である．また，$dict.$ の候補としてはコーパスに含まれる全ての N-gram が対象になり得る為に全探索を行なうには膨大な計算量がかかる．そこで，分節化前の N-gram の頻度情報に基づき，近似的に初期 $Score$ を求め，全ての N-gram を辞書に登録した後に，初期の分節を決定する．なお，辞書に含まれない分節については十分に大きな負の値を $Score$ として与える．初期 $Score^{initial}$ は以下で定める．

$$Score^{initial} = -\log(\#(w_j|\boldsymbol{d}) / \sum_i \#(a_i|\boldsymbol{d}/f_0)) \quad (13)$$

[7]一文字単位で分かち書きされていると見なしても同等である．

ここで，f_0 は期待される単語の平均長さを設定する．本研究では $f_0 = 2.0$ と設定する．この $Score$ に基づき対象コーパスを分節化した後に，文中に現れた単語を数え上げその $Score$ を改めて式 (12) に従い計算し辞書を作成する．また，無駄な計算量の削減の為に初期辞書の単語に N-gram 出現頻度が F_{min} 以上のもののみ登録する．本研究では $F_{min} = 3$ としている[8]．

3.3.2 辞書の逐次的探索

次に，作成された辞書から順に一単語削除した後に，一単語削除された辞書を使って対象の文章を分節化し $L^{double}(\boldsymbol{d})$ を再計算する．記述長が減少した場合にはその削除を採用する事により，無駄な単語を辞書から減らす．記述長が減少しなかった場合には消去せずに次の単語の検討に移る．この操作を消すべき単語が無くなるまで繰り返す．これにより，最適性の保証はないが記述長を減少させる妥当な辞書を作成することが出来ると考えられる．作成された辞書から自明である 1-gram(一文字) の単語を除いたものを高次の動作要素に相当する単語であるとする．

獲得された単語はそれに含まれる文字に相当する AR モデルを順次再生することで実際のロボットの動作系列を生成できる．このとき生成される状態量の初期位置を決める必要がある．本研究では初期位置には学習データにおいてその単語に相当すると認識された時系列の初期座標の平均値を用いた．

4 実験

提案手法の妥当性を検証するために実験を行なった．人間の自然な動作から仮想的なロボットに動作抽出を行わせるとともに，被験者実験を行いロボットによる動作抽出の結果と人間による動作抽出の結果を比較した．

4.1 実験条件

模倣を行なう対象を人間の上半身動作とし，機械式のモーションキャプチャである Meta motion 社製の Gypsy 5 Torso を用い人間の上半身の各関節を取得した (図 2)．自由度は頭，首，両肩，両鎖骨，腰，両肘，両手首，胸，各 3 次元（オイラー角）の 36 次元である．本研究と類似した手法を用いている Taniguchi [28] では 3 次元の時系列を扱っており，それに比べると格段に大きい次元を扱っている．取得したデータは図 3 に示すように簡単な

[8] f_0 の設計指針としては最終的に獲得される辞書に含まれる単語の想定平均長が適切な初期値と考えられる．本実験では隠れ状態が 12 程度であるので 2〜3 程度が適当と考えられ，実際その周辺の初期値で良い探索解を得られた．また，f_0 については探索初期値を決めるだけので十分な探索時間を掛けられる場合は初期値を網羅的に設定し大域的な最適解を探索すればよい．F_{min} については 1 にすると出現する全ての N-gram が対象となり計算量が膨大となるので，通常 2, 3 程度にするのが効率的である．しかし，文書数が膨大な場合はより大きくしても構わない．

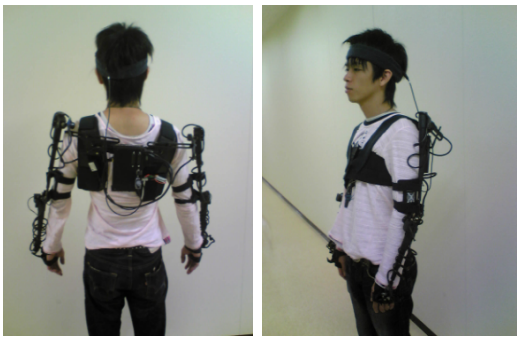

図 2: モーションキャプチャ Gyspsy 5 Torso の概観

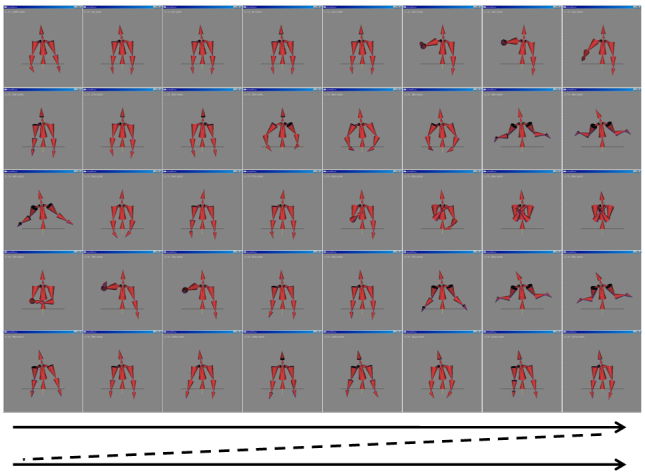

図 3: 被験者によって提示された分節化されていない時系列の例

3D 描画で再生可能である．

1 セッション 20[s] として，5 セッションの動作系列を記録した．フレームレートは 60[Hz] である．1 セッションは 1200[$step$] で構成される．被験者は成人男性である．記録している間，被験者には「Hi(オッス)」と「Shurug(肩すぼめ)」の主に二つの動作を提示するように求めた．ただし，他の動作も自然に挿入する事が許されており，実際に両腕をブラブラさせる動きや，「腰に手をやる」仕草などが上記二つの動作以外にも提示された．図 3 に一例として 3 セッション目の動作系列を 0.5[s] 毎にサンプリングしたものを示す．また，モーションキャプチャで取得されたデータを解析する前に分散一定の白色ノイズを付加した．ノイズの標準偏差は 0.3 とした．

4.2 実験結果

本実験では 36 次元の時系列情報を 3.1 の手法に基づいて 4 次元まで低次元化した．これに定数項 1 を付加し 5 次元のベクトルを SARM の解析対象とした．また，

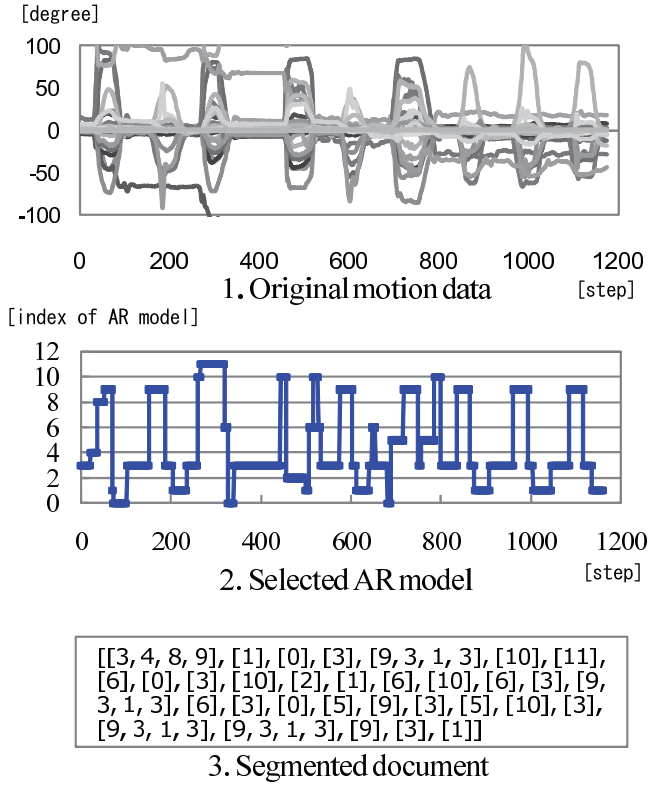

図 4: 多次元時系列情報から予測モデルの遷移による N-gram 系列への変換

特異値分解で低次元化すると，低次元化された時系列情報がなめらかで無くなる事があるので，指数平滑化係数 $\lambda = 0.8$ で平滑化を行なった[9]．

これの時系列情報を SARM を用いて解析し，AR モデルのパラメータとその選択確率を各時刻で求めた．ここで AR モデルの数は 12 とした[10]．またパラメータ推定においては，初期値を時系列データ全体を k-means 法でクラスタリングし，その中心ベクトルを指す一定関数を各 AR モデルの初期値とした[11]．EM アルゴリズムは 10 回繰り返し，単語抽出において N-gram は 10-gram までを考慮した．第 1 回目のセッション情報について元の時系列情報から SARM の隠れ状態系列に変換され，そこから分節化された文字列に変換される様子を図 4 に示す．

隠れ状態列に対して $N \leqq 10$ の N-gram 全てを考慮

[9]時系列 x_t を指数平滑化した結果 z_t は $z_t = \lambda z_{t-1} + (1-\lambda)x_t$ として求められる．λ は大きくしすぎると動作がなまる．サンプリング時間を $\triangle t$ とすると時定数 τ により $\lambda = \exp(-\triangle/\tau)$ とできるが，$\tau < 1$[s] 程度には納める必要があると考えられる．

[10]AR モデル数を単語抽出を含めた確率モデルにおいてロボット自身が適切に定めるモデル選択手法の適用は今後の課題であるが，本研究では 8〜12 程度で SARM の尤度の増加がほぼ止まった事を理由に AR モデル数を定めた．

[11]定数項にかかる行列の最終列に列ベクトルでクラスタ中心ベクトルが割り振られ，他の成分は 0 という行列

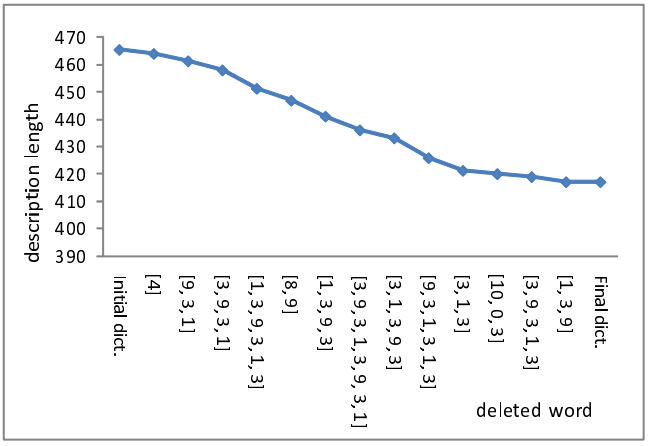

図 5: 初期の辞書から削られた単語と，それらが削られた時点での記述長

すると 1332 単語となったが，初期スコアに基づき対象の文書群を分節化し，そこに現れたものだけを初期辞書に登録すると合計 25 単語，1-gram を除くと 14 の単語が登録された．その後，一つの単語を削っては二段階符号化長が減少するかを逐次的に調べる逐次探索により順に 12 個の単語が削除されていった．削除した単語と記述長の変化を図 5 に示す．最終的には 1-gram を除くと，$[3,4,8,9]$ と $[9,3,1,3]$ という二つの単語が抽出された．

5 セッションに対応する分節化された 5 つの文書中には $[3,4,8,9]$ は 3 回，$[9,3,1,3]$ は 12 回現れていた．これらの部分に相当する時系列の内の一つを再生した結果を図 6 と図 7 に示す．これより，$[3,4,8,9]$ が “Shrug” の部分に，$[9,3,1,3]$ が “Hi” の部分に相当している事がわかる．

4.3 抽出の評価

実際にこれらの抽出結果が人間が行う認識とどれだけ符号しているかを評価する．人間の動作からロボットに教師無しで分節化と行動学習を行わせる先行研究の多くでは，ロボットがいくつかの動作を獲得した事を示すのみで，実際にロボットが行った分節化と人間が行った分節化がどれだけ近いかという評価を行っていない．ここでは，５名の被験者にロボットが学習に用いたモーションキャプチャーで採取した５セッションの動作列を図 3 のようなアニメーションにより提示し，そこから “Shrug” 及び “Hi” の二つの全動作についてその動作の開始時間と終了時間を各被験者の感覚に基づいて記入させた．つまり，ロボットが行ったのとほぼ同様のタスクを行わせた．５名の被験者が同じ５セッションの動作系列を分節化する事で，人間の平均的な分節化とそのばらつきを得るのが目的である．ここで実映像ではなくアニメーションに変換して提示させた理由であるが，この被験者実験はロボットの分節化能力を人間のそれと比較するためであり，ロボットが関節角しか観測できていないという条件を出来る限り被験者とそろえる為に実映像でなくアニメーションに変換して提示した．人間の表情など付加的な情報を観察出来た場合，被験者はこれに基づき分節化を行う可能性がある[12]．

被験者は必要に応じて一時停止，フレーム単位での送り，戻しを行う事が出来る．また，動作には明確な境界は存在しないため被験者毎に揺らぎが存在する．標準的な観察者の視点では 5 セッション全体で Shrug が 11 回，Hi が 13 回観測されたが，平均するとそれぞれ一回の長さは Shrug が 1.52[s]，Hi が 1.39[s] であった．被験者 5 人間での揺らぎは標準偏差で Shrug で 0.42[s]，Hi で 0.58[s] と，人間の間でもかなり大きな揺らぎが見られた．そこで，5 人中 3 人以上がその動作の提示中であると判断した区間を真のモーション提示区間であるとし，これとロボットによる認識の比較を二段階に分けて行い，提案手法の評価を行った．

4.3.1 認識率の評価

上記より，Hi の提示区間に対して $[9,3,1,3]$ が重複部分を持っている，もしくは，Shrug の提示区間に対して $[3,4,8,9]$ が重複部分を持っていれば認識出来たとし，その評価のために再現率 (recall) と適合率 (precision) を計算した．全 Hi,Shrug の内どれだけを単語として発見できたかを表わす再現率 (recall) は 54.2%，出力した全単語の中でどれだけが正しい Hi,Shrug であったかを表わす適合率 (precision) は 93.9%となった．また，総合的な評価指標として再現率と適合率の調和平均で計算される F 値は 68.4%となった．対応箇所を余すところ無く見つける検索課題では再現率が重要であるが，模倣学習の元となる教師事例を見つける意味では適合率が重要である．よって，本手法は非運動系列からの模倣学習手法としては適当な動作抽出を行っていると言える．

4.3.2 抽出区間の評価

次に，認識できた動作に対して，その抽出区間がどれだけ人間の抽出区間と近いかを評価した．評価のために提案手法の行う区間抽出を，モーション提示区間に含まれるフレームを検索する課題と見なして F 値を計算し

[12] これにより，人間が人間の動作を摸倣する際に用いる情報が過剰に欠落するのではないかという危惧もあるが，バイオロジカルモーションについての一連の研究でも知られているように，人間は角関節につけられた光点の運動を観察することだけからでも性別や持っている物の重などを推定する事が出来ると考えられている [29].

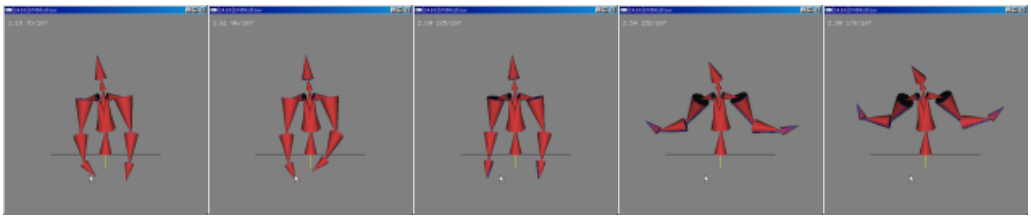

図 6: [3,4,8,9] として分節化された動作 (Shrug モーションに相当)

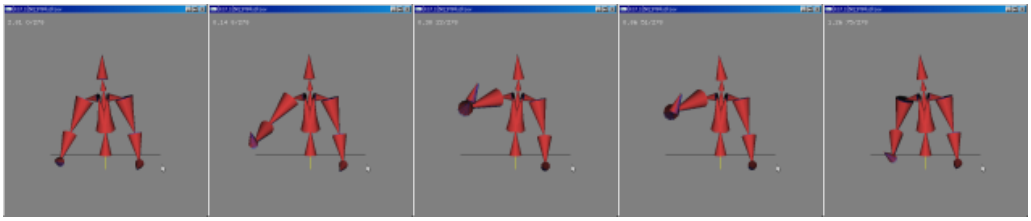

図 7: [9,3,1,3] として分節化された動作 (Hi モーションに相当)

た．ある動作に対して F 値は

$$F = \frac{2 \times \text{正しく抽出された区間長} \times 100[\%]}{(\text{モーション提示区間長} + \text{抽出された区間長})} \tag{14}$$

で定められる．提案手法の F 値は 54.7% であった．一方で，人間の行う認識にもばらつきがあるので，5 人の被験者の全組み合わせに対して，一方の認識結果が一方の認識結果の予測結果であると見なして F 値を計算し，全組み合わせについての平均をとったところ 72.9%であった．一方，被験者間で最も認識結果の悪かった組み合わせでの F 値は 56.6%であったため，提案手法は最も動作抽出の感覚に開きのある人間間のずれ程度の性能で動作抽出を行う事が出来たと言える (図 8)．

これらより，提案手法は関節角度の時系列情報のみしか扱わないにも関わらず，教師無しで人間の繰り返される特徴的な動作をある程度抽出する事が出来たと考えられる．

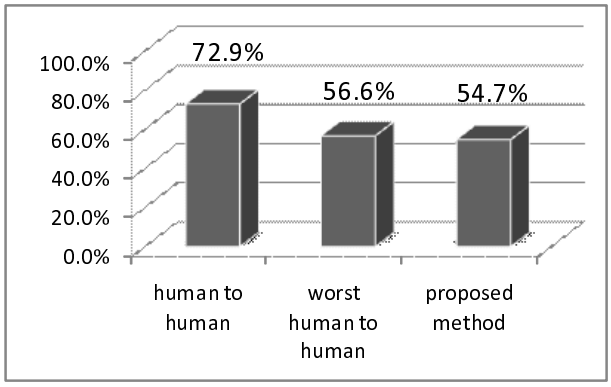

図 8: 区間の抽出についての評価．human to human は被験者の全組み合わせに対する F 値の平均，worst human to human はその中で最も F 値の低かった被験者間での F 値，proposed method は被験者の平均的な認識に対する提案手法の F 値．

5 まとめと議論

本研究では人間の日常的な動作からロボットが模倣学習を通じて自律的に様々な動作を学習するための手法について提案した．提案手法では，人間の特徴的動作が低次元化されているという特徴と，ある程度決まった系列が繰り返し生起するという特徴を用いて動作抽出を行なった．当手法で用いた単語抽出手法は最小記述長原理という明確な基準に基づいている為に閾値のような特に恣意的なパラメータ設定を要しない．先行研究 [28] では単語抽出だけで 4 つのパラメータ設計を要したが，本手法ではほぼ探索の初期値を決定する f_0 のみである．しかし，この f_0 についても探索の初期値を決定するのみであるので，複数設定しその中で記述長が最小の解を得られるものにすればよい．他のパラメータ設計としては，SARM において AR モデルの数を選択設定する部分と，低次元化を行なう際に何次元にするかという点に恣意性があるが，これらについては，学習プロセス全体に対してモデル選択手法の適用を考える事で解決する必要があると考える．

模倣学習後の動作生成の段階から考えると，動作を構成する要素となる AR モデルは生成系としては，特に状態量を状態空間の一定の領域に安定的に出力するわ

けではなく，しばしば関節角を過剰に変化させてしまう事が見受けられた．各動作要素をガウス分布で表現する隠れマルコフモデル [4, 20] や，安定性についての制約を加えた線形システムなど，他の手法を用いる事も検討に値する．

本稿では提案手法により，特徴的動作が低次元化された時系列情報から複数の予測器により記号化され，そのN-gram 統計量を用いて単語として抽出出来る事が示された．しかし，これらの単語の具体的なラベルは EM アルゴリズムの初期値によっても変化する．このため，本論文では実験は一事例に留めたが，複数回の実験を行った際に何を不変として比較するかも重要な課題である．本研究では被験者に実際に動作の区間抽出を行わせ，真値とすることにより提案手法の妥当性を評価したが，この評価手法自体についてもより検討される必要がある．

当手法はヒューマノイドロボットが人間動作の観察を通じて，人間の指示を受けることなく自律的に人間の動作を模倣するというプロセスを実現する為に提案したが，技術的な応用としてはそれに留まらない．ヒューマノイドロボットのみならず VR 空間上のアバターなどにも適用可能であるし，自動的な人間動作の解析という視点からの応用も考えられる．人間の動作を高次の動作要素に分解し解析することはサーブリッグ分析など IE(Industrial Engineering) における動作研究で人手によって為されており，このような動作分析を自動化する為の解析手法としても応用の可能性が考えられる．

また，本研究では関節角という単一の時系列からの特徴的動作抽出を行なったが，人間の動作を振り返ると人間の身体座標の系列情報だけで毎回変わらず意味を成す動作はバイバイやお辞儀などのジェスチャ動作だけであり，物体操作などの環境との相互作用を含んだ動作の模倣では操作対象物体との関係性やシチュエーションとの関係性を含んだ議論が必要である．これにどう展開するかも今後の課題である．

謝辞

本研究は科学研究費補助金 若手研究（スタートアップ）20800060「非分節な人間機械相互作用を通じた自己組織化型模倣学習機構の構築」，科学研究費補助金 学術創成「記号過程を内包した動的適応システムの設計論」19GS0208 及び国立情報学研究所共同研究助成「能動的ハンドインタラクションによる実世界言語コミュニケーションの学習に関する研究」の 一部支援を受けた．

参考文献

[1] 明和政子. 心が芽ばえるとき コミュニケーションの誕生と進化 (叢書コムニス). NTT 出版, 2006.

[2] C.L. Nehaniv and K. Dautenhahn. The correspondence problem. *Imitation in Animals and Artifacts*, pp. 41–61, 2002.

[3] A. Billard, Y. Epars, S. Calinon, S. Schaal, and G. Cheng. Discovering optimal imitation strategies. *Robotics and Autonomous Systems*, Vol. 47, No. 2-3, pp. 69–77, 2004.

[4] K. Sugiura and N. Iwahashi. Learning object-manipulation verbs for human-robot communication. In *Workshop on Multimodal Interfaces in Semantic Interaction at the International Conference on Multimodal Interfaces*, pp. 32–38, 2007.

[5] 谷口忠大, 岩橋直人, 中西弘門, 西川郁子. ヒューマン・ロボットインタラクションを通した役割反転模倣に基づく実時間応答戦略獲得. 第 23 回人工知能学会全国大会 in CD-ROM, 2009.

[6] J. Barbič, A. Safonova, J.Y. Pan, C. Faloutsos, J.K. Hodgins, and N.S. Pollard. Segmenting motion capture data into distinct behaviors. In *Proceedings of Graphics Interface 2004*, pp. 185–194, 2004.

[7] J.M. Rubin and Richards. Boundaries of visual motion. *A.I. Memo 835, Massachusetts Institute of Technology, Artificial Intelligence Laboratory*, 1985.

[8] A. Fod, M.J. Matarić, and O.C. Jenkins. Automated derivation of primitives for movement classification. *Autonomous robots*, Vol. 12, No. 1, pp. 39–54, 2002.

[9] 櫻井啓智, 中西弘明, 堀口由貴男, 椹木哲夫. 特異スペクトル変換による行動の分節化. 第 24 回ファジィシステムシンポジウム講演論文集, pp. 291–296, 2008.

[10] Y. Li, T. Wang, and H.Y. Shum. Motion texture: a two-level statistical model for character motion synthesis. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pp. 465–472, 2002.

[11] D.M. Wolpert, K. Doya, and M. Kawato. A unifying comuputational framework for motor control and social interaction. *Phil Trans R Soc Lond B*, Vol. 358, pp. 593–602, 2003.

[12] K.P. Murphy. Switching Kalman filters. *Technical reports, DEC/ Compaq Cambridge Research Labs*, 1998.

[13] H. Kawashima and T. Matsuyama. Multiphase learning for an interval-based hybrid dynamical system. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, Vol. E88-A, No. 11, pp. 3022–3035, 2005.

[14] M. Ito, K. Noda, Y. Hoshino, and J. Tani. Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model. *Neural Networks*, Vol. 19, No. 3, pp. 323–337, 2006.

[15] J. Tani, M. Ito, and Y. Sugita. Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using RNNPB. *Neural Networks*, Vol. 17, No. 8-9, pp. 1273–1289, 2004.

[16] 藤田雅博, 下村秀樹（編）. 発達する知能 - 知能を形作る相互作用 (インテリジェンス・ダイナミクス). シュプリンガージャパン, 2008.

[17] M. Okada, D. Nakamura, and Y. Nakamura. Self-organizing Symbol Acquisition and Motion Generation based on Dynamics-based Information Processing System. In *Proc. of the second International Workshop on Man-Machine Symbiotic-Systems*, pp. 219–229, 2004.

[18] A. J. Ijspeert, J. Nakanishi, and S. Schaal. Learning attractor landscapes for learning motor primitives. In *Neural Information Processing Systems 15*, pp. 1547–1554, 2003.

[19] 門根秀樹, 中村仁彦. パターンの相関と連想記憶に基づく運動パターンの分節化・記憶・抽象化. 第 24 回日本ロボット学会学術講演会 1D32, 2006.

[20] 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.

[21] L. Fadiga, L. Fogassi, G. Pavesi, and G. Rizzolatti. Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, Vol. 73, No. 6, pp. 2608–2611, 1995.

[22] 中村仁彦. 岩波講座 物理の世界 （物理と数理１）ロボットの脳を創る　脳科学から知能の構成へ. 岩波書店, 2003.

[23] 柏野牧夫. 音声知覚の運動理論をめぐって (＜小特集＞人間の音声情報処理機構の解明に向けて). 日本音響学会誌, Vol. 62, No. 5, pp. 391–396, 2006.

[24] 池上嘉彦. 自然と文化の記号論 (放送大学教材). 放送大学教育振興会, 2002.

[25] 梅村恭司.　未踏テキスト情報中のキーワードの抽出システム開発.　Technical report, http://www.ipa.go.jp/archive/NBP/12nendo/12mito/mdata/10-36h/10-36h.pdf, 2000.

[26] 松原勇介, 秋葉友良, 辻井潤一. 最小記述長原理に基づいた日本語話し言葉の単語分割. 言語処理学会第 13 回年次大会, 2007.

[27] 三嶋博之. エコロジカル・マインド - 知性と環境をつなぐ心理学 (NHK ブックス). 日本放送出版協会, 2000.

[28] T. Taniguchi, N. Iwahashi, K. Sugiura, and T. Sawaragi. Constructive approach to role-reversal imitation through unsegmented interactions. *Journal ref: Journal of Robotics and Mechatronics*, Vol. 20, No. 4, pp. 567–577, 2008.

[29] 佐々木正人. アフォーダンス-新しい認知の理論 (岩波科学ライブラリー (12)). 岩波書店, 1994.

谷口忠大（正会員）

1978 年 6 月 24 日生．2006 年京都大学工学研究科博士課程修了．2005 年より日本学術振興会特別研究員(DC2)，2006 年より同(PD)．2007 年より京都大学情報学研究科にて(PD)再任．2008 年より現在，個体と組織に於ける記号過程の計算論的な理解や共生社会に向けた知能情報学技術の応用研究についての研究に従事．京都大学博士（工学)．計測自動制御学会学術奨励賞，システム制御情報学会学会賞奨励賞，論文賞，砂原賞など受賞．計測自動制御学会，日本人工知能学会，システム制御情報学会，日本神経回路学会などの会員．

岩橋直人（非会員）

1962 年 1 月 13 日生. 1985 慶応義塾大学理工学部計測工学科卒. 1985～1998 ソニー. 1990～1993 ATR 自動翻訳研究所に出向．1998～2004 ソニーコンピュータサイエンス研究所．現在，情報通信研究機構，ATR メディア情報科学研究所，神戸大学大学院．博士（工学)．人間とロボットのインタラクション，ロボットによる言語獲得の研究に従事．人工知能学会，電子情報通信学会，日本認知科学会各会員．