

## RNN 予測器を用いた人物行動の自動分節化

牛 童<sup>†</sup> 岩井 儀雄<sup>†</sup> 谷内田正彦<sup>†</sup>

<sup>†</sup> 大阪大学大学院基礎工学研究科  
〒560-8531 大阪府豊中市待兼山町 1-3  
E-mail: niu@yachi-lab.sys.es.osaka-u.ac.jp

あらまし 近年、犯罪の多様化により、セキュリティの強化に対する意識が高まっている。そのため、本研究では未知行動を例外行動として自動検出するシステムを開発することを目標としている。このようなシステムでは新規行動を追加学習するときモデル全体を再学習する必要がある。本研究では、リカレントニューラルネットワークを用い、人物行動を予測可能性に基づいて自動分節化することで、モデル更新すべきサブモデルのみを再学習する手法を構築することを目的としている。本稿では、その予備実験として行った、モデルの予測能力評価について報告する。

キーワード RNN, 予測器, 自動分節化

## Automatic Segmentation of Human Actions based on Predictor Model using RNN

Tong NIU<sup>†</sup>, Yoshio IWAI<sup>†</sup>, and Masahiko YACHIDA<sup>†</sup>

<sup>†</sup> Graduate School of Engineering Science, Osaka University  
1-3 Machikaneyama-cho, Toyonaka, Osaka, 560-8531 Japan  
E-mail: niu@yachi-lab.sys.es.osaka-u.ac.jp

**Abstract** Recently, diversity of crimes has become a major social problem, calling for reinforcements in security systems. The objective of the author's research is to create a system that automatically detects unknown behaviors as exceptions. Such systems require retraining of the whole system when supplementing a new behavior into the model. In this research, the authors aim to segment the behaviors, based on predictability of the recurrent neural network, for retraining only the sub-model for supplementation. In this paper, the authors present the results of the preliminary experiments evaluating the model's prediction capability.

**Key words** RNN, Predictor Model, Segmentation

### 1. はじめに

近年、犯罪の多様化が深刻な社会問題となり、セキュリティの確保と安全性がますます重要になっている。安全で安心な社会を実現するには、監視システムの強化が必要である。しかし、一般的な防犯カメラ監視システムでは、監視者がカメラの映像を監視し続ける必要があり、目視という単純作業による疲労や怠惰などから起こる人為的なミスや、コスト、効率の面で問題がある。そのため、人物の行動を認識し、未知の行動を例外行動として自動的に検出するようなセキュリティシステムが望まれる。このようなシステムによって、監視者の負担を軽減し、人為的なミスを避けることができ、コストの削減や異常事態の早期発見などにつながる事が期待できる。

このような理由から、例外行動の自動検出に関する研究が従来から盛んに行われている [2], [3]。このような自動検出システ

ムは学習部分と認識部分の二つから構成される。学習には、正常行動を学習し、行動モデルを構築する。一般的にまず、人物の行動軌跡を観測し、体の特徴点のデータを取得する。取得した行動データをパターン分類し、コンピュータシステムで学習することで行動モデルを構築する。認識時には、観測した行動に対して、行動モデルの適合度を計算し、適合度が低いモデルに不適合な行動を例外行動として検出する。

行動モデルを構築する際、従来は行動軌跡全体を学習したため、学習コストが大きくなるという問題点があった。これを解決するため、軌跡を部分軌跡で分割・類似行動を統合して学習することで、学習コストを減らすという手法 [4] が提案された。しかし、一般的に行動の分節箇所を手動で設定する 경우가多く、システム構築者にとって、手間のかかる作業である。自動的に行動素を抽出する従来研究 [5] もあるが、行動を分割する箇所は静止点で設定するものであった。しかし、人物の行動は静止

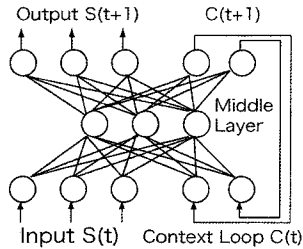


図1 RNNの構成

点で分節化できるものと、静止点ではなくても分節化できるものがある。本研究ではより効率的に分割を行うため、軌跡のモデル適合度をもとに分割を行う手法を開発することを目標としている。本稿では、その予備実験として行った、モデルの予測能力の評価について報告する。

## 2. 人物行動モデルの構築

行動の自動分節化に関する研究は、ロボティクスの分野でも行われており、リカレントニューラルネットワーク (Recurrent Neural Network) (以下 RNN と略記) によるロボットの部分行動分節化が提案されている [1]。この研究では RNN を局所的に配置した Mixture of Experts モデルが用いられている。本研究ではこのモデルに基づき、RNN の出力誤差をもとに人物行動パターンの自動分節化を行うシステムの構築を目標とする。以下に本研究で目標とするシステムの構築方法について説明する。

### 2.1 RNN 予測器

本研究では RNN を用いて、行動の学習と予測を行う。RNN(図 1) は階層型ニューラルネットワークに回帰構造を付け加えたものであり、現状態  $S(t)$  を入力することで次状態  $S(t+1)$  を出力する予測器として働く。また、コンテキスト層によって過去の履歴も考慮した学習が可能である。学習方法としては BPTT(Back Propagation Through Time) [7] を用いて、出力誤差が小さくなるように、各層の結合重みを修正していく。

### 2.2 Mixture of Experts によるモデル構築

Mixture of experts (以下 MoE と略記) は階層型のネットワークである。MoE は入力データ空間をいくつかの小領域に分割し、その分割された各領域に対してひとつのニューラルネットワークを割り当てることによって複雑な問題の解を求めため手法である。複雑な問題を分割して小領域に区切ることによって、1つの大きなニューラルネットワークを学習させるよりも効率的な学習をさせようというモデルである。本研究では MoE モデルを用いて、一連の行動系列のモデルを構築する。具体的に一連の行動系列をいくつかの部分行動に分割し、その分割された各領域に RNN を用いてモデル学習する。

人物の行動モデルを考える際、複雑な一連の行動も、単純な部分的パーツの組み合わせから構成されると考えられる。例えば、食べ物を口に入れるという行為は、箸を持つ、食べ物を掴む、食べ物を口に入れるといった三つのパーツに分解できると考えられる。これらのパーツが部分行動にあたる。多様で複雑

な行動は、手持ちのいくつかの部分行動をうまく組み合わせることにより、構成できると考えられ、これを行動的分節化という。行動を分節化し、類似行動を同一の部分行動として統合、それぞれを独立に別々の RNN で学習する。ここで、RNN で学習された部分行動のことをサブモデルと呼ぶ。行動をサブモデルに分割し、学習することで学習コストを削減することができる。例えば、図 2 では、サブモデルが 3 つあり、モデルの数は 2 である。サブモデル RNN1a で分節化を行い、RNN1a から RNN2a へ遷移するような行動がモデル 1 であり、RNN1a から RNN2b へ遷移するような行動はモデル 2 である。このように、モデル 1 とモデル 2 で共通のサブモデル RNN1a を用いることにより、学習コストを減らすことができる。

### 2.3 行動モデルの分節化と追加学習

本節では目標とする、RNN の予測可能性に基づく分節化、について説明する。行動モデルを構築する際、逐次的に新しい学習パターンを追加学習する必要がある。そこで、追加学習データに対して予測を行い、これまで構築した行動モデルにどこまで適応するかを測る。予測誤差が閾値を越えるところを検出し、そこで追加学習データを前半 (予測可能) と後半 (予測不可能) の二つに分ける。予測不可能部分を新しい RNN で学習し、予測可能部分はもとの RNN 学習器に入れ、再学習を行い、新しい行動モデルを作る。追加学習を含めた行動モデル構築の全体的な流れは図 3 に示す。処理の手順は主に以下の 4 ステップとなっている。

- (1) 学習：初期行動パターンを RNN で学習し、初期行動モデル  $M_n$  を構築する。
- (2) 予測：追加学習データ  $NewData$  に対して、予測を行い、分節点を検出する。
- (3) 分節：追加学習データ  $NewData$  を予測可能な  $NewData_{known}$  と予測不可能な  $NewData_{unknown}$  の二つに分ける。
- (4) 再学習：予測不可能の  $NewData_{unknown}$  に対して RNN で学習し、新しい行動モデル  $M_{n+1}$  を構築する。

RNN で学習し、新しい行動モデルを構築する自動分節化の例を図 4, 5, 6 に示す。これまで学習済みのモデル構造 (図 4) は、RNN1 から RNN2a, RNN2b に分節し、RNN2b から RNN3a, RNN3b を分節し、サブモデルが 5 つ、モデルが 3 つある。これに対して、新しい学習データが入るとする。学習済

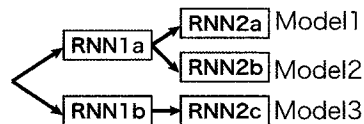


図2 サブモデルの遷移

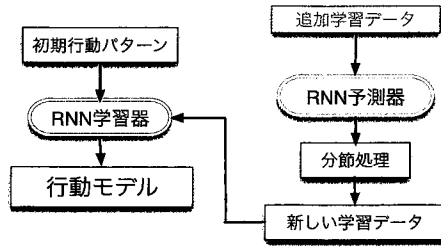


図3 行動モデル構築の処理流れ

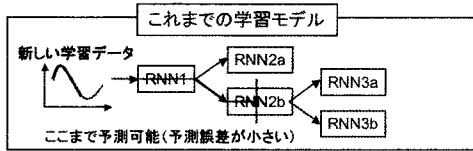


図4 学習済みのモデル構造

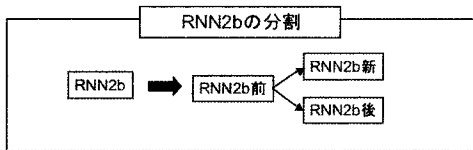


図5 モデルの分節化

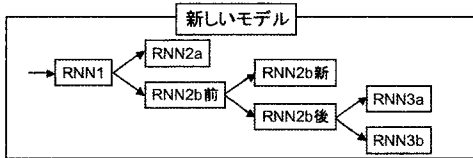
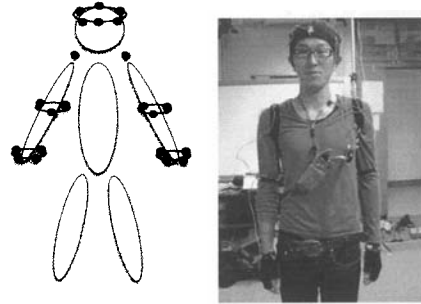


図6 新しい行動モデル

みのモデルで予測を行うと、サブモデル RNN2b の途中で適合不可になった、つまり予測値と実際の学習データの差が大きくなった、とする。この分節点で新しい学習データを予測可能な部分と予測不可能の部分に分ける。予測不可能の部分进行学习し、新しいサブモデル RNN2b 新を作製する。それと同時に学習済みの RNN2b を前と後ろに分節し (図5)、分節された RNN2b 前と RNN2b 後について再学習する。最後に新しいサブモデル RNN2b 新と RNN2b の後半部を RNN2b の前半部につける。すると新しい行動モデル (図6) はサブモデルが7つ、モデル4つとなる。このように分節された箇所のみ再学習自動分節をすることで、分節に関係ない RNN1, RNN3a と RNN3b を考慮しなくてもいいため、学習コストを削減した追加学習ができる。

### 3. 汎化性評価実験

より正確に分節化を行うためには、行動モデル適合性の評価は重要であり、同じ行動を同一のものとして判断しなければなら



(a) マーカの位置 (b) 実際の装着

図7 マーカ装着の様子

い。そこで、提案システムが類似行動に対してどれぐらいの予測性能があるかを調べ、モデル適合度の閾値設定するための指標とする。本稿では得られた予測性能の評価結果について報告する。評価方法として Jack-Knife 法を用い、類似行動の予測誤差を調べた。Jack-Knife 法とは全サンプルから、1つのサンプルを除き、残りのサンプルでモデルを構築したときに、あらかじめ抜いていたサンプルをどれぐらいモデルに適合するかを検定する方法である。今回の実験では体形の異なる被験者7人に対しそれぞれ3つずつの行動パターンを取得した。各行動パターンに対し、6人のデータを学習データとして RNN で学習し、行動モデルを構築した。残り1人をテストデータとして予測を行い、予測誤差を測った。1パターンの行動につき、7通りの計測を行った。

#### 3.1 実験環境

PhaseSpace という光学式 Motion Capture System を用いて、時間軸に沿った体のモーションデータを取得した。被験者の上半身にマーカーを22点 (頭:6点, 肩:2点, 両肘:6点, 両手首:8点) を付け、マーカー装着の様子を図7に示す。各マーカーの位置情報 (x,y,z) を実験室に設置させた4台のカメラから読み取った。カメラの測定範囲は 4m×4m である

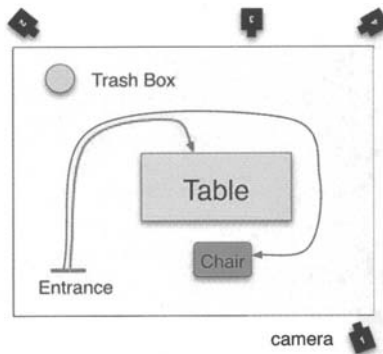


図8 実験室のレイアウト

人がオフィス内での行動を想定し、実験室を図8のように設置した。行動パターンを秘書行動(P1)、事務作業(P2)、テーブル掃除(P3)の3つを取った。実験の様子を図9、図10、図11に示す。各行動パターンの概要は以下である。

- P1 (秘書行動)：入室 → お茶をテーブルの上に置く → 退室
- P2 (事務作業)：入室 → 椅子に座る → 新聞をめくる → 手元のゴミをゴミ箱に捨てる → 退室
- P3 (テーブル掃除)：入室 → 椅子に座る → ノートを整理する → 一冊のノートをゴミ箱に捨てる → 雑巾をかける → 退室

各被験者・行動パターンにつき、4回ずつ実験を行った。被験者によって、身長・歩き方・行動時間が異なり、被験者データと各行動パターンの平均実験時間を表1に示す。表1により、汎化性実験に用いる実験データはばらつきのあるものといえる。



(a) お茶を持った状態 (b) お茶を置く行動  
図9 秘書行動



(a) 新聞をめくる行動 (b) ゴミを捨てる行動  
図10 事務作業



(a) ノートを整理する行動 (b) 雑巾をかける行動  
図11 テーブル掃除

被験者 (性別)	身長 (cm)	P1(秒)	P2(秒)	P3(秒)
1 (女)	167	23.3	47.7	77.3
2 (男)	167	21.3	45	77
3 (男)	170	23.7	49.7	82.3
4 (男)	172	20.7	46	73.7
5 (女)	157	22	49.3	76.7
6 (男)	175	22	49.3	76.7
7 (女)	158	21	47.7	77.3

### 3.2 実験結果

22点のマーカを用いて、動きの位置情報のデータを取得した。その後各部位のマーカ群の中心点を求め、それぞれを各部位の特徴点の動き軌跡とする。今回は学習用の特徴点として、右肘、左肘と頭の3つを選び、合計9次元の位置情報データをRNNに入力し、行動モデルを構築した。Jack-Knife法を用いて、未学習被験者の動きに対して予測を行い、予測誤差を測った。予測結果の例として、図12にP1(秘書行動)を示す。ここでは各特徴点(右肘、左肘、頭)の $x$ 座標値のみプロットした。破線をテストデータの入力値、実線をRNNでの出力値を表す。図12により、破線と実線はほぼ一致しており、未学習の行動に対しても正しく予測していることがわかる。各行動パターンに対して、各次元での平均予測誤差を図13、14、15に示す。縦軸は各ステップでの予測誤差を足し合わせ、ステップ数で割った平均予測誤差となる。横軸をそれぞれ入出力次元(右肘の $x, y, z$ 、左肘の $x, y, z$ 、頭の $x, y, z$ )を表す。各行動パターンの平均予測誤差を以下のようにになった。

- P1 (秘書行動)：平均予測誤差は6.3cm以下。
- P2 (事務作業)：平均予測誤差は3.9cm以下。
- P3 (テーブル掃除)：平均予測誤差は3.9cm以下。

3行動パターンの平均予測誤差を見ると、すべての特徴点(左肘、右肘、頭)の $y$ 座標(高さ)方向の予測誤差が比較的に小さいことがわかる。それは高さ方向へ移動が少ないと考えられる。これらの行動でモデルを構築する際、 $y$ 方向の閾値を $x, z$ 方向より小さく設定すればよいと思われる。

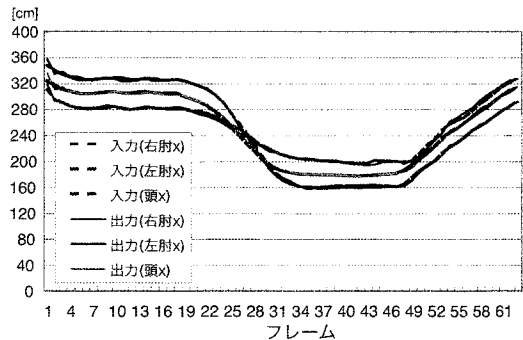


図12 P1(秘書行動)の予測結果

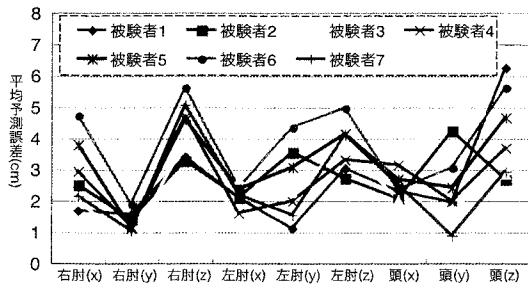


図 13 P1 の平均予測誤差

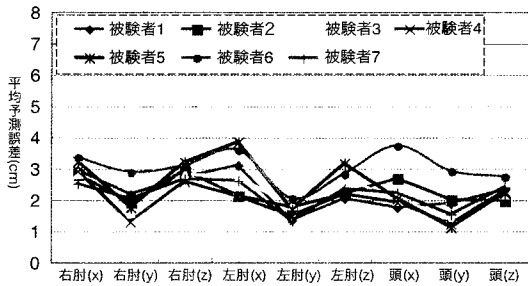


図 14 P2 の平均予測誤差

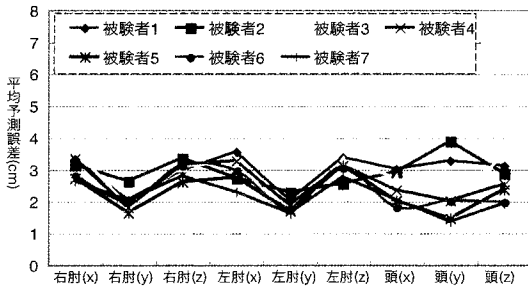


図 15 P3 の平均予測誤差

#### 4. まとめと今後の課題

本研究では人物行動モデルの自動分節化を目標としている。そのアプローチとして、RNNを用いて人物行動モデルを構築する方法を用い、モデル適合度に基づく自動分節手法を提案した。RNNモデル適合度の閾値を決めるため、汎化性評価を行った。その結果、平均予測誤差が3.9cm-6.3cmの精度が得られた。動きが小さい特徴点については予測誤差が小さくなる傾向があることが分かった。

今後は予測誤差に基づく閾値の設定、自動分節化アルゴリズムを実装する。行動パターンと特徴点を増やし、追加学習の性

能評価を行う予定である。

#### 文 献

- [1] 郷古学, 伊藤宏司, "環境の分節化と予測に基づく内部モデルの獲得," 第19回自律分散システム・シンポジウム資料, pp. 43-48 2007
- [2] N.M.Oliver, B. Rosario and A. P. Pentland, "A Bayesian Computer Vision System for Modeling Human Interactions," *IEEE Trans.PAMI*, Vol. 20, No. 8, pp. 831-843, 2000.
- [3] Thomas O. Meservy, Matthew L. Jensen, John Kruse, Douglas P. Twitchell, Gabriel Tsechpenakis, Judee K. Burgoon, Dimitris N. Metaxas, Jay F. Nunamaker Jr., "Deception Detection through Automatic, Unobtrusive Analysis of Non-verbal Behavior," *IEEE Intelligent Systems*, Vol. 20, No. 5, September/October 2005.
- [4] 青木康洋, 岩井儀雄, 谷内田正彦, "全方位画像による例外行動の検出," 情報処理学会研究報告, *CVIM*, Vol. 2006, No.51, pp. 91-98, 2006.
- [5] 北橋忠宏, 福永邦雄, 小島篤博, 長田典子, 樋口未来 "人間行動とその対象物体との機能に着目した協調的認識機構と認識結果の自然言語表現" 「特定領域研究」17年度報告書
- [6] J. Tani, M. Ito, "Self-Organization of Behavioral Primitives as Multiple Attractor Dynamics: A Robot Experiment," *IEEE Trans. On Systems, Man, and Cybernetics, Part A: Systems and Humans*, Vol. 33, No. 4, pp. 481-488, 2003.
- [7] D. Rumelhart, G. Hinton, and R. Williams: Learning internal representation by error propagation, In D.E.Rumelhart and J.L.McClelland, eds., *Parallel Distributed Processing*, MIT Press, Cambridge, MA, 1986.