

リカレントニューラルネット予測器を用いた 人物行動のモデル化と例外行動検出

牛 童[†] 岩井 儀雄[†] 谷内田正彦[†]

[†] 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 本研究では監視カメラによる例外行動の自動検出を目標とする。通常の人間行動モデルを構築し、それに当てはまらないものを例外行動として報告する。人物行動モデルの構築にはリカレントニューラルネット (RNN) 予測器による手法を用いる。認識時には、観測特徴量と RNN の予測値との予測誤差を計算する。予測誤差をもとにモデルの適合性を計算し、全てのモデルが不適合となると例外行動として検出する。本システムは、追加学習が容易に行え、認識時には例外行動発見を迅速に行うことができ、不審行動に対するより素早い対応・対処が可能になる。

キーワード RNN, 予測器, 例外検出

Modeling Human Actions and Detection of Suspicious Actions based on Predictor Model using RNN

Tong NIU[†], Yoshio IWAI[†], and Masahiko YACHIDA[†]

[†] 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 Our purpose is to detect suspicious actions automatically from a security camera. In this paper, we propose a system for tracking people and detecting a suspicious action when it does not match with the learned human action models. Human actions are modeled by a predictor model using recurrent neural networks. A suspicious action is recognized when not the all human action models match with it. The validity of the action model is measured by the error between a predicted value and an input trajectory. The proposed system can deal with additional learning easily and detect a suspicious action rapidly.

Key words RNN, predictor model, detection of suspicious actions

1. はじめに

近年、犯罪の増加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。そこで、便利で安心できる社会を実現するため、監視システムの開発が求められている。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られることができ、コストの削減や異常事態の早期発見などが期待できる。

上記のような理由から、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 [1], [2]。例外行動を検出するためには、人物を抽出し、追跡する追跡モジュールと、人物の行動パターンを学習し、認識する認識モジュールが必要である。追跡モジュールでは人物行動に関する特徴量を抽出し、認識モジュールでは追跡モジュールで取得した特徴量を用いて、行動学習と認識を行う。人物の抽出と追跡の処理としては、動き情報と色情報などを用いる手法が提案されている [3]。一方、行動認識の手法としては、時系列の特徴量を入力し、識別クラスを出力するパターンクラスタリングが一般的である [4]。これは行動軌跡全体をモデル学習するため、新しいパターンを追加学習する際、モデル全体を再学習する必要があるといった問題点がある。

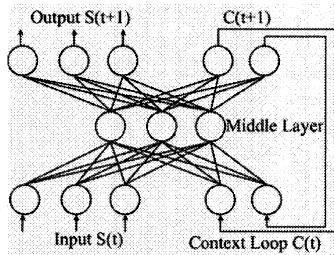


図1 RNNの構成

これに対して近年では、脳の中の神経回路網を数学的にモデル化したニューラルネットワークを用いる認識手法が注目されつつある [5], [6]。これらの研究ではリカレントニューラルネットワーク (図1) を予測器として行動モデルを構築している。応用例として、音声認識 [7] やロボットの行動生成 [5] などが挙げられる。特に Haruno ら [5] が提案するロボットの行動生成での応用例ではモデル学習時に行動軌跡を分割し、各分割軌跡を独立に予測器で学習し、各予測器を上位モジュールで制御することで追加学習を容易に行えるシステムを実現している。

このような背景を受け、本研究では行動認識モジュールに、予測器を用いたシステムを提案する。具体的には、リカレントニューラルネットワークを用いて通常の人物行動モデルを構築し、学習したモデルに当てはまらない例外行動を自動検出する。リカレントニューラルネットワークは現状態から次状態を予測する予測器として働き、コンテキスト層によって過去の履歴も考慮した予測が可能である。提案手法は Haruno らの手法 [5] と同様にリカレントニューラルネットワークを独立に学習させるため、パターンクラスタリングと比較すると、追加学習が容易に行えると考えられる。本論文ではこれまで構築した基礎システムの構築方法とその評価について報告する。

2. 人物行動のモデル化

2.1 人物行動モデル

本研究では、青木らの手法 [9] と同様に人物の重心位置 (x, y) を観測特徴量として追跡し、学習することで行動モデルを構築する。学習データとして、正常な歩行軌跡をリカレントニューラルネットワークによって学習する。リカレントニューラルネットワークは、現時刻の状態を入力し、次時刻の状態を出力する予測器とする。学習方法としてはBPTT(Back Propagation Through Time) [11] を用いて、各層の結合重みの誤差が小さくなるように修正していく。

複数の動作系列を一つのモデルで学習する手法として複数のリカレントニューラルネットワークを用いてモデルを構築する「局所表現」[5] と一つのリカレントニューラルネットワークに制御用のパラメータを付加する「分散表現」[10] が存在する。これらのモデルでは動作を分割できる最小単位 (プリミティブ) ごとにリカレントニューラルネットワークの挙動を変化させる。「局所表現」のモデルではリカレントニューラルネットワーク間の遷移によってこれを実現し、「分散表現」のモデルでは制御

用パラメータを変更することでこれを実現する。「分散表現」のモデルは全ての動作系列を単一のリカレントニューラルネットワークで学習するため、新しい軌跡を追加学習することが困難である。そこで、本研究では図2に示すような「局所表現」のモデルを用いる。

2.2 サブモデル分割による学習データの削減

近年、行動プリミティブに基づいた行動認識、生成に関する研究が盛んになりつつある。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。これらのパーツが行動プリミティブにあたる。多様で複雑な行動は、手持ちのいくつかの行動プリミティブ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、これを行動の分節化という。行動を分節化し、類似行動を同一のプリミティブとして統合、それぞれを独立に別々のリカレントニューラルネットワークで学習する。ここで、リカレントニューラルネットワークで学習された行動プリミティブ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。

図2では、サブモデルが5つあり、モデルの数は3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。パターンクラスタリングでは主に確率的モデルが用いるため、連結するサブモデル間の遷移は確率的に定義されてきたが、リカレントニューラルネットワークを用いる場合は、遷移する際、連結するサブモデル全てに遷移することになる。すなわち、認識時に RNN1a の予測が終了すると、RNN2a と RNN2b を同時に利用して、行動を予測していくことになる。

2.3 追加学習

「局所表現」のモデルを用いることにより、該当するリカレントニューラルネットワーク (サブモデル) のみを更新することで追加学習をすることが可能である。本モデルでは学習時にプリミティブに分割された動作系列は独立に学習され、その結合としてモデル全体が表現される。一方、パターンクラスタリングでは結合部も確率的に遷移するため、追加学習する際には全モデルを更新する必要がある。

図3に「局所表現」のリカレントニューラルネットワークを用いた追加学習の一例の概略図を示す。新しい学習データが2

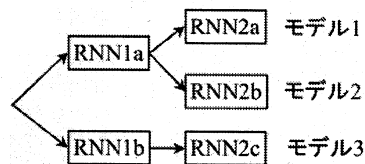


図2 局所表現の行動モデル

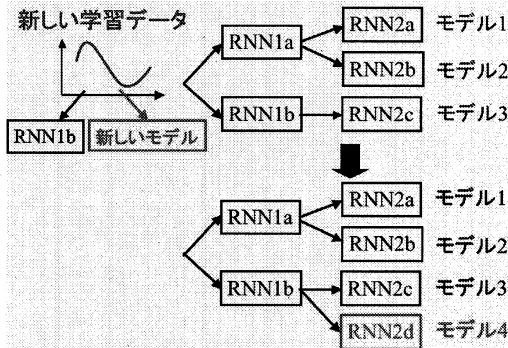


図3 追加学習の概略図

つのプリミティブに分割でき、その前半部がRNN1bによって予測可能であったとする。この時、サブモデルRNN1bに新しい学習データ用のサブモデルRNN2dを付加することで新しいモデルを構築することが可能である。このように、新たなサブモデルを付加し、それを学習することで追加学習を実現することが可能である。

2.4 例外行動検出手法

本研究では予測可能性を定量的に評価することでモデルの適合度を判定する。全てのモデルが不適合になると例外行動として判定する。観測された行動軌跡を予測する時、RNNに現在の観測特徴 $S(t)$ を入力することで次時刻の予測観測特徴量 $S'(t)$ が出力される。この予測値と次の観測特徴量 $S(t+1)$ を比較することでモデル適合度を判定する。具体的には数ステップ分の予測値と観測特徴量を(1)式で評価し、閾値より大きくなると例外行動とする。ここで、 f は評価関数、 T_e は閾値、 h は評価するステップ数、 $S(t)$ は観測特徴量、 $S'(t)$ はRNNの出力である。

$$f(S(t+1-h), S'(t-h), \dots, S(t+1), S'(t)) > T_e \quad (1)$$

本稿では簡単のため、評価関数は各ステップにおける予測値と観測特徴量の差の絶対値を取る。また、これらの差があらかじめ設定したステップ分、閾値 T_e より大きくなるとモデル不適合として判定する。すなわち、式(2)が満たされると、モデル不適合とする。

$$\sum_{\tau=0}^h |S(t+1-\tau) - S'(t-\tau)| > T_e \quad (2)$$

3. システムの構成

本システムの概要を図4に示す。本システムは、追跡モジュールと認識モジュールによって構成される。追跡モジュールではカメラからの入力画像をもとに人物の抽出と追跡を行う。認識モジュールでは、追跡モジュールで抽出された特徴量をもとに人物の行動モデルを学習する部分と、そのモデルをもとに行動を予測し、例外行動を検出する部分からなっている。

3.1 追跡モジュール

追跡モジュールでは入力画像から人物領域を抽出し、特徴量

を計算する。特徴量 $S(t)$ として、画像内の重心位置 $(x(t), y(t))$ を用いる。画像内の人は常に一人とし、人物以外の移動物体は存在しないと仮定する。また、一連の行動の途中で、人が画面から隠れてしまうようなオクルージョンの発生やある位置に止まってしまうような場合も考慮しないことにする。このような設定により、行動軌跡 $\{x(t), y(t)\}$ は、人物が画像内に現れてから消えるまでの重心位置の変化として観測できる。

人物領域抽出は背景差分によって行う。まず前処理として、背景画像を取得しておく。次に、背景画像と入力画像との差分をとり、固定閾値 T_b で二値化処理を行うことで、人物候補領域の抽出を行う。このとき、光の変化や背景のわずかな動きの影響で、人物以外のところも検出されてしまう。人物領域を抽出するためには、これらの領域を除去することが必要である。ノイズ除去の手法として、二値化で得られたすべての候補領域の面積を計算し、その中一番大きいものを人物領域とする。

3.2 認識モジュール

認識モジュールは学習部分と予測部分から成る。認識時には、追跡モジュールで取得した観測特徴量とリカレントニューラルネットワークの予測値との予測誤差をもとにモデルの適合度を評価する。全てのモデルが不適合となると例外行動とし、いずれかのモデルの予測が終了すると正常行動とする。

3.2.1 行動モデルの実装

本研究では学習動作として比較的単純な動きを用いる。そこで、構成したりカレントニューラルネットワークの中間層の層数は10とし、コンテキスト層の層数は10とした。予備実験により、この層数で学習が成功することを確認した。実際には学習パターン数、複雑さに合わせて、最適な中間層とコンテキスト層の層数の組み合わせを調べることが必要になる。また、リカレントニューラルネットワークの各層の出力にはシグモイド関数をかけるため、重心位置はシグモイド関数の値域である $[0, 1]$ に正規化する。

学習時に人物の特徴量に加え、行動軌跡をどこまで予測したかを示す「フラグ」もリカレントニューラルネットに入れて一緒に学習する(図5)。予測器の特徴として、学習時に学習軌跡全体の長さに関する情報が失われてしまう。本システムでは予測中の箇所をフラグとして一緒に学習し、認識時には予測箇所を内部連想することでこれを解決した。サブモデル間の遷移はフラグによって実現できる。特徴量は重心位置 (x, y) の2次元を用いており、フラグと合わせて3次元の学習データを作成する。このフラグは0.1から0.9までの値を学習データに合わせて均等に分割し、均等に分割されたそれぞれの値を学習データの各ステップに割り当てる。例えば、ある重心のデータが

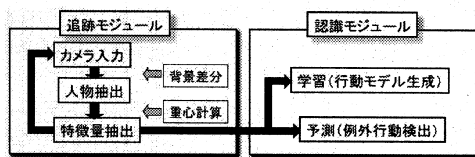


図4 システムの概要

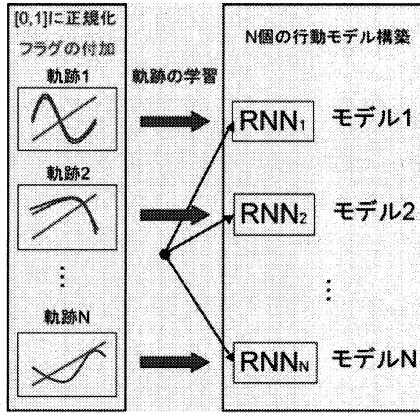


図5 フラグを入れた行動モデル

N 個あるとき、第 i 番目のデータに割り当てられるフラグは $0.1 + 0.8 \times (i - 1) / (N - 1)$ となる。このように 0.1 は 1 番目の学習データを表し、 $0.8 / (N - 1)$ の刻みで徐々に増えていき、 0.9 が最後の学習データに対応する。これにより現在予測中の箇所を内部状態としてフラグで予測することができ、 0.9 になると現サブモデルでの予測が終了したことを意味する。

学習時、同一の軌跡は同じリカレントニューラルネットワークで学習してモデルを構築する。各モデルは別々のリカレントニューラルネットワークで学習する。作成した重心とフラグのデータを教師信号とし、誤差が小さくなるように重みを修正していく。

3.2.2 認識処理の実装

認識時には特徴量である重心とフラグは別々の方法で処理する。重心は追跡モジュールで観測可能であるが、内部状態であるフラグは観測不可能である。各ステップで、重心は観測値を、フラグは前ステップの出力値を入力する。ただし予測開始時、フラグの値として、 0.1 を入力する。このように、このモデルではリカレントニューラルネットワークが軌跡のどの部分を予測しているかということ、フラグによって自己連想しているのである。

例外行動の判定は N 個の行動サブモデルであるリカレントニューラルネットワークのモデル適合性に基づいて行う。モデル適合性は現ステップの重心の予測値 $(x'(t), y'(t))$ と次ステップの観測値 $(x(t+1), y(t+1))$ をもとに評価する。本稿では、各モデルに対応するリカレントニューラルネットワークを学習しているため、 N 個のモデルに対して N 個のリカレントニューラルネットワークがある。不適合となっていないリカレントニューラルネットワークについて、それぞれの差の絶対値を計算し、過去 h ステップ分の差について閾値処理する。式 (3) に示すように、 h ステップ連続で差の絶対値が閾値より大きくなるとそのリカレントニューラルネットワークを不適合と判定する。不適合となったリカレントニューラルネットワークについては、それ以降の予測処理は省き、全てのリカレントニューラルネットワークが不適合になると例外行動として報告する。

$$\begin{aligned} & \bigwedge_{\tau=0}^h (x(t+1-\tau), y(t+1-\tau)) \\ & - (x'(t-\tau) - y'(t-\tau)) | > T_e \end{aligned} \quad (3)$$

正常行動の判定はフラグの値をもとに行われる。フラグの値は学習軌跡のどの箇所を予測しているかを表しており、これが 0.9 になる、すなわち学習軌跡の終わりまで来るとそれ以降はリカレントニューラルネットワークで予測することが不可能である。一方、予測が終了したことはそのモデルに適合していることも示している。そこで、フラグの値が 0.9 近くの閾値 T_f 以上になり、過去 h ステップでモデル不適合の判定がされていないときに正常行動として報告する。ここで、閾値はリカレントニューラルネットワークの学習時に多少の誤差 μ があり、内部連想のため、誤差が蓄積するため、 0.9 より多少小さい値とする。また、フラグの値が T_f 以上になる前に、追跡が終了すると（観測特徴量が得られなくなると）、不適合となっていないモデルが存在すれば、正常行動とする。

4. 実験

リカレントニューラルネットワークを用いた人物行動モデルの評価実験と例外行動検出実験を行った。ここでは分節化なしと分節化ありの二通りの実験を行った。分節化をせず、行動軌跡を一つのモデルとして学習させ、行動モデルの精度について評価した。また、分節化を行い、予測の正確性を確認した。

4.1 全方位カメラによる認識実験と評価

4.1.1 実験環境

照明の比較的安定している室内を人物が動く様子を実際に撮影し、得られた動画画像を用いて処理を行い、追跡実験を行った。実験環境の詳細を表 1 に示す。撮影した背景画像を図 6 に示す。

表 1 実験環境 (全方位カメラ)

撮影背景	室内
カメラ	全方位 SONY 3CCD カメラ XC-009
画像サイズ	512 × 440 画素
CPU	Pentium4 2.40B GHz
メモリ	1GB
OS	Red Hat Linux 7.3

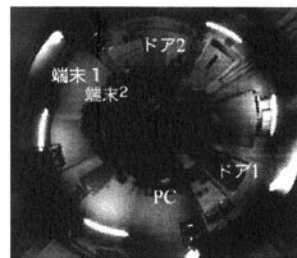


図 6 研究室内の全方位画像

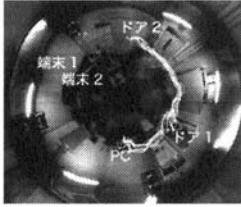


図7 学習軌跡

表2 行動パターン

経路	行動
d1	ドア1→ドア2
d2	ドア2→ドア1
d3	ドア2→PC
d4	PC→ドア2
d5	ドア1→PC
d6	PC→ドア1

表3 例外行動

経路	行動
e1	ドア2→端末1→ドア2
e2	ぐるっと一周
e3	ほふく前進で入室
e4	ドア2→端末1

4.1.2 RNNの初期値と学習精度の評価

学習データとして表2に示すような6つの行動パターンを選んだ。学習に用いたすべての人物軌跡を図7に示す。一つの行動パターンに3つのデータを用いて学習させた。それぞれの学習データをリカレントニューラルネットワークで学習し、行動モデルとした。

リカレントニューラルネットワークの学習は最急降下法で重みを更新するため、初期値によっては、重みが局所解に収束する場合がある。そこで、重みの初期値と学習精度の関係について、次のような評価実験を行った。

データとして、行動モデルd1(ドア2→ドア1)を構築するとき用いた3つの学習データを利用した。それらをd1-1,d1-2,d1-3とする。実験方法は次の通りである。まず、3つの学習データを3組に分ける。すなわち、単一のデータで学習する場合(d1-1のみ、d1-2のみ、d1-3のみ)、2つのデータで学習する場合(d1-1とd1-2、d1-1とd1-3、d1-2とd1-3)、3つのデータで学習する場合(d1-1とd1-2とd1-3)に分け、計7通りの学習パターンとする。それぞれの学習パターンについて、重みの初期値をランダムに設定し、20回ずつ学習した。それぞれの学習結果について学習成功・失敗をフラグの学習誤差によって判定する。この最大誤差が閾値より大きいときを学習失敗とする。フラグは予測時、リカレントニューラルネットワークの内部で連想されていくため、誤差は蓄積し、予測誤差の一番大きな要因となる。学習成功の閾値を0.1と設定したときの平均成功率は比較的高く、約84%である。学習失敗と学習成功の区別が分かりやすいため、ここでは、d1-1のみを学習した場合で生成したモデルを失敗と成功に分け、それぞれ予測実験を行った。図8と図9はそれぞれ学習失敗モデルと学習成功モデルでの予測結果を示す。赤点は予測データの軌跡を表し、実線はそれぞれのモデルでの予測軌跡を表す。図8より、学習が失敗した場合、予測値は観測値と大きく離れていることが分かる。また、図9より、学習が成功した場合、学習モデルによりほぼ正確に観測値を予測できることが分かった。

次に、学習が成功したのものについて、観測値d1-1,d1-2,d1-3を用いて、予測を行い、予測誤差 $\|S(t) - S'(t)\|$ とその分散を算出した。その結果を図10に示す。横軸の数字はその学習試行の番号を示し、縦軸は予測誤差を表す。これにより、学習に成功すれば、初期値による性能の違いは平均予測誤差 ± 0.0101 以下に収まることが分かった。

4.1.3 モデルの精度

行動モデルの評価方法として、表2に示した行動モデルd1-d6について、学習データを用いて予測を行い、予測誤差を算出した。行動モデルd1-d6には各モデルに3つのデータ(合計18)を用いて学習と予測を行った。平均予測誤差を図11に示す。これにより、これまで作成した行動モデルd1-d6は平均誤差 1.54×10^{-2} (約8画素)以下の精度で予測を行うことができることが示された。

4.2 例外検出実験

表2の学習データをもとに生成した各行動モデルについて例外行動検出実験を行った。例外行動として、表3に示すものを用いた。例外行動の例を図12に示す。モデル適合度の閾値 T_e を0.045と設定し、2ステップで評価する。すなわち、予測誤差がこの閾値 T_e を連続に2回越えたら、モデル不適合と判定し、次時刻以降そのモデルでの予測を行わないことにする。すべてのモデルが不適合となると、その行動軌跡は例外行動とする。実験結果として、すべての例外行動を正確に判定することができた。

例外検出は閾値 T_e の影響を受けると考えられる。そこで、例外行動を判定する閾値 T_e と検出精度の関係を調べた。その結果を図13に示す。構成した行動モデルd1-d6に例外行動と正常行動の軌跡を入力し、識別を行った。太線R1は表3の6個の例外行動データのうち例外として判定できた割合である。細線R2は表2の18個の正常行動のうち正常として判定できた割合である。これらの二つが共に1になるのは全てのデータが正確に判定できた閾値である。本実験では閾値 T_e を0.044から0.051の間にするると全てのデータが正確に判定できる。

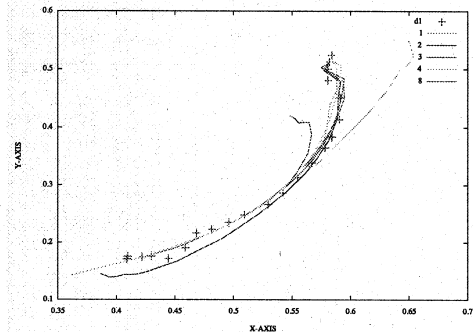


図8 学習失敗モデルでの予測結果

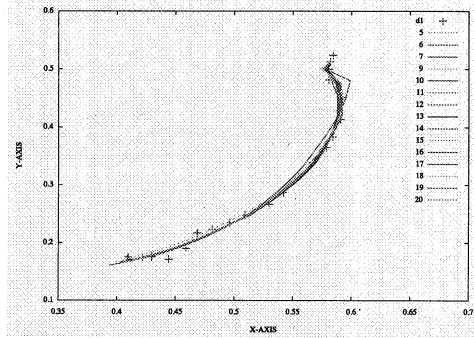


図 9 学習成功モデルでの予測結果

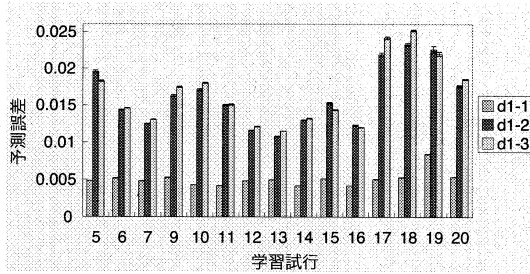


図 10 予測誤差と分散

4.3 分節化実験

4.3.1 実験環境

照明の比較的安定し、広い講義室内人物が動く様子を通常カメラを用いて取り、得られた動画像について追跡処理を行った。実験環境の詳細を表 4 に示す。講義室の様子は図 14 に示す。

4.3.2 分節化行動モデル

学習データとして、図 15 に示すように、「左から右まで歩く」、「左から歩き、中央でターンして、左に戻る」、「右から左まで歩く」、「右から歩き、中央でターンして、右に戻る」という

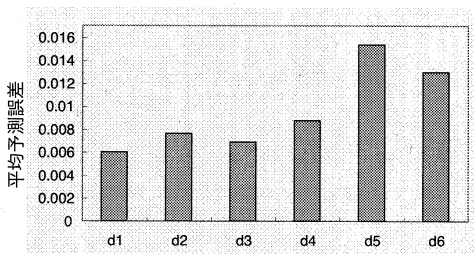
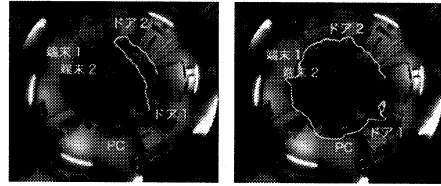


図 11 平均予測誤差



(a) ほふく前進で入室 (b) ぐるっと一周

図 12 例外行動の軌跡

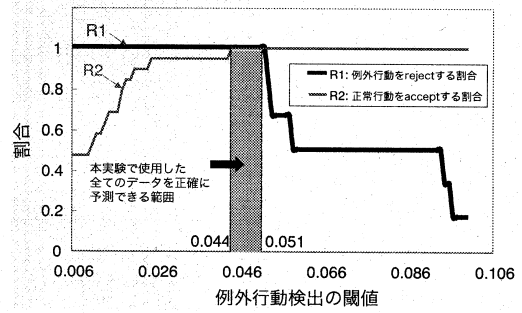


図 13 閾値と検出誤差

4つの行動パターンを撮影した。一つの行動パターンに2つのデータを用いて学習させた。サブモデルと行動モデルの定義は表 5 と表 に示す。このように $M1$ と $M2$ で共通のサブモデル $m1$, $M3$ と $M4$ で共通のサブモデル $m4$ を用いることにより、学習コストを減らすことができた。生成された行動モデルを図 16 に示す。

認識時には図 17 の矢印に示すように、まずサブモデル $m1$

表 4 実験環境 (通常カメラ)

撮影背景	講義室内
カメラ	通常画角 SONY 3CCD カメラ DCR-VX2000
画像サイズ	720 × 480 画素
CPU	PowerPC G4 867MHz
メモリ	1.12GB
OS	Mac OS X10.4.8



図 14 講義室の様子

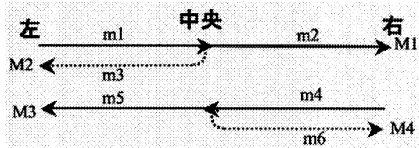


図 15 学習軌跡

表 5 サブモデル

経路	行動
m1	左 → 中央
m2	中央 → 右
m3	中央でターンする → 左
m4	右 → 中央
m5	中央 → 左
m6	中央でターンする → 右

表 6 行動モデル

モデル (経路)	行動
M1: m1→m2	左 → 右
M2: m1→m3	左 → 中央 → 左
M3: m4→m5	右 → 左
M4: m4→m6	右 → 中央 → 右

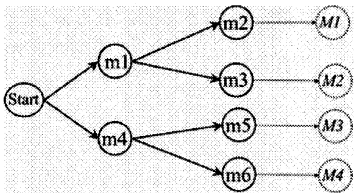


図 16 行動モデル

と m_4 を予測していき、適合度を計算する。行動パターン (左→右) を予測データとすると、 m_4 は不適合となり、 m_4 と m_4 から分節されたサブモデル m_5, m_6 を削除する。つまり、 m_4 - m_6 での予測を行わないことにする。一方、サブモデル m_1 は適合となるので、次に、 m_1 から分節されたサブモデル m_2, m_3 での予測を同時に行っていく。 m_2 は適合となると、その行動パターンはモデル $M1$ に属すると判定できる。表 6 に示す 4 つの行動パターンそれぞれの予測結果を図 18, 19, 20, 21 に示す。縦軸は正規化した重心位置の値、横軸は重心データの番号を示す。マークなしの曲線は観測特徴量を表し、マーク付きの曲線は前時刻の観測特徴量より得られた予測値を表す。丸で囲んだ部分は分節化の境を示す。図 18-21 により、分節化モデルにより、ほぼ正確に観測特徴量を予測できたことが分かった。今後は全方位カメラでより複雑な行動パターンを取り、分節化を行う予定である。

5. おわりに

本稿では、人物の異常行動を例外行動として自動的に検出可能な監視システムへの応用を目的として構築した基礎システムについて報告した。リカレントニューラルネットワークによる

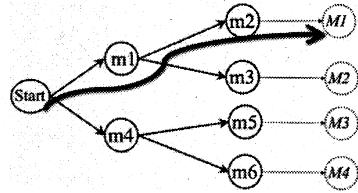


図 17 認識プロセス

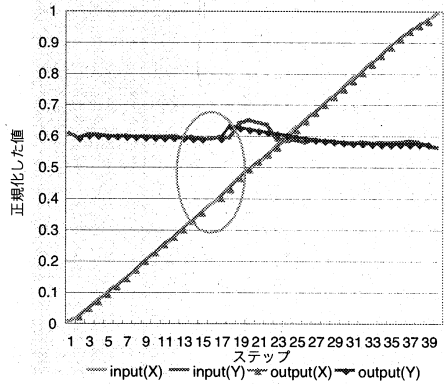


図 18 左 → 右

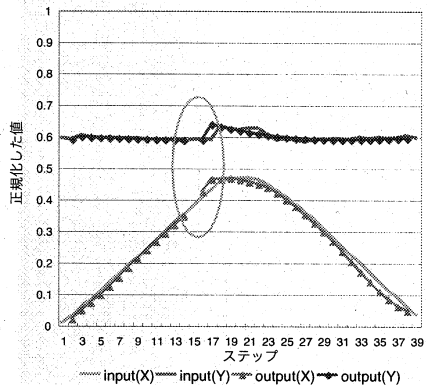


図 19 左 → 中央 → 左

予測器を用いて人物行動モデルを構築し、例外行動を検出するシステムを構築した。

追跡モジュールは簡単のため、背景差分によって、人物を抽出し、特徴量として重心位置を算出した。認識モジュールでは、人物行動として重心の軌跡を学習し、行動モデルを構築した。本モデルは追加学習が容易にできるように別々のリカレントニューラルネットワークで構成されており、それぞれが独立に学習される。認識時には、現時刻のリカレントニューラルネットワークの予測値と次時刻の観測特徴量との予測誤差でモデル適合度を評価した。予測誤差が閾値より大きくなると不適合とした。全てのモデルが不適合になると例外行動とした。そうでない場

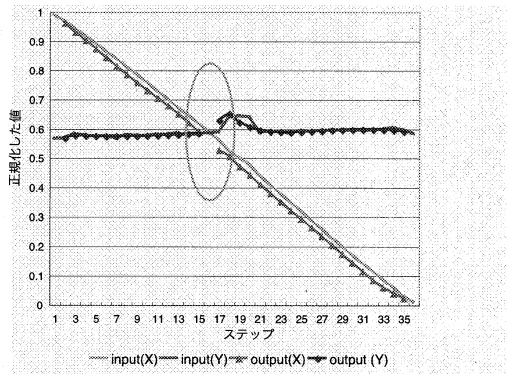


図 20 右 → 左

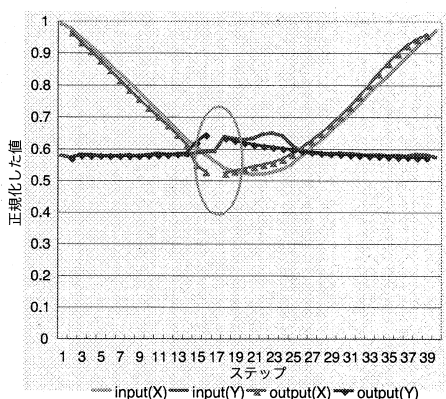


図 21 右 → 中央 → 右

合は正常行動と判定した。評価実験により、本システムが正確に例外行動検出できることを確認した。行動モデルの精度評価として、学習データを予測し、予測誤差を算出する。その結果、平均予測誤差は 1.54×10^{-2} (約 8 画素) 以下の精度で予測を行うことができた。また、リカレントニューラルネットワークの重みの初期値を変えて結果にどう影響するかを調べた。初期値による性能の違いも平均誤差で ± 0.0101 以下に収まることが分かった。よりパフォーマンスを高めるには例外検出の閾値を調整する必要がある。また、分節化実験により、学習データの削減ができ、予測も正確に行うことを確認できた。

今後の課題は主に本システムを発展させることである。まず、例外行動の評価関数の改良することによってより複雑な軌跡にも対応できるようにする。次に例外行動の判定について、より有効な方法を開発する。それから、本システムの利点である追加学習の性能を評価する。さらに行動モデルをサブモデルに自動分割する手法を実装し、オクルージョン時の内部連想機能を評価することが今後の予定である。

文 献

[1] N.M.Oliver, B. Rosario and A. P. Pentland, "A Bayesian Computer Vision System for Modeling Human Interac-

tions," *IEEE Trans.PAMI*, Vol. 20, No. 8, pp. 831-843, 2000.

- [2] 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.
- [3] 志賀泰博, 海老根秀之, 池田美鈴, 中村納, "色情報および動き情報に基づく人物顔領域と顔部品の抽出およびその動きの検出," *電子・情報・システム部門誌*, Vol. 121-c, pp. 912-920, 2001.
- [4] Heisele, B. Woehler, C. Res. Center, Daimler-Benz AG, Ulm, "Motion-based recognition of pedestrians," *Pattern Recognition, Proceedings. Fourteenth International Conference*, Vol. 2, pp. 1325-1330, 1998.
- [5] M. Haruno, D. Wolpert, M. Kawato, "MOSAIC Model for Sensori-motor Learning and Control," *Neural Computation*, Vol. 13, pp. 2201-2220, 2001.
- [6] 三好誠司, 原一之, 岡田真人, "オンライン学習理論に基づく単純パーセプトロンのアンサンブル学習の解析," *電子情報通信学会論文誌 D*, Vol. J87-D2, no. 7, pp. 1391-1401, 2004.
- [7] 内山徹, 高橋治久, "リカレントニューラル予測モデルを用いた不特定話者単語音声認識," *信学論 D*, Vol. J83-D2, No.2, pp. 776-783, 2000.
- [8] A. Matsumura, Y. Iwai, M. Yachida, "Stochastic Action Recognition from Omnidirectional Images," *Proc. ACCV*, Vol. 1, pp. 120-125, 2004.
- [9] 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.
- [10] 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.
- [11] D. Rumelhart, G. Hinton, and R. Williams: Learning internal representation by error propagation, In D.E.Rumelhart and J.L.McLlelland, eds., *Parallel Distributed Processing*, MIT Press, Cambridge, MA, 1986.