

反辞書確率モデルを用いた監視カメラの 人物異常行動の検出

酒井 啓正¹ 陳韜^{†1} 森田 啓義¹ 眞田 亜紀子¹

Abstract: 本稿では、反辞書確率モデルを用いて監視カメラなどに映り込んだ人物の異常行動を検出する方法を提案し、実験を行うことでその検出精度を求める。異常行動を検出する場合、大きく分けて二つの手法が存在する。一つは異常な行動をあらかじめモデルデータとして定義し、モデルとのマッチングで検出する手法で、モデルデータにない行動を検出できない恐れがある。もう一つは観測頻度の低い行動を異常として検出する手法で、観測頻度の高い異常行動が検出できない恐れがある。そこで、人物の軌跡データに現れないパターンの集合である反辞書の作成と軌跡データの出現確率に基づくスコアリングを行なう手法を新たに提案する。提案手法を監視カメラで撮影された人物の歩行映像に適用することで、正常行動データに対しての再現率・適合率がそれぞれ 84%・49%、異常行動データに対しての再現率・適合率がそれぞれ 82%・96% となった。

1. はじめに

2020 年東京オリンピックに向けてのセキュリティ市場の拡大にともない、監視カメラ市場の拡大も予測される。しかし、監視カメラが増えればそれを監視するために多数の人員が必要となる。また、一人の人が監視できる監視カメラの台数や時間は限られている。そのため、監視カメラの映像を自動解析するシステムの必要性は高い。

本稿では、反辞書確率モデルを用いて監視カメラが捉えられた人物の異常行動を自動検出する方法を提案する。

異常行動を検出する場合、大きく分けて二つの手法が存在する。一つは異常な行動をあらかじめモデルデータとして定義し、モデルとのマッチングによって異常行動を検出する手法である。[1], [2]。もう一つは観測頻度の低い行動を異常として検出する手法である [3], [4], [5]。前者はモデルデータとして定義されていない行動を検出できず、後者は観測頻度の高い異常行動を検出できないという問題点がある。

本稿ではこれらを解決するために、正常行動の軌跡データに出現しない軌跡パターンを集めた反辞書確率モデルに基づくスコアリングを行なう手法を提案する。さらに、提案手法を監視カメラで撮影された人物の歩行映像に適用した実験結果を報告する。

反辞書確率モデルは、有限マルコフ過程の一種であり、元来、テキストや計算機ファイルなどのファイル圧縮技法においてエントロピー符号化を行うために考案された [6]。反辞書確率モデルは算術符号と組み合わせることによって、シンボル単位でエントロピー符号化を効率よく行える。その効果は心電図波形データの無歪み圧縮においても確認されている [7]。

反辞書確率モデルでは、データに出現しないパターンを十分考慮したモデル構築がなされるため、通常マルコフ確率モデルで生じるゼロ確率問題（未出現のシンボルに一定以上の正の確率を付与せざるをえないため結果的に圧縮効果の劣化を招く）を回避できる点が優れている。

一方で、反辞書確率モデルを用いたデータ圧縮法では、データに出現しないはずのパターンがなんらかの理由で生じた場合には、一時的に圧縮効果が著しく劣化する。この際立った性質を正常行動から外れた異常行動の検出に適用しようというのが本研究の狙いである。筆者らのグループはすでに、反辞書確率モデルを用いて心電図波形に含まれる不整脈の検出を行っており、既存の手法と比べても、遜色ないか、あるいはそれ以上の検出性能を得ている [8]。

以下、次節では従来の異常行動検出法をより詳しく紹介する。3 節では、本研究に用いる反辞書について紹介し、反辞書確率モデルを用いた異常行動検出法を提案する。4 節では、提案法を用いた異常行動検出実験について述べ、実験結果について考察する。5 節で本稿の内容をまとめ、今後の課題について論じる。

2. 異常検知手法

2.1 モデルデータと合致する行動を検出する手法

文献 [1], [2] では、不審行動を予めモデルデータとして定義し、モデルデータと合致する行動を不審行動として検出している。文献 [1] では、夜間等の無人となる時間帯に扉の開閉を不審行動と定義し、あらかじめ扉に付けておいた扉開閉センサにより侵入を検知している。また、異常行動を観測した場合 Web カメラから得られた映像を携帯電話やパソコンで確認できるシステムを提案している。文献 [2] では、高解像度撮影可能な全方位ビデオカメラを用いて得られた画像から注視すべき対象物を追跡し、追跡結果から出現・退出・静止・作業などのイベントに分類している。また、対象が多地点で起こすイベントを同時に検出することができる。

2.2 観測頻度の少ない行動を不審行動として検出する手法

文献 [3], [4], [5] では、初めて観測された行動や観測頻度の低い行動を不審行動として検出している。文献 [3] では、人物の行動モデルを行動素と呼ばれる要素で表し、予め頻繁に観測される行動を行動モデルとして学習し行動認識を行なうことで、未知の行動を観測したときに不審行動として検出して

¹ 電気通信大学大学院情報システム学研究所

^{†1} 現在、インフォメーション・ディベロップメント

いる．文献 [4] では，周囲 360 度の撮影が一度に可能な全方位視覚センサにより撮影された画像を用いて人物追跡を行ない，得られた歩行軌跡から人物の行動モデルを作成し，未知の行動を行なう人物を不審人物として検出している．また，従来研究での人物追跡の処理時間が膨大であるといった欠点を改良し，人物追跡を 0.3 秒で行なっている．文献 [5] では，カメラ映像から得られた人物の追跡結果に対して，主成分分析を行うことで，通常の行動から逸脱した行動を検出している．

2.3 従来研究の問題点

文献 [1], [2] の手法は，予め定義した不審行動を正しく検出することが出来るが，多様な行動特徴を有している不審行動を予め全て定義しておくことは難しい．

文献 [3], [4] の手法は，不審行動を予め定義する必要がないが，頻繁に観測されるような不審行動の検出が難しい．また，観測頻度の低い行動が必ずしも不審行動とは限らないという問題がある．

文献 [5] では，主成分分析を用いて，正常な人物歩行軌跡から逸脱したものを検出しているが，あくまで検出の可能性を示唆している段階であり，実際に多数の人物歩行軌跡を用いて実験を行い精度を求めるには至っていない．

3. 反辞書確率モデルによる異常検知

与えられたデータ列に現れないパターンを表す反辞書の作成と，データの出現頻度に基づくモデリングを行うことで，反辞書確率モデルを作成する．作成された反辞書確率モデルを用いて，未知のデータに対してスコアの計算を行なうことで，より精度の高い検出を目指す．概要を Fig. 1 に記す．

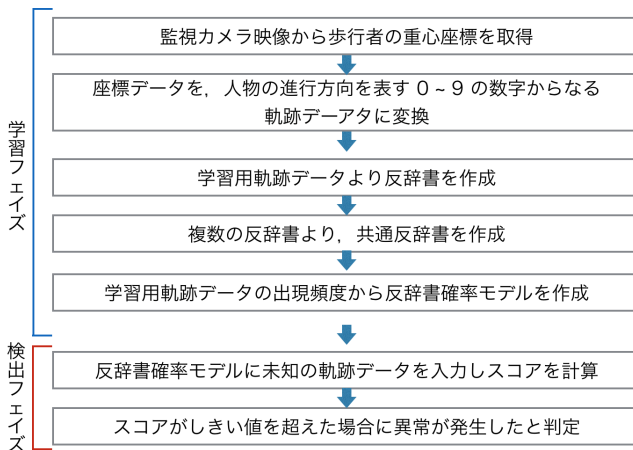


Fig. 1 異常検出の概要

3.1 人物追跡方法について

監視カメラによる人物追跡の方法は，筆者らの研究グループと熊平製作所が共同開発したシステムを用いる [9]．このシステムでは，H.264 ビデオ圧縮標準規格エンコーダーを搭載したネットワークカメラ (8fps) を使い，カメラ出力である圧縮ビデオデータから，フレームごとに抽出された動きベクトルと圧縮に用いる符号器の種類を表すマクロブロックタイプ移動体を取り囲む長方形領域とその重心座標を算出することによって人物追跡を行なっている．

3.2 軌跡データ

監視カメラ映像から抽出された重心座標列の差分をとり，人物の進行方向ベクトルを算出する．つぎに進行方向ベクトルを Fig. 2 に示す方法で 0 から 7 までの数字に量子化する．さらに，ベクトルの大きさに関するしきい値 α と β ($\alpha < \beta$) を設定することで，ベクトルの大きさが α より小さい場合は数字 8 を， β より大きい場合は数字 9 を，ベクトルの進行方向に関わらず割り当てる．このようにしてフレームごとに量子化された進行方向ベクトルを表す数字の列を，ここでは軌跡データと呼ぶ．

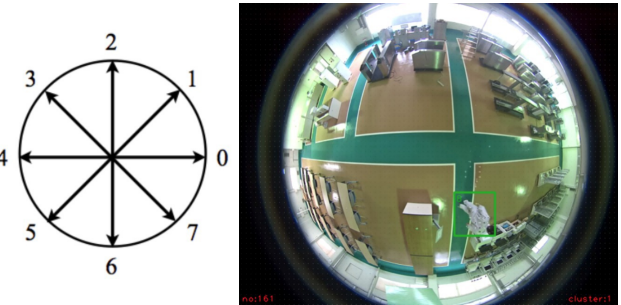


Fig. 2 監視カメラの様子 (右) と人物の移動方向割り当て (左)．真上から見た人物の移動方向を計算し，それに対応する数字を軌跡データとして出力する．例外的に，移動量が一定の値より少ないときは 8，多いときは 9 を出力する．

3.3 反辞書

軌跡データはもとより，テキスト文字列や観測デジタル波形などは，一般に，与えられた有限アルファベット \mathcal{X} に値をとるデータ列 x として表される．データ x の反辞書 $\mathcal{A}(x)$ とは， x に現れない部分列の中で，極小なもの集合である．反辞書 $\mathcal{A}(x)$ の元を極小禁止語 (Minimam Forbidden Word: 略して MFW) と呼ぶ．

MFW が極小であるというのは，MFW 自身は与えられたデータ列に含まれないが，MFW から先頭文字もしくは末尾文字を取り除いた，MFW より長さが 1 短い部分列は，いずれもデータに含まれることを意味する．例えば， $x = 01011$ に対して， $w = 110$ は x に現れないが， w から先頭文字もしくは末尾文字を取り除いて得られる部分列 10, 11 はいずれも x に現れるので， w は x の MFW である．

アルファベットが $\mathcal{X} = \{0, 1\}$ の場合，上例の x の反辞書は，

$$\mathcal{A}(x) = \{00, 110, 111, 1010\}$$

である．MFW の数学的な定義は文献 [6], [7] に詳しく記載されている．

3.4 反辞書オートマトン

有限アルファベット \mathcal{X} に含まれる文字から構成された文字列 x の反辞書 $\mathcal{A}(x)$ に含まれるすべての MFW を受理系列とする決定性有限オートマトン $F(\mathcal{A}(x)) = \{Q, \mathcal{X}, \delta, s_0, \mathcal{F}\}$ を x の反辞書オートマトンと呼ぶ．ここで， Q はオートマトンのノード集合， \mathcal{X} は入力文字列のアルファベット， $\delta: Q \times \mathcal{X} \rightarrow Q$ は遷移関数， $s_0 \in Q$ は開始状態，そして， \mathcal{F} は $\mathcal{A}(x)$ の MFW に対応する受理ノードである．

反辞書オートマトン $F(\mathcal{A}(x))$ に， \mathcal{X} に含まれる文字から構成された別の文字列 y を入力した場合，オートマトンの規則からすると，もし y に $\mathcal{A}(x)$ の MFW w が一つでも含まれ

ていれば、最初の w までを入力した時点で、 $F(A(x))$ は y を受理して停止するが、そうでなければ、 y は受理されない。

ただこれでは、“異常な” 入力文字列にいくつの MFW が出現するかを知ることができないので、受理ノード $s_a \in \mathcal{F}$ から遷移が可能になるように遷移関数 δ をつぎのように修正する。

各 s_a には、ある MFW $w = w_1 w_2 \cdots w_m$ が 1 対 1 に対応しており、 s_0 からシンボル w_1, w_2, \dots, w_m に対応する枝を次々に辿ることによって s_a に到達する。 s_a からシンボル $x \in \mathcal{X}$ で遷移する遷移先のノードは、 w から先頭シンボルを取除いた部分列 $w_2 \cdots w_m$ の末尾に x を加えた文字列に対して、 s_0 からシンボル w_2, \dots, w_m, x に対応する枝を次々に辿って到達するノード w とする。すなわち、

$$\delta(s_a, x) = w$$

と定める。

例として、上述の $A(x)$ から構成された反辞書オートマトン $F(A(x))$ を Fig. 3 に示す。図中の太い円が受理ノードを表し、そこから修正した遷移関数によって定まる次ノードへの枝とそのラベルを赤で描いている。

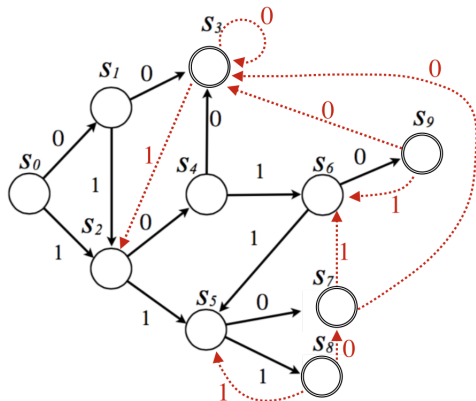


Fig. 3 反辞書オートマトン $F(A(x))$

反辞書オートマトンの構築アルゴリズムは文献 [9] に与えられている。

3.5 反辞書確率モデル

反辞書オートマトン $F(A(x))$ に学習用のある軌跡データ $y = y_1 \cdots y_n$ を入力する。すると、オートマトンは、あらかじめ定められた初期ノード s_0 から始まり、そのノードから出ている枝の中から入力シンボル y_1 に対応する枝を選択して、次のノードへと移る。さらに、シンボル y_k ($1 < k \leq n$) が入力されるごとにノード遷移を繰り返す。

このとき、ノード s からシンボル y に対応した枝 e_y が選ばれて次ノードへ遷移する頻度を $n(y|s)$ とおく。つぎに、この $n(y|s)$ を用いて、現在のノードが s であるという条件の下で入力シンボルが y である条件付き確率を次式で定義する。

$$P(y|s) = n(y|s) / \sum_{\xi \in \mathcal{X}} n(\xi|s) \quad (3.5.1)$$

反辞書確率モデルとは、あるデータ列 x から作成された反辞書オートマトン $F(A(x))$ と、 $F(A(x))$ にある学習用の軌跡データ y を入力することによって算出される条件付き確率 $\{P(y|s), y \in \mathcal{X}, s \in \mathcal{Q}\}$ の組である。

3.6 スコア値計算

未知の軌跡データ $z = z_1 \cdots z_t \cdots z_n$ を反辞書確率モデルに読み込ませ、各シンボル z_t ($1 \leq t \leq n$) ごとにスコア値 $s(z_t)$ を次式で計算する。

$$s(z_t) = \begin{cases} C & \text{if } p(z_t|s_{t-1}) = 0, \\ -\log p(z_t|s_{t-1}) & \text{otherwise.} \end{cases} \quad (3.6.1)$$

ここで C はある一定値である。スコア値の総和 $S(z) = \sum_{t=1}^n s(z_t)$ は反辞書確率モデルを用いて z を算術符号化したときの理想的な符号長に相当している [8]。

実際の追跡実験では、局所的なスコアの変動を抑えるため、時点 t ($1 \leq t \leq n$) におけるスコア値は t を含む直前の T 個のスコア値を算術平均した値 $\tilde{s}(z_t)$

$$\tilde{s}(z_t) = \begin{cases} \frac{1}{t} \sum_{k=1}^t s(z_k) & (t \leq T) \\ \frac{1}{T} \sum_{k=1}^T s(z_{t-T+k}) & (T < t \leq n) \end{cases} \quad (3.6.2)$$

に修正する。そして、修正スコアの総和 $\tilde{S}(z) = \sum_{t=1}^n \tilde{s}(z_t)$ がある与えられたしきい値を超えた場合、データは異常であると判断する。

式 (3.6.1) における C と式 (3.6.2) における T は選ぶ値によって異常判断の検出精度が変化する。適切な値の選び方には文献 [9] で検証されており、次節の実験では、 $C = 25, T = 5$ とした。

4. 実験方法

4.1 実験用のデータ

実験に用いたデータは大きく分けて学習用のデータと検出用のデータの 2 種類が存在する。さらに、検出用データには、walk, fall, run, stop and run, U turn の 5 種類の人物行動データが存在する。学習用のデータは walk のみで、検出用の walk のデータとは異なるものとなる。Table 2 と 3 に実験に用いた軌跡データの内訳と本数を掲げる。

Table 1 学習用データ

種類	本数	分類
walk	121 本	正常

Table 2 検出用データ

種類	本数	異常判定数	分類
walk	50 本	5 本	正常
fall	44 本	37 本	異常
run	108 本	94 本	異常
stop and run	46 本	28 本	異常
U turn	40 本	35 本	異常

4.2 実験の流れ

Table 1 の 121 本の学習用データを 1 本の記号列 x とみなし、反辞書 $A(x)$ を作成する。 x の長さは 8,484 であり、 $A(x)$ は約 3.5 万個の MFW からなる。先行研究から長さ n の文字列は高々 $(|\mathcal{X}| - 1)(n - 1) + |\mathcal{X}|$ 個の MFW を持つことが知られているので、この値は妥当である。しかし、すべてを検出に用いるには多すぎるので、この中から、長さ 2 以下の MFW 26 個を選び出したものを反辞書とした。また、反辞書を作成する段階で、“888” の MFW を加えた。これは、倒

れこみに関する検出精度を向上させるためである．作成された反辞書より反辞書オートマトンを作成し，Table 1 のデータを入力して，(3.5.1) から遷移確率を求めて，反辞書確率モデルを作成した．

4.3 評価基準

評価基準には，正常行動の評価にあたる再現率 1・適合率 1 と，異常行動の評価基準にあたる再現率 2・適合率 2 を次式から求める．

$$\text{再現率 1} = |D_n \cap N_n| / |N_n| \quad (4.3.1)$$

$$\text{適合率 1} = |D_n \cap N_n| / |D_n| \quad (4.3.2)$$

$$\text{再現率 2} = |D_a \cap N_a| / |N_a| \quad (4.3.3)$$

$$\text{適合率 2} = |D_a \cap N_a| / |D_a| \quad (4.3.4)$$

ここで， N_n, N_a, D_n, D_a はそれぞれ以下のデータ集合を表す．

- N_n : 真の正常行動軌跡データの集合
- N_a : 真の異常行動軌跡データの集合
- D_n : 正常行動と判断された軌跡データの集合
- D_a : 異常行動と判断された軌跡データの集合

これらの集合の関係を Fig. 4 に示す．

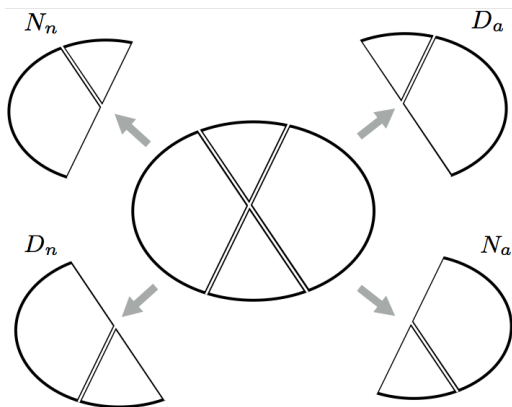


Fig. 4 評価用データの集合

5. 実験結果

異常行動を検出する際に用いるスコアのしきい値を 3 から 7 に変えた場合の再現率・適合率 (すべて% 表記) の挙動を Table 3 に示す．しきい値を大きくしていくと

- (1) 再現率 1 は増加するが，適合率 1 は減少し，逆に，
 - (2) 再現率 2 は減少するが，適合率 2 は増加する
- という関係があることが分かる．これは，しきい値が大きいと真に異常行動である軌跡データが異常とは判断されず，逆に正常行動と判断されるケースが増えることから予想しうる結果であるが，その予想を越えて適合率 1 が著しく低い．これは，検出用に用意した 288 本の軌跡データの中で異常行動に対して正常行動のデータ数が 50 本と少なかったことが一つの原因である．

また，しきい値 3 未満では，ほぼ全ての軌跡データが異常と判断されたため，しきい値としては不相当と判断し表には掲載していない．

Table 3 実験結果

しきい値	再現率 1	適合率 1	再現率 2	適合率 2	平均
3	84	49	82	96	78
4	88	40	72	97	74
5	90	39	70	97	74
6	90	39	70	97	74
7	98	26	42	99	66

今回最も良い精度となったしきい値 3 に設定した場合の検出結果を Table 4 に掲げる．正常行動である walk に関して

Table 4 実験結果 (しきい値 = 3)

種類	本数	異常判定数	分類
walk	50 本	8 本	正常
fall	44 本	37 本	異常
run	108 本	94 本	異常
stop and run	46 本	28 本	異常
U turn	40 本	35 本	異常

は，50 本の軌跡データのうち 16% の 8 本を誤検出している．さらに，walk 以外の異常行動軌跡データでは，それぞれの本数から異常判定数を引いたものが誤検出として加わっている．この表をみると stop and run の場合，正常と判定している軌跡データが 40% 近くあり，このことが適合率 1 を大きく劣化させていると考えられる．

6. おわりに

本稿では，反辞書確率モデルを用いて人物の異常行動を検知する手法を提案した．この手法は，異常行動の検出に関して従来研究とは異なるアプローチをとっている．この手法を用いることで，正常行動データに対する再現率・適合率はそれぞれ 82%・49%，run, fall, stop and run, U turn などの異常行動データに対する再現率・適合率はそれぞれ 82%・96% という結果が得られた．

現在は未知の行動軌跡データを正常行動と異常行動の 2 パターンにのみ分類を行っているが，今後は異常行動に関してさらに細かく分類を行う予定である．

References

- [1] 古谷雅理 他: NCS(ネットワーク カメラシステム) による監視システムの構築と運用, 情報処理学会論文誌, Vol. 46, No. 4, pp. 965-973 (2005).
- [2] 市川徹 他: 高解像度全方位ビデオカメラを用いた遠隔監視システムにおけるイベント検出, PRMU2000-213, pp. 87-94 (2001).
- [3] 岩井儀雄 他: 行動素の混合分布に基づく行動認識と例外行動検出, 電気学会論文誌 C, Vol. 130, No. 4, pp. 546-556 (2010).
- [4] 小坂裕紀 他: 全方位画像を用いた歩行軌跡による確率的人物行動認識システム, 情報処理学会研究報告 CVIM, Vol. 38, pp. 163-170 (2005).
- [5] 鶴飼正人 他: 異常行動検知のための画像処理手法の開発, 鉄道総研報告, Vol. 21, No. 11, pp. 17-22 (2007).
- [6] M. Crochemore, et. al: Data com- Data compression using antidictionaries, Proc. IEEE, Vol. 88, No. 11, pp. 1756-1768 (2000).
- [7] Ota, T. and Morita, H.: On-Line Electrocardiogram Lossless Compression Using Antidictionary Codes for a Finite Alphabet, IEICE Trans. on Information and Systems, Vol. E93-D, No. 12, pp. 3384-3391 (2010).
- [8] Ota, T. and Morita, H. et. al: Real-Time and Memory-Efficient Arrhythmia Detection in ECG Monitors Using Antidictionary Coding, IEICE Trans. on Fundamentals, Vol. E96, No. 12, pp. 2343-2350 (2012).
- [9] 萩原竜太郎: 反辞書確率モデルを用いた監視カメラ映像における不審者検出, 修士論文, 電気通信大学 (2014).