

異常検知のための自己組織化ネットワークと その監視映像への適用

菅沼 雅徳^{1,2,a)} 長尾 智晴¹

概要: 本論文では、監視映像中における異常検知のための環境に適応的な自己組織化ネットワークを提案する。提案するネットワークでは、監視映像からの入力刺激に応じてネットワーク内のノードを生成、削除、またノードの更新を行うことで映像内の正常性を表現するネットワークを構築する。そして構築されたネットワーク内のノードの状態に基づき、映像内の正常度を算出することで異常検知を行う。実験では、固定カメラと一定の周期で視点が移動する巡回カメラを想定した2種類の監視映像を用いて、映像内に現れる歩行者と車両を異常として定義した侵入物体検知問題を扱い、提案するネットワークの性能評価を行った。

キーワード: 異常検知, 侵入物体検知, 監視カメラ, 動画像処理

A Self-Organising Network for Anomaly Detection and its Application to Video Surveillance

MASANORI SUGANUMA^{1,2,a)} TOMOHARU NAGAO¹

Abstract: In this paper, we propose a self-organising network for anomaly detection in surveillance videos. The network is used to model normal patterns in video scenes and detects deviations from them as abnormal patterns. The network can add nodes into its map and delete nodes from its map for representation of normal patterns more accurately. To demonstrate the effectiveness of our method, we have applied our method to intrusion detection tasks in surveillance videos obtained by a fixed camera and a pseudo rotating camera.

Keywords: anomaly detection, intrusion detection task, surveillance camera, video processing

1. はじめに

近年の監視カメラの普及に伴い、不審者検知や侵入物体検知などの防犯システムの要求が高まっている。その一方で、監視カメラ映像の多くは人間の目視によって確認されているため、膨大な労力が必要である。そのため、計算機を用いた映像内の人の異常動作や、映像内への侵入物体な

どの異常検知に関する研究が数多く行われている。異常検知の自動化が実現すれば、監視者の負担軽減や人為的ミスの抑制など多くの有用性が期待できる。

異常検知に関する先行研究では、映像内での出現頻度が少ないパターンや、正常パターンと特徴が大きく異なるパターンを異常と定義して、それらを検出する手法が多い [1], [2], [3]。そのため、多くの手法では事前に異常を含まない正常パターンから映像内の正常モデルを構築し、その正常モデルにおける生起確率が低いパターンや、モデルを逸脱するパターンを異常として検出している。本論文でも多くの先行研究と同様に、映像内での出現頻度が少ないパターンや、正常パターンから逸脱するパターンを異常と定義する。

¹ 横浜国立大学大学院環境情報学府
Graduate School of Environment and Information Sciences,
Yokohama National University, Yokohama, Kanagawa 240-
8501, Japan

² 日本学術振興会特別研究員 DC
JSPS Research Fellow

^{a)} suganuma-masanori-hf@ynu.jp

南里らは立体高次局所自己相関特徴 (Cubic High-order Local Auto-Correlation; CHLAC) 特徴 [4] を用いて、人物の歩行動作を正常動作、走行動作と転倒動作を異常動作と定義した実験において、複数の人物が映る動画像から異常動作を検出することに成功している [5]. Xiang らは屋内と屋外で撮影された動画像内での人物の行動パターンについてのモデルを学習し、学習データにはなかった行動パターンを異常として検出する手法を提案している [6]. また、近年では人が多く存在する混雑シーンにおける異常検知の研究 [7], [8], [9] も盛んに行われており、Mahadevan らは mixture of dynamic textures [10] を用いて時空間特徴に関する正常モデルを構築し、歩道の混雑シーンにおいて車両などの異常を検出することができることを示している [11]. これら多くの先行研究では事前に構築された正常モデルが固定であり、異常検知の適用中に更新されることはない。しかし、現実世界での運用を考えると、天候によって照明条件が変わることで正常パターンが変化したり、学習データには出現しなかった正常パターンが出現する場合など、事前に構築した固定の正常モデルだけでは正確に映像内の正常性を表現することが難しいと考えられる。したがって、異常検知をより頑健に行うには、適用中の環境に応じてモデルの更新が行われる環境に適応的な手法が望まれる。そこで本論文では、異常検知の適用中にモデルの更新が行われる手法を提案する。

環境に適応的な異常検知手法の例としては Masland らの Grow When Required (GWR) ネットワークが挙げられる [12], [13], [14]. GWR ネットワークでは頻繁に観測される入力刺激に対して、馴化モデルである Stanley モデル [15] を用いてネットワークの出力を次第に減少させていくことで環境の正常性を表現する。GWR ネットワークを移動ロボットに搭載し、ソナーセンサ情報を入力刺激として扱った実験では、ロボットがそれまで観測していた環境とは異なる環境に置かれるとネットワークが強い出力を示すことが確認されている。また、Nehmzow らは GWR ネットワークの入力刺激に画像情報を用いた異常検知実験を行っている [16], [17]. しかし、GWR ネットワークは単純な構造のため、環境に変動が含まれるような場合に対して適用することは難しいと考えられる。武田らは環境からの入力パターンに対する反応とそのパターンが生じる領域に対する反応の抑制によって、環境に変動を含む監視映像内の侵入物体を検知するネットワークモデルを提案している [18]. Staufer らは画素ごとに混合ガウス分布を用いて環境の背景のモデル化を行っている [19]. この手法では、混合ガウス分布を用いることで背景の揺らぎなどを多峰性の分布で表現しつつ、パラメータ更新によって環境変化への適応も実現している。これらの手法は固定カメラからの映像を用いた侵入物体検知問題に対して優れた性能を示している [18]. しかし、武田らの手法では映像内の位置関係を

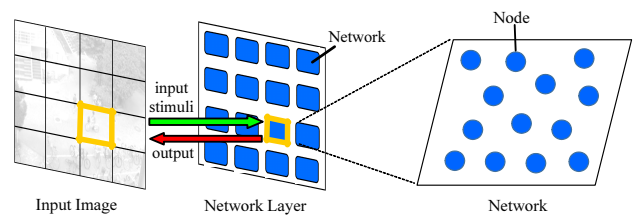


図 1: 提案手法の構造例

Fig. 1 The structure of the proposed method.

環境の正常性を表現するために利用していること、Staufer らの手法では正常性の表現力に乏しいことから、映像内の背景が周期的に変動する旋回カメラなどで撮影された映像に対して適用することは困難が予想される。旋回カメラによる異常検知が可能となれば、固定カメラより広範囲な監視を行うことができたり、決まったルートを巡回する警備ロボットへの応用が行えるなど、より多種類の場面で適用することが可能となりさらなる有用性が期待できる。

そこで本論文では、固定カメラと視点が移動する旋回カメラを想定した 2 種類の監視映像を対象に異常検知を行う、より環境への適応性が高い自己組織化ネットワークを提案する。提案するネットワークは、環境からの入力刺激に応じてノードの生成、削除、またノード状態の更新を行うことで環境の正常性を表現する。実験では固定カメラと一定周期で視点が移動する旋回カメラを想定した 2 種類の監視映像を用いて、映像内に現れる歩行者と車両を異常として定義した侵入物体検知問題に提案手法を適用し、従来手法と比較することで性能の検証を行う。

2. 自己組織化ネットワークによる異常検知

2.1 概要

提案する異常検知モデルの構造例を図 1 に示す。提案手法では、同一構造の自己組織化ネットワークが入力画像中に格子状に整列している。本論文では $m \times n$ 画素に対して 1 つのネットワークを配置しており、例えば入力画像サイズが $M \times N$ 画素の場合、入力画像中には $\frac{M}{m} \times \frac{N}{n}$ のネットワークが配置される。

各ネットワークは複数のノードによって構成される。ノード i は、重みベクトル \mathbf{v}_i 、時刻 t における馴化係数 $h_i(t)$ 、年齢 $age_i(t)$ をもち、重みベクトルは環境からの入力刺激と同次元数の $[0.0, 1.0]$ の実数値、馴化係数はノードがどの程度環境に馴化しているかを表す $[0.0, 1.0]$ の実数値、年齢はノードの年齢を表す非負の整数値である。

提案手法では、入力画像中の各格子領域から算出した画像統計量を入力刺激として対応する各ネットワークに入力する。そして、ネットワーク内でその入力刺激と類似した重みベクトルをもつノードの選択を行い、選択されたノードの年齢と馴化係数の更新を行う。さらに選択されたノードの入力刺激との類似度と馴化係数の値に応じて、ネット

ワーク内にノードを追加することで環境の正常性を表現する構造を構築していく。そして、ノードの類似度を用いて環境内の刺激に対する応答値をネットワークごとに非負の実数値で出力することで異常検知を行う。

2.2 処理の流れ

図 2 に提案手法の処理の流れを示す。処理の詳細については次の通りである。

- 1 全てのネットワークを次の手順で初期化する。
 - (a) ネットワークに I 個のノードを生成する。
 - (b) 各ノードの重みベクトルをあらかじめ算出した入力刺激からランダムに選択する。
 - (c) 各ノードの馴化係数 $h_i(0)$ を 1.0 にする。
 - (d) 各ノードの年齢 $age_i(0)$ を 0 にする。
- 2 全てのネットワークについて、入力動画像の 1 フレーム毎に以下の (a) から (h) の処理を繰り返す。
 - (a) ネットワークが配置された格子領域から入力刺激を算出し、ネットワークに入力する。入力刺激は格子領域内の画素値から算出した平均、中央値などの画像統計量を用いる。全ての入力刺激は $[0.0, 1.0]$ に正規化される。
 - (b) ネットワーク内のすべてのノードについて、入力刺激との類似度が大きい上位 S 個のノードを選択する。ノード i における類似度 D_i の算出には式 (1) を用いる。

$$D_i = \exp(-|\mathbf{x} - \mathbf{v}_i|) \quad (1)$$

ここで、 D_i はノード i の入力刺激 \mathbf{x} との類似度、 \mathbf{v}_i はノード i がもつ重みベクトルである。式 (1) は、類似度 D_i が大きいノードほど入力刺激と類似した重みベクトルをもっていることを示す。

- (c) 選択された S 個のノードの年齢 $age_i(t)$ を 0 にする。式 (2) における i は選択された S 個のノードのインデックスを表す。

$$age_i(t) = 0 \quad (2)$$

- (d) 各格子領域に対するネットワーク l の出力値 O^l は選択された S 個のノードの類似度を用いて、式 (3) によって算出される。

$$O^l = \begin{cases} D_{c(1)} & (S = 1) \\ \sum_{i=1}^{S-1} D_{c(i)} D_{c(i+1)} & (otherwise) \end{cases} \quad (3)$$

$c(i)$ は現フレームで選択された上位 S 個のノードのインデックスのうち、 i 番目に類似度が大きいノードのインデックスである。本論文では映像内における出現頻度が少ない刺激や特徴が大きく異なる刺激を異常として定義しているため、出力値 O^l がしきい値 O_{thr} より小さい場合、その格子領域を異常であると

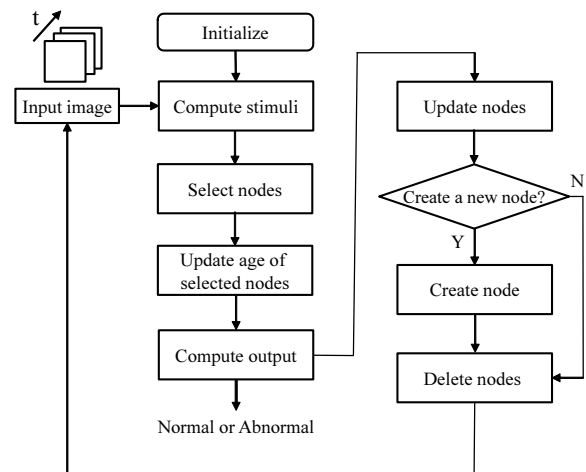


図 2: 提案手法の処理の流れ

Fig. 2 Flowchart of the proposed method.

判定する。

- (e) ネットワーク内で最大の類似度 D_{max} であるノード \max の馴化係数 $h_{max}(t)$ を式 (4) によって更新する。なお、 $\max = \arg \max(D_i)$ 、 γ は定数である。

$$h_{max}(t+1) = h_{max}(t) - \gamma D_{max} \quad (4)$$

ここで、馴化係数 $h(t)$ はそのノードが環境の入力刺激に対してどの程度馴化しているかを表す指標であり、処理ステップ (g) において新たなノードをネットワークに追加するか否かの指標となる。馴化係数の更新式 (4) には入力刺激との類似度を用いており、環境からの入力刺激と類似した重みベクトルをもつノードの馴化係数は 0 に近づいていく。

- (f) 現フレームで選択された上位 S 個以外のノードの age_i を式 (5) によって更新する。

$$age_i(t+1) = age_i(t) + 1 \quad (5)$$

- (g) 式 (6) を満たす場合、新たなノード k をネットワークに追加する。

$$\begin{cases} D_{max} < D_{thr} \\ h_{max} < H_{thr} \end{cases} \quad (6)$$

このとき、ノード k の重みベクトルは入力刺激 \mathbf{x} とノード \max の重みベクトル \mathbf{v}_{max} との平均値とし、馴化係数の初期値 $h_k(0)$ は 1.0、年齢の初期値 $age_k(0)$ は 0 とする。ネットワークが式 (6) を満たすことは、ある程度ネットワークの更新が行われているにもかかわらず環境からの入力刺激を表現するノードがネットワーク内に存在しないことを示している。このことから、環境の刺激をより正確に表現するために、環境の入力刺激を利用して新たなノードをネットワークに追加する操作を行っている。

- (h) age_i の値がしきい値 age_{thr} より大きい場合、ノード i をネットワークから削除する。

3. 固定カメラと旋回カメラを想定した映像からの侵入物体検知実験

3.1 概要

提案する自己組織化ネットワークの有効性を検証するため、屋外監視映像内に現れる歩行者と車両を異常として定義した侵入物体検知問題に提案手法を適用し、先行研究との比較を行う。対象とした屋外監視映像は、固定カメラから撮影された監視映像と一定周期で視点が移動する旋回カメラを想定した2種類の監視映像である。比較に用いた手法は、文献[12]で提案されたGWRネットワーク（以下、GWR）、Staufferらの手法[19]に対して直近の入力フレームを重視した更新を行う混合ガウス分布を用いた背景モデルの手法[20]（以下、MOG）を用いた。

各実験で用いた提案手法のパラメータを表1に示す。これらのパラメータは次節3.2で説明するF値が最大となるように、事前に行った実験によって決定した。提案手法とGWRの入力刺激には、格子領域内の画素値の平均、最大値、最小値、レンジ、中央値、第一四分位数、第三四分位数の7種類の統計量をRGBカラー画像、RGBエッジ画像のそれぞれから算出した計42次元の特徴量を用いた。

3.2 評価方法

まず、異常検知性能の定量評価を行うために、監視映像内に現れる歩行者と車両を対象に正解画像を作成した。さらに、提案手法とGWRでは画像内の格子領域単位の出力であるため、作成した正解画像から正解格子画像を作成した。正解格子画像は、正解画像を提案手法およびGWRと同様のサイズの格子領域に分割し、格子領域内に占める正解画素の割合が0.1以上の格子領域を正解格子領域とすることで作成した。1つの格子領域サイズが10×10画素とした場合の正解格子画像の例を図3b、図4bに示す。また、画素単位で出力するMOGと評価方法を揃えるために、MOGによる検出画像を同じサイズの格子領域で分割し、格子領域内に占める検出画素の割合が同じく0.1以上の格子領域をMOGによる検出格子領域と定義する。

本論文では作成した正解格子画像をもとに、式(7)で示される再現率(R)、適合率(P)、F値(F)の指標を用いて定量評価を行う。なお、過剰な過検出結果をノイズとして除外するために、物体の侵入によって生じた過検出格子領域は検出格子領域から除外した。

$$R = \frac{C}{C_a}, \quad P = \frac{C}{A}, \quad F = \frac{2 \cdot R \cdot P}{R + P} \quad (7)$$

C は正しく検出した格子領域数、 C_a は正解格子領域数、 A は検出した格子領域数である。

表 1: 提案手法に関するパラメータ
Table 1 Parameter settings for the proposed method.

パラメータ	固定	旋回
初期化時のノード数 I		20
ノードの選択数 S		1
馴化係数の更新係数 γ		0.1
ノード追加の馴化係数しきい値 H_{thr}		0.1
最大年齢 age_{thr}		800
ノード追加の類似度しきい値 (モデル適用時) D_{thr}		0.1
ノード追加の類似度しきい値 (モデル構築時) D_{thr}	0.6	0.925
出力値のしきい値 O_{thr}	0.76	0.7

3.3 固定カメラからの監視映像による侵入物体検知

3.3.1 実験設定

固定カメラからの監視映像には、PETS2001 データセット中の DATASET2 TESTING CAMERA1 を使用した。用いた画像は320×240画素のカラー画像である。提案手法とGWRでは10×10画素に1つのネットワークを配置したため、画像サイズからネットワーク数は768である。この動画像では手前の樹木が風によって揺れており、また背景にちらつきなどのノイズがみられる。固定カメラ映像の例を図3aに示す。

歩行者と車両が現れない開始から300フレームまでを提案手法と比較手法のモデル構築期間とし、301フレームから1800フレームまでをモデル適用期間とする。なお、正解格子画像は301フレームから1800フレームまで10フレームおきに作成し、計150フレーム分を用意した。

3.3.2 実験結果

図3に各手法による検出結果例を示す。なお、MOGは格子領域単位に変換した結果を示している。図3の結果から、提案手法では手前の揺れている樹木領域への過検出を抑制しつつ、歩行者を検出できていることがわかる。MOGでは樹木領域で生じる多様な入力パターンを適切にモデル化できなかったため、樹木領域に対して過検出してしまっている。GWRは提案手法と同様に樹木領域に対して過検出を抑制できているが、歩行者に対して検出漏れをしてしまう場合が多くみられた。

再現率、適合率、F値による定量評価結果を表2に示す。なお、正解格子領域数は正解画像150フレーム分に対して1732である。表2から提案手法はGWRと比べて再現率、適合率において優れており、MOGと比べると適合率で同等程度の性能を示していることがわかる。しかし、再現率において提案手法はMOGより低い値となっている。これは、背景と類似した色情報をもつ歩行者領域に対して検出漏れをしてしまっていることが原因だと考えられる。提案手法では映像内の色情報をもとに映像内の正常性を記述するため、背景の色情報と類似している侵入物体に対しては検出が抑制されてしまう。これは今後、入力刺激に時空間

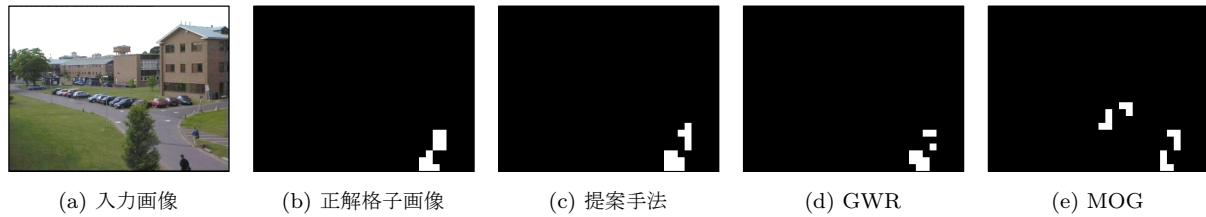


図 3: 各手法の検出結果例 (固定カメラ映像)

Fig. 3 Detection results of each method in the fixed camera.

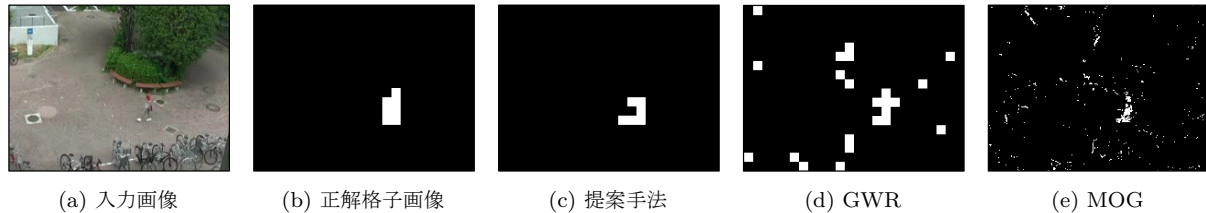


図 4: 各手法の検出結果例 (旋回カメラ映像)

Fig. 4 Detection results of each method in the rotating camera.

表 2: 検出結果に対する定量評価 (固定カメラ映像)

Table 2 Recall, Precision and F-measure on the fixed camera.

	Proposed	GWR	MOG
再現率	0.838	0.656	0.897
適合率	0.868	0.758	0.874
F 値	0.853	0.703	0.885

表 3: 検出結果に対する定量評価 (旋回カメラ映像)

Table 3 Recall, Precision and F-measure on the rotating camera.

	Proposed	GWR	MOG
再現率	0.806	0.553	0.567
適合率	0.825	0.639	0.684
F 値	0.815	0.593	0.620

特徴を使用し、入力刺激の時間的な変化も考慮することで解決が可能であると考えられる。

3.4 旋回カメラを想定した映像による侵入物体検知

3.4.1 実験設定

本論文では、固定カメラから撮影された映像を用いて擬似的な旋回カメラ映像を作成した。具体的には、一定速度のウィンドウによって固定カメラ映像内から画像を切り出すことで、擬似的な旋回カメラ映像（以下、旋回カメラ映像）を用意した。今回は 320×240 画素の動画像から 240×180 画素の動画像を作成した。対象とした環境では映像内の左右上下から歩行者が映像内に出現する。図 4a に旋回カメラ映像の例を示す。

3.3 の実験と同様に提案手法と GWR では 10×10 画素に 1 つのネットワークを配置したため、画像サイズからネットワーク数は 432 である。歩行者が出現しない開始から 1000 フレームまでを各手法のモデル構築期間とし、歩行者が出現する 1001 フレームから 2000 フレームまでをモデル適用期間とする。正解格子画像は 1001 フレームから 2000 フレームまで 10 フレームおきに作成し、計 100 フレーム分を用意した。

3.4.2 実験結果

図 4 に各手法における検出結果例を示す。図 4 の結果から、提案手法では変動する背景領域への過検出を抑制し

つつ、侵入物体である歩行者を検出できていることがわかる。一方で、GWR と MOG では変動する背景領域に対して過検出するケースが多くみられた。これは変動する背景の多様な状態を適切にモデル化できなかったためだと考えられる。

次に再現率、適合率、F 値による定量評価結果を表 3 に示す。なお、正解格子領域数は正解画像 100 フレーム分に対して 566 である。表 3 から提案手法は再現率、適合率、F 値において他の 2 手法と比べて高い値を示していることがわかる。再現率は比較手法と比べて約 24% 以上、適合率は約 14% 以上優れている。以上の結果から、提案手法は旋回カメラ映像において比較手法より高い異常検知性能を示すことが確認された。

3.5 考察

固定カメラでの実験において、提案手法は図 3 に示したように風によって揺れている樹木領域への過検出を抑制しつつ、侵入物体を検出できている。このことを定量的に評価するため、手前の樹木領域における過検出数を 301 フレームから 1800 フレームの 1500 フレーム分について調べた。その結果、過検出数が提案手法では 179、GWR では 301、MOG では 402 となった。このことから、提案手法は他の 2 手法と比べて変動が生じる領域への過検出を抑制で

きていることがわかる。固定カメラ映像における樹木領域への過検出の抑制結果と巡回カメラ映像における評価結果から、提案手法は他の2手法と比べてより環境の変化に頑健に異常検知を行うことができることが確認された。

このように提案手法が環境の変化に対して頑健に異常検知を行えるのは、環境からの入力刺激との類似度を考慮して環境の正常性を表現している点にある。まず、提案手法では入力刺激と類似しているノードがネットワーク内に存在しない場合、入力刺激と類似したノードをネットワーク内に追加することで環境の変化に適応する。固定カメラ映像における樹木領域や、巡回カメラ映像における背景領域のように様々な入力刺激が観測される環境では、このノード追加の操作によって環境の状態を表現するネットワークを構築していく。次に、提案手法ではこの構築されたネットワーク内のノードと入力刺激との類似度を用いて出力値を算出するため、環境内で特徴の異なる入力刺激に対して出力値の正常度を抑制することができる。GWRでは入力刺激との類似度ではなく、出現頻度を考慮して異常判定を行うため、モデルの入力刺激に対する表現力が弱いといえる。またMOGについては、変動領域で観測される様々な入力パターンを背景として適切にモデル化できなかったことが過検出結果の原因だと考えられる。

4. まとめ

本論文では、環境から入力される刺激に基づいて環境の正常性を表現し、異常検知を行うネットワークモデルを提案した。提案手法では、環境からの入力刺激との類似度に応じてノードの追加、削除、状態の更新、また出力値を算出することで環境の正常性を表現する。提案手法を固定カメラと巡回カメラを想定した2種類の監視映像からの侵入物体検知問題に適用し、先行研究と比較を行った結果、環境内の日常的な変動や周期的な変動がある領域への過検出を抑制しつつ、異常である侵入物体を検知できることが確認された。

今後の課題として、長時間の監視映像への適用実験と、その実験を通じての精度向上が挙げられる。さらに、提案手法の拡張として人物の異常動作検知や、混雑シーンにおける適用などを検討したい。

謝辞 本研究は、株式会社半導体理工学研究センター(STARC)の助成を受けて行われたものである。

参考文献

- [1] Adam, A., Rivlin, E., Shimshoni, I. and Reinitz, D.: Robust real-time unusual event detection using multiple fixed-location monitors, *IEEE Trans. PAMI*, Vol. 30, No. 3, pp. 555–560 (2008).
- [2] Basharat, A., Gritai, A. and Shah, M.: Learning object motion patterns for anomaly detection and improved object detection, *Proc. IEEE International Conference on CVPR*, pp. 1–8 (2008).
- [3] Zhong, H., Shi, J. and Visontai, M.: Detecting unusual activity in video, *Proc. IEEE International Conference on CVPR*, Vol. 2, pp. 819–826 (2004).
- [4] Kobayashi, T. and Otsu, N.: Action and simultaneous multiple-person identification using cubic higher-order local auto-correlation, *Proc. the 17th International Conference on Pattern Recognition*, Vol. 4, pp. 741–744 (2004).
- [5] 南里卓也, 大津展之: 複数人動画からの異常動作検出, 情処学論, コンピュータビジョンとイメージメディア, Vol. 46, No. 15, pp. 43–50 (2005).
- [6] Xiang, T. and Gong, S.: Video Behavior Profiling for Anomaly Detection, *IEEE Trans. PAMI*, Vol. 30, No. 5, pp. 893–908 (2008).
- [7] Mehran, R., Oyama, A. and Shah, M.: Abnormal crowd behavior detection using social force model, *Proc. IEEE International Conference on CVPR*, pp. 935–942 (2009).
- [8] Saxena, S., Brémond, F., Thonnat, M. and Ma, R.: Crowd behavior recognition for video surveillance, *Advanced Concepts for Intelligent Vision Systems*, pp. 970–981 (2008).
- [9] Saligrama, V. and Chen, Z.: Video anomaly detection based on local statistical aggregates, *Proc. IEEE Conference on CVPR*, pp. 2112–2119 (2012).
- [10] Chan, A. B. and Vasconcelos, N.: Modeling, clustering, and segmenting video with mixtures of dynamic textures, *IEEE Trans. PAMI*, Vol. 30, No. 5, pp. 909–926 (2008).
- [11] Mahadevan, V., Li, W., Bhalodia, V. and Vasconcelos, N.: Anomaly detection in crowded scenes, *Proc. IEEE International Conference on CVPR*, pp. 1975–1981 (2010).
- [12] Marsland, S., Shapiro, J. and Nehmzow, U.: A self-organising network that grows when required, *Neural Networks*, Vol. 15, No. 8, pp. 1041–1058 (2002).
- [13] Marsland, S., Nehmzow, U. and Shapiro, J.: Detecting novel features of an environment using habituation, *Proc. Simulation of Adaptive Behavior*, pp. 189–198 (2000).
- [14] Marsland, S., Nehmzow, U. and Shapiro, J.: Environment-specific novelty detection, *From Animals to Animats, the 7th International Conference on Simulation of Adaptive Behaviour*, pp. 36–45 (2002).
- [15] Stanley, J. C.: Computer simulation of a model of habituation., *Nature*, Vol. 261, pp. 146–147 (1976).
- [16] Nehmzow, U. and Neto, H. V.: Novelty-based visual inspection using mobile robots, *Proc. TAROS* (2004).
- [17] Nehmzow, U. and Neto, H. V.: Visual attention and novelty detection: Experiments with automatic scale selection, *Proc. TAROS*, pp. 139–146 (2006).
- [18] 武田真人, 矢田紀子, 長尾智晴: 映像監視のための環境に適応的な異常検知ネットワーク, 電子情報通信学会論文誌, Vol. J94-D, No. 10, pp. 1631–1639 (2011).
- [19] Stauffer, C. and Grimson, W. E. L.: Adaptive background mixture models for real-time tracking, *Proc. IEEE International Conference on CVPR*, Vol. 2 (1999).
- [20] KadewTraKuPong, P. and Bowden, R.: An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection, *Proc. 2nd European Workshop on Advanced Video-Based Surveillance Systems* (2001).