

時空間特徴に基づくエスカレータシーンにおける人の異常行動検知

村井 泰裕† 藤吉 弘亘† 数井 誠人††

† 中部大学大学院 工学研究科 情報工学専攻
†† 株式会社 日立製作所 日立研究所

E-mail:†yasu@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp, ††masato.kazui.bq@hitachi.com

あらまし 本稿では、画像中の物体のアピアランスとモーションの情報を持った時空間特徴を用いて、動的な背景であるエスカレータシーンにおける人の異常行動検知を行う手法を提案する。提案手法は、混合正規分布による動的背景モデリング、時空間特徴の RealAdaboost に基づく人領域検出、重み付き時空間勾配を用いた動きベクトルの定常度算出の3つのステップから構成されている。提案手法は、立体高次局所自己相関特徴を用いた従来法では検知が困難である動的背景シーンにおいても、人の異常行動検知が可能となる。評価実験より、エスカレータシーンに対し、提案手法では従来法と比較して誤検出率10%のとき検出率を約27%向上させることができ、その有効性を確認した。

Abnormal Motion Detection at Escalator Scene based on Spatio-temporal Features

Yasuhiro Murai† Hironobu Fujiyoshi† Masato Kazui††

†Department of Computer Science, Chubu University
††Hitachi, Ltd., Hitachi Research Laboratory

E-mail:†yasu@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp, ††masato.kazui.bq@hitachi.com

Abstract This paper presents a method for detecting abnormal motion at an escalator scene, in which the movement of the escalator provides a dynamic background. This method is based on the use of spatio-temporal features obtained by space-time patches. Our approach consists of three steps; dynamic background modeling by using the Gaussian mixture model, human region detection based on RealAdaboost, and calculation of irregularity measure by using weighted space-time gradients. The proposed method can detect abnormal motions from a scene with a dynamic background, that would be difficult to detect with the conventional method using CHLAC (Cubic Higher-order Local Auto-Correlation) features. Our experimental results show that using our method has about 27% higher performance than that of the conventional method.

1 はじめに

近年、大規模な監視カメラ導入とともに、物体検出や追跡などの動画理解技術によるセキュリティシステムの要求が高まりつつある。最近では、これらの動画理解技術を用い、人の不審な行動・転倒等の異常行動検知に関する研究が数多く行われている。人の行動を認識し、異常行動を自動的に検知するセキュリティシステムによって、監視者の負担の

軽減、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が期待できる。

人の異常行動検知を目的とした最も代表的な手法として、南里らにより提案されている立体高次局所自己相関 (Cubic Higher-order Local Auto-Correlation: CHLAC) 特徴と部分空間法を組み合わせた異常行動検知手法 [1] が挙げられる。南里らの手法では、フレーム間差分により検出した時系列2値化画像に対して変位パターンを当てはめ、時間

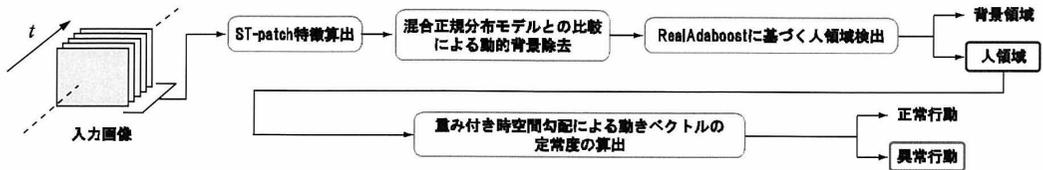


図 1: 提案手法の流れ

方向に積算した特徴量である CHLAC 特徴により、正常行動の部分空間を構築する。そして、構築した部分空間からどれだけ逸脱しているかを表す値である射影成分距離が閾値以上であるものを異常行動として検知する。これにより、個々の人物の切り出しやトラッキング無しにロバストな異常行動検知を可能にしている。また、この CHLAC 特徴は、各種のイベントシーンから CHLAC 特徴を算出し、判別分析法によってイベント検出を行う手法 [2] にも用いられている。

関らは、エレベータ内の限定空間において、算出したフローベクトルの向きと大きさの分散、及び個数からなる統計量を評価することで人の暴れ検知をリアルタイムに行う手法 [3] を提案している。関らの手法では、異常行動の指標としている特徴量を画像全体から求めるため、異常行動が画像中の何処で発生しているかを特定することができないという問題点が挙げられる。このような問題に対し都築らは、SIFT 特徴量を用いた Mean-Shift 探索により得られた特徴点追跡結果のフローを用いて移動方向モデルを作成し、モデルに逸脱するフローを異常行動として検知する手法 [4] を提案している。

CHLAC 特徴やフローベクトルを用いた従来法における問題として、背景が動的である場合、人の異常行動検知が非常に困難になるという点である。関らや都築らの手法では、フローの情報のみを使用して異常行動検知を行うため、人以外のフローベクトルが観測された場合に対応することができない。また、南里らの手法で用いている CHLAC 特徴は、物体の「動き」の「形状」を表す特徴量であるため、動的背景からも特徴量が抽出されてしまう。そのため、人のみを捉えた異常行動検知を行うことが困難である。

そこで本稿では、画像中の物体の「アピアランス（見え）」と「モーション（動き）」の 2 つの情報を持った時空間特徴を用いて、動的な背景を持つエスカレーターシーンより人の異常行動検知を行う手法を提案する。本稿では、2 章にて提案手法で用いる時

空間特徴である Space-Time Patch 特徴について述べ、3 章にて Space-Time Patch 特徴を用いた人の異常行動検知について述べる。そして、4 章にて評価実験を行い、提案手法の有効性を示す。

2 Space-Time Patch 特徴

提案手法で用いる Space-Time Patch (ST-patch) 特徴とは、Shechtman らにより提案されている物体の局所的な「アピアランス」と「モーション」の時間的変化を捉えた特徴 [5] である。この ST-patch 特徴はイベント検出 [6] や移動方向を考慮した物体検出とセグメンテーション手法 [7] など、様々な手法に応用されている。

ST-patch 特徴は、画像を時間方向に重ねた時空間画像の局所領域 (e.g, $20 \times 20[\text{pixel}] \times 5[\text{frame}]$) における x 軸, y 軸, t 軸の勾配値より求められる。ある画素 i からは、 P_{x_i} , P_{y_i} , P_{t_i} の 3 つの勾配値を持つ時空間勾配が得られる。この時空間特徴を式 (1) のように、ある局所領域の画素数 n でスタックする。

$$\begin{bmatrix} P_{x_1} & P_{y_1} & P_{t_1} \\ P_{x_2} & P_{y_2} & P_{t_2} \\ \vdots & \vdots & \vdots \\ P_{x_n} & P_{y_n} & P_{t_n} \end{bmatrix}_{n \times 3} \quad (1)$$

この時空間勾配値からなる $n \times 3$ の行列を \mathbf{G} とし、行列 \mathbf{G}^T を掛ける。そして、行列 $\mathbf{G}^T \mathbf{G}$ を \mathbf{M} とすると、行列 \mathbf{M} は式 (2) で表すことができる。

$$\mathbf{M} = \mathbf{G}^T \mathbf{G} = \begin{bmatrix} \sum P_x^2 & \sum P_x P_y & \sum P_x P_t \\ \sum P_y P_x & \sum P_y^2 & \sum P_y P_t \\ \sum P_t P_x & \sum P_t P_y & \sum P_t^2 \end{bmatrix} \quad (2)$$

式 (2) より求められる行列 \mathbf{M} が 1 つの ST-patch から得られる特徴量となる。行列 \mathbf{M} には、左上の 2×2 の行列に「アピアランス」の情報、3 行目と 3 列目には「モーション」の情報を持つ。この時、ST-patch 特徴は対称成分が同一値のベクトルとなっているため、提案手法ではこの ST-patch 特徴を式 (3) の 6 次元ベクトル \mathbf{v} として扱う。

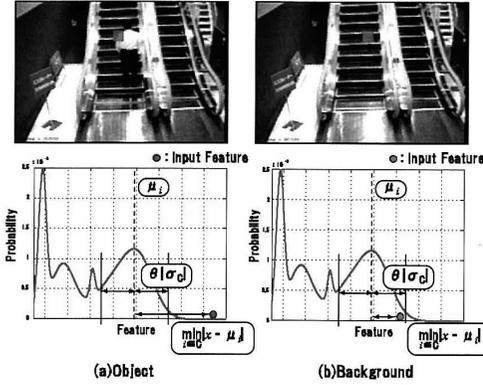


図 2: 混合正規分布モデルとの比較方法

$$\begin{aligned} \mathbf{v} &= \left(\sum P_x^2, \sum P_x P_y, \dots, \sum P_t^2 \right) \\ &= (v_1, v_2, \dots, v_6) \end{aligned} \quad (3)$$

そして、6次元の特徴ベクトル \mathbf{v} を式(4)により、ベクトルの総和で正規化する。

$$\mathbf{v}_N = \frac{v}{\left(\sum_{i=1}^6 |v_i| + \epsilon \right)} \quad (4)$$

提案手法では、この6次元の正規化特徴ベクトル \mathbf{v}_N を時空間特徴として用いる。

3 ST-patch 特徴を用いた異常行動検知

人の異常検知を行うためには、対象物体の動きの時間的変化を観測し、行動の認識を行う必要がある。しかし、背景が動的な場合、対象物体の動きの特徴を得ることが困難となる。そこで提案手法では、動的な背景を持つシーンに存在する人の異常行動検知を実現するため、まず動的背景における時空間特徴をモデリングすることにより、動的背景除去を行い、物体領域のみを検出する。次に、検出された物体領域から RealAdaboost により構築した識別器に基づき、人領域の検出を行う。そして、検出された人領域より得られる動きベクトルの定常度により、異常行動の検出を行う。図 1 に提案手法の流れを示す。本章では、人の異常行動検知のフレームワークである混合正規分布による動的背景モデリング、RealAdaboost に基づく人領域検出、動きベクトルの定常度による異常行動検知方法について述べる。

3.1 混合正規分布による動的背景モデリング

混合正規分布とは、用意された複数の正規分布を組み合わせた確率モデルである [8]。データ \mathbf{x} の確率



図 3: 動的背景モデリングによる物体領域検出例

密度分布 $p(\mathbf{x})$ を C 個のクラスを持つ確率密度分布 $p(\mathbf{x}|i)_{i=1, \dots, C}$ の重み付き線形結合とすると、 $p(\mathbf{x})$ は式 (5) で表される。

$$p(\mathbf{x}) = \sum_{i=1}^C \alpha_i p_i(\mathbf{x}|C) \quad (5)$$

$$p_i(\mathbf{x}|C) = \frac{1}{\sqrt{(2\pi)^M |\sigma_i|}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mu_i)^T \sigma_i^{-1} (\mathbf{x} - \mu_i) \right\} \quad (6)$$

この時、 M は次元数、 α_i は重み、 μ_i は平均、 σ_i は共分散行列を表す。このように、混合正規分布では各正規分布ごとに α_i 、 μ_i 、 σ_i の 3 つのパラメータが必要となるため、各パラメータを EM アルゴリズムにより推定する。

提案手法では、6次元ベクトルとして表される時空間特徴 \mathbf{v}_N を混合正規分布を当てはめることで動的背景のモデリングを行う。この確率密度分布から、式 (7) の判定式によって物体/背景領域の判別を行う。

$$\min_{i \in C} |\mathbf{x} - \mu_i| > \theta |\sigma_i| \quad (\theta > 0) \quad (7)$$

ここで、 \mathbf{x} は入力特徴、 θ は閾値を表す。この時、閾値 θ の値が大きいくほど、物体検出判定が厳しくなる。この式 (7) を指定領域に適用し、観測された特徴量が判定式を満たした場合、物体領域として検出する。図 2 に混合正規分布モデルとの比較方法を示す。図 2(a) は入力特徴量が物体であった場合、図 2(b) は背景であった場合を表す。

図 2 に示す混合正規分布モデルとの比較を格子状に配置した ST-patch に適用することで、動的背景の除去が可能となる。図 3 に ST-patch サイズを $20 \times 20[\text{pixel}] \times 5[\text{frame}]$ 、配置する ST-patch 間隔を $5[\text{pixel}]$ 、混合正規分布のクラス数を 5 としたときの動的背景モデリングによる物体領域検出例を示す。このとき、動的背景の画像 $3000[\text{frame}]$ を学習データとして使用した。

3.2 RealAdaboostに基づく人領域検出

時空間特徴のモデリングによる物体領域検出では、個々の ST-patch が物体/背景であるかの判別を行っているため、人の領域がどこに存在しているかをすることはできない。そこで提案手法では、人領域から得られる時空間特徴から RealAdaboost[9]により識別器を構築し、人領域検出を行う。

RealAdaboost とは、Positive クラスの特徴量と Negative クラスの特徴量における各次元の確率密度分布から類似度を求め、最も識別に有効である特徴量を学習する手法である。提案手法では、Positive クラスとして正常行動時の人領域、それ以外の領域を Negative クラスとして用いる。

ST-patch サイズを $20 \times 20[\text{pixel}] \times 5[\text{frame}]$ 、配置する ST-patch 間隔を $5[\text{pixel}]$ 、検出ウィンドウのサイズを $140 \times 260[\text{pixel}]$ (故に、ウィンドウ内に存在する ST-patch は 1152 個) としたとき、検出ウィンドウの領域より得られる時空間特徴は $1152 \times 6 = 6912$ 次元となる。この 6912 次元の特徴量を RealAdaboost により学習し、識別器を構築する。ここで、RealAdaboost による学習過程では、先の手法により背景除去された画像より得られる特徴量を使用し、背景として判別された領域の特徴量は 0 として扱う。これにより、背景の影響を受けることなく特徴量の学習を行うことが可能であり、検出精度の向上が期待できる。

構築された識別器を用いて人領域検出をする際には、動的背景除去画像に対して検出ウィンドウのラスト走査を行い、検出ウィンドウ内の特徴量が人領域/非人領域であるかの判別を行う。そして、人領域として検出されたウィンドウに対しては、Mean-Shift クラスタリング [10] によりウィンドウの統合処理を行う。

3.3 動きベクトルの定常度による異常行動検知

通常、転倒による異常行動は定常状態と比較し、動きベクトルが大きく変化する傾向がある。そこで正常/異常行動の判定指標として、ST-patch より求められる CRIM (Continues Rank-Increase Measure) [5] を用いる。

CRIM とは、ST-patch 内に存在する動きベクトルの定常度を表す値である。図 4 に正常/異常行動における CRIM の性質を表現した例を示す。図 4(a) に示す正常行動時には、CRIM は低い値となり、図

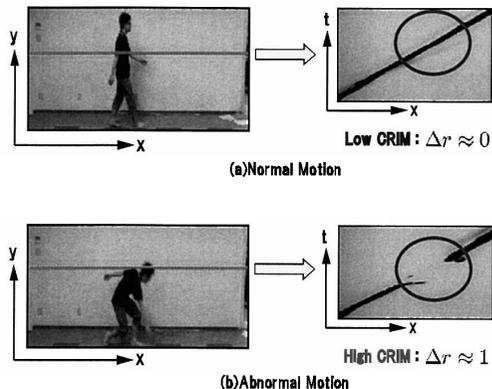


図 4: CRIM の性質

4(b) に示す異常行動時、即ち転倒による動きベクトルの変化が発生した時には、高い値となる性質がある。提案手法では、検出された人領域を 1 つの ST-patch とみなして算出した CRIM の値を異常行動検知の指標とする。

3.3.1 CRIM の算出

ST-patch 特徴を表す式 (2) の行列 \mathbf{M} と、式 (8) の行列 \mathbf{M}^\diamond より求められる固有値を用いて $\text{CRIM} \Delta r$ を算出する。

$$\mathbf{M}^\diamond = \begin{bmatrix} \sum P_x^2 & \sum P_x P_y \\ \sum P_y P_x & \sum P_y^2 \end{bmatrix} \quad (8)$$

$\lambda_1 \geq \lambda_2 \geq \lambda_3$ を行列 \mathbf{M} の固有値、 $\lambda_1^\diamond \geq \lambda_2^\diamond$ を行列 \mathbf{M}^\diamond の固有値とする。行列 \mathbf{M} 、行列 \mathbf{M}^\diamond は共に対称行列であり、行列 \mathbf{M} 、行列 \mathbf{M}^\diamond の固有値は全て実数となる。また、それぞれの固有値の関係は $\lambda_1 \geq \lambda_1^\diamond \geq \lambda_2 \geq \lambda_2^\diamond \geq \lambda_3$ となる。このような関係から、 Δr を式 (9) のように定義する。

$$\Delta r = \frac{\lambda_2 \cdot \lambda_3}{\lambda_1^\diamond \cdot \lambda_2^\diamond} \quad (0 \leq \Delta r \leq 1) \quad (9)$$

ここで求められた Δr の値は、ST-patch 内の動きベクトルが一定のとき $\Delta r \approx 0$ となり、動きベクトルの変化が激しいとき $\Delta r \approx 1$ となる。これにより、ST-patch 内の動きベクトルの定常度を求めることができる。

3.3.2 重み付き時空間勾配を用いた CRIM の算出

RealAdaboost では、特徴量の値からヒストグラムのビンを算出することで確率密度分布を表現し、その確率密度分布を LUT (Look-Up Table) として弱識別器を構成している。これにより、RealAdaboost

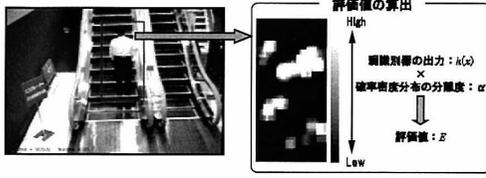


図 5: 評価値

の弱識別器の出力値から、入力された特徴が物体であるかを表す信頼度を連続値で得ることができる。提案手法では、RealAdaboost より得られる値に着目し、重み付き時空間勾配を用いて CRIM の算出を行うことにより、異常行動検知性能の向上を図る。

RealAdaboost に基づく人領域検出時には、各弱識別器より弱識別器出力、確率密度分布の分離度の 2 つの値を得ることができる。弱識別器出力 $h(x)$ は、RealAdaboost を用いて作成した Positive クラスと Negative クラスの確率密度分布 W_+^j , W_-^j より算出される。入力された ST-patch 特徴を x としたとき、その特徴量の値に対応した 1 次元ヒストグラムのビン番号 j の確率密度分布 W_+^j , W_-^j から弱識別器出力を式 (10) により算出する。

$$h(x) = \frac{1}{2} \ln \frac{W_+^j + \epsilon}{W_-^j + \epsilon} \quad (10)$$

ここで、 ϵ は分母が 0 になることを防ぐための係数である。

次に、確率密度分布 W_+^j , W_-^j を用いて、分布の分離度 α を式 (11) により算出する。

$$\alpha = 1 - \sum_j \sqrt{W_+^j W_-^j} \quad (11)$$

この分布の分離度の値が大きいくほど Positive サンプルと Negative サンプルの分布が分離していることを意味する。

算出した弱識別器出力と確率密度分布の分離度を用いて、評価値 E を算出する。評価値には、弱識別器出力と確率密度分布の分離度を掛けた値を用いる。また、提案手法では、評価値を $h(x) > 0$ の時、即ち Positive クラスであると出力された値のみを用いる。よって、評価値は式 (12) で表現される。

$$E = \begin{cases} Ch(x)\alpha & \text{if } h(x) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

ここで、 C は定数を表す。 C の値が高いほど、先に述べた CRIM が評価値に対する依存性が高くなる。

評価値を算出することにより、図 5 に示す分布を得ることができる。得られた分布より、物体の形状

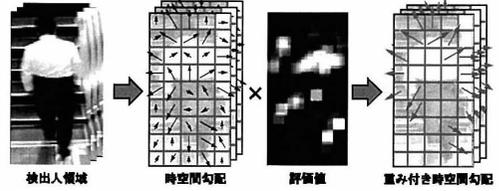


図 6: 重み付き時空間勾配を用いた CRIM 算出

を最も表す領域に高い値が得られていることが分かる。このように、評価値により物体形状を表現した局所的な情報を得ることが可能となる。

提案手法では、この評価値を重みとし、時空間勾配に掛け合わせることにより、重み付き時空間勾配を算出する。ある画素 i からは、 P_{x_i} , P_{y_i} , P_{t_i} の 3 つの勾配値を持つ時空間勾配が得られる。ここで、画素 i における評価値 E_i を掛け合わせた特徴量を式 (13) のように、画素数 n でスタックする。

$$\begin{bmatrix} E_1 P_{x_1} & E_1 P_{y_1} & E_1 P_{t_1} \\ E_2 P_{x_2} & E_2 P_{y_2} & E_2 P_{t_2} \\ \vdots & \vdots & \vdots \\ E_n P_{x_n} & E_n P_{y_n} & E_n P_{t_n} \end{bmatrix}_{n \times 3} \quad (13)$$

このように、評価値によって重み付けされた時空間勾配からなる $n \times 3$ の行列より、先に述べた手順で CRIM の算出を行う。図 6 に重み付き時空間勾配を用いた CRIM 算出の例を示す。

4 評価実験

提案手法の有効性を評価するため実験を行う。本章では、まず、実験概要について述べた後、各手法による異常行動検知結果、異常行動検知性能評価について示す。そして、得られた結果より提案手法の有効性について述べる。

4.1 実験概要

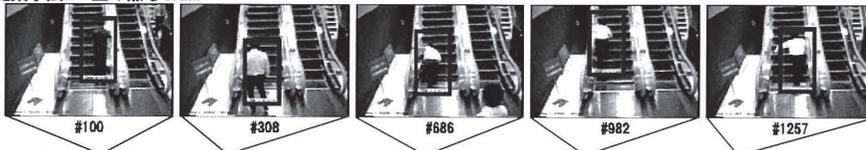
実験では、提案手法である CRIM を用いた異常行動検知手法と、従来法である CHLAC 特徴を用いた異常行動検知手法を、エスカレーターシーンに対して適用し、異常行動検知性能の評価を行う。

CRIM を用いた提案手法としては、重み付き時空間勾配により CRIM を算出して異常行動検知を行う手法 (提案手法 1) と、重み付き時空間勾配無しに CRIM を算出して異常行動検知を行う手法 (提案手法 2) とする。この提案手法 1 と提案手法 2 により、重み付き時空間勾配の有効性を評価する。

提案手法1：重み付きCRIM



提案手法2：重み無しCRIM



□ : 異常行動
□ : 正常行動

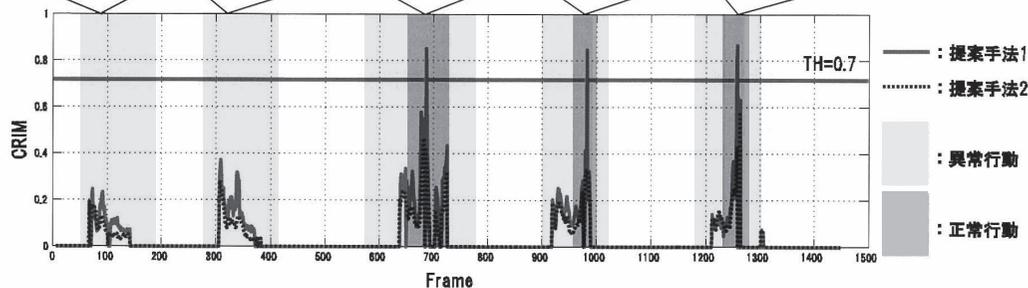


図 7: CRIM による異常行動検知結果

CHLAC 特徴を用いた従来法としては、CHLAC 特徴を用いて異常行動検知を行う文献 [1] の手法 (従来法 1) と、動的背景除去により検出された物体領域における CHLAC 特徴を用いて異常行動検知を行う手法 (従来法 2) とする。

従来法 1 では、正常行動をしている人の学習データ 3000[frame] より算出した CHLAC 特徴 (20[frame] 積算) を使用し、部分空間を構築する。実験では、部分空間構築における累積寄与率を 0.99 としたとき、得られた部分空間の次元は 14 次元となった。この従来法 1 により、提案手法と文献 [1] の手法の異常行動検知性能の評価を行う。

従来法 2 では、提案手法により動的背景除去を行った物体領域から CHLAC 特徴 (20[frame] 積算) を算出し、部分空間を構築する。これにより、動的背景による影響を除去した提案手法と同じ条件で異常行動検知を行う。従来法 1 と同様に、部分空間構築における累積寄与率を 0.99 としたとき、得られた部分空間の次元は 5 次元となった。この従来法 2 により、提案手法で用いる CRIM と従来法で用いる CHLAC 特徴との、異常行動検知の指標に関する有効性を評価する。

4.2 提案手法 (CRIM) による異常行動検知結果

図 7 に提案手法 1 と提案手法 2 による異常行動検知結果を示す。図 7 より、CRIM の閾値を $TH=0.7$

としたとき、提案手法 1 では正常/異常行動を正しく判別できることが分かる。ここで、提案手法 2 に対して提案手法 1 の方がより異常行動時に高い CRIM の値を観測しており、さらに正常行動時には低い CRIM 値を観測している。故に、提案手法 1 の方が異常行動検知性能が高いといえる。

CRIM を用いた異常行動検知手法について考察を行う。まず、提案手法 2 では、背景を含む人の領域全体から CRIM を算出している。これにより、提案手法 2 で算出される CRIM は大域的に捉えた値となる。これに対し、提案手法 1 では重み付き空間勾配を用いることにより、人の局所的な形状特徴を捉えた CRIM の算出を可能にしているため、検知性能が向上したと考えられる。

4.3 従来法 (CHLAC) による異常行動検知結果

図 8 に従来法 1 と従来法 2 による異常行動検知結果を示す。図 8 より、従来法 1 では、正常行動と異常行動における部分空間の逸脱度の差があまり表れていないことが分かる。また、従来法 2 では、動的背景による影響がないため、従来法 1 と比較して正常/異常行動を判別することができている。しかし、未検知の異常行動シーンや、人の横切りによる誤検知が発生していることが分かる。

CHLAC 特徴を用いた異常行動検知手法について考察を行う。図 9 に従来法 1、従来法 2 で用いられ

従来法1：CHLAC特徴



従来法2：動的背景除去+CHLAC特徴

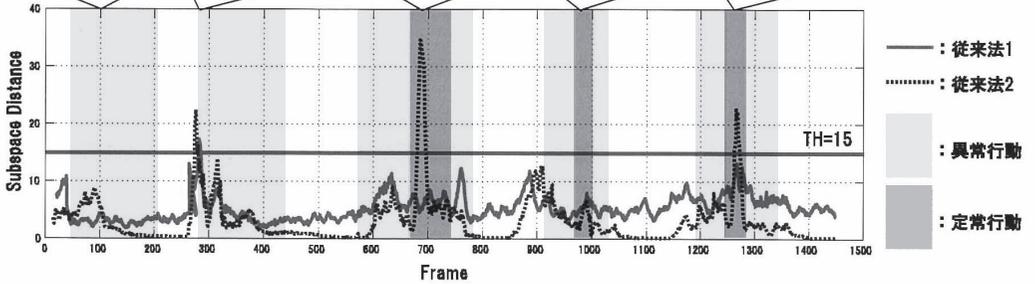


図 8: CHLAC による異常行動検知結果

る時系列 2 値化画像を示す。図 9 より、従来法 1 では、エスカレータのステップからも動き特徴が検出されていることが分かる。よって、人の動き特徴のみを捉えることができないため、異常行動検知が困難となったと考えられる。従来法 2 では、人の領域のみの動き特徴を得ることができている。しかし、この時系列 2 値化画像より得られる CHLAC 特徴では、人の転倒に対する特徴を上手く捉えることができないため、異常行動検知が困難となったと考えられる。また、従来法 1、従来法 2 では、正常行動時に算出される CHLAC 特徴を学習データとし、部分空間を構築しているため、人の横切りによる動き特徴が検出された場合、部分空間から逸脱した特徴であると判定され、異常行動として検知する場合がある。

提案手法と CHLAC 特徴に用いる時系列フレーム数について考察する。提案手法で用いている ST-patch では、時系列フレームを 5[frame] 使用し、異常行動検知を行っている。これに対し CHLAC 特徴では、20[frame] 積算した特徴である。そのため、提案手法の方が短いフレームバッファで異常行動を検知することが可能である。

4.4 定量的評価

提案手法における異常行動検知性能の定量的な評価を行う。評価には、閾値を変化させたときの

従来法1：CHLAC特徴



従来法2：動的背景除去+CHLAC特徴

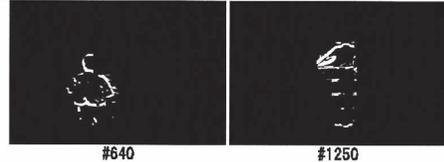


図 9: 時系列 2 値化画像

検出率と誤検出率をプロットした ROC (Receiver Operating Characteristic) カーブを用いる。ここで示す検出率は、異常行動として正しく検知した割合を表し、誤検出率は正常行動を異常行動として誤検知した割合を表す。評価データには、ステップ上り、下り方向のエスカレータシーンの正常行動 28 シーン、異常行動 18 シーンの計 46 シーンを用いる。

図 11 に提案手法 1、提案手法 2、従来法 1、従来法 2 の ROC カーブを示す。結果より、4つの手法の中で、提案手法 1 のプロット点が最も左上に近いため、最も異常行動検知の良い手法であるといえる。また、提案手法 1 では、従来法 1 と比較して誤検出率 10% のとき検出率を約 27% 向上させることができた。



図 10: 提案手法による複数人の異常行動検知例

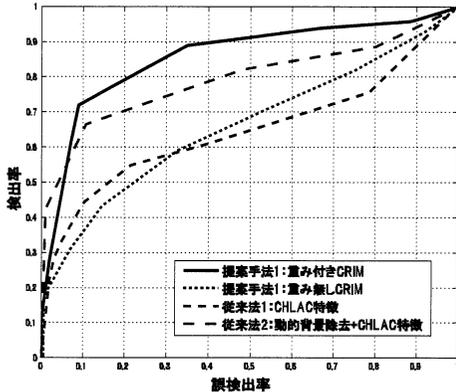


図 11: ROC カーブによる定量的評価

以上の結果より、動的な背景であるエスカレータシーンの場合、動的背景除去と、対象物体の形状特徴を捉えた CRIM を指標とした異常行動検知の有効性を確認することができた。

4.5 複数人の異常行動検知

提案手法では、時空間特徴を用いた人領域検出を行い、その領域ごとに異常行動検知を行っているため、複数人のシーンに対しても異常行動検知を行うことができる。図 10 に提案手法による複数人の異常行動検知例を示す。図 10 に示すように、提案手法では動的な背景上から異常行動が画像中の何処で発生しているかを特定することが可能である。

5 おわりに

本稿では、動的背景モデリングにより検出された物体領域より、統計的学習である RealAdaboost に基づく人領域検出・異常行動検知を行う手法を提案した。提案手法では、動的背景を持つエスカレータのシーンにおいて、異常行動の検知を行うことが可能であり、評価実験によりその有効性を確認した。今後は、ST-patch サイズの可変によるマルチスケール人領域検出を行う予定である。

参考文献

- [1] 南里卓也, 大津展之. 複数人動画からの異常動作検出. 情報処理学会論文誌. コンピュータビジョンとイメージメディア, Vol. 45, No. SIG_15, pp. 43-50, 2005.
- [2] T. Kobayashi and N. Otsu. Action and Simultaneous Multiple-Person identification Using Cubic Higher-Order Local Auto-Correlation. In *Proc. International Conference on Pattern Recognition*, pp. 741-744, 2004.
- [3] 関真規人, 林健太郎, 谷口博康, 橋本学, 笹川耕一. リアルタイム人物暴れ検出システム. 第 10 回 画像センシングシンポジウム, pp. E-02-1-6, 2004.
- [4] 都築勇司, 西村孝, 藤吉弘亙, 森次郎. 特徴点追跡結果を用いた動きのモデル化と例外行動検出. 第 14 回 画像センシングシンポジウム, No. IN1-07, pp. 1-8, 2008.
- [5] E. Shechtman and M. Irani. Space-Time Behavior Based Correlation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 405-412, 2005.
- [6] Y. Ke, R. Sukthankar and M. Hebert. Event Detection in Crowded Videos. In *Proc. IEEE International Conference on Computer Vision*, pp. 8-15, 2007.
- [7] Y. Murai, H. Fujiyoshi and T. Kanade. Combined Object Detection and Segmentation by Using Space-Time Patches. In *Proc. 8th Asian Conference on Computer Vision*, Vol. Part I, No. LNCS 4843, pp. 915-924, 2007.
- [8] A. P. Dempster, N. M. Laird and D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society*, Vol. 39, No. 1, pp. 1-38, 1977.
- [9] B. Wu, H. Ai, C. Huang and S. Lao. Fast Rotation Invariant Multi-View Face Detection based on Real Adaboost. In *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 79-84, 2004.
- [10] D. Comaniciu and P. Meer. Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 603-619, 2002.