

1. はじめに

近年、動画理解技術によるセキュリティシステムを目的とした人の不振な行動・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。

このような背景を受け、本研究では画像中の物体の「アピランス（見え）」と「モーション（動き）」の情報をを持った時空間特徴を用いて、動的な背景上の人の異常行動検知を行うことを目的とする。

2. Space-Time Patch 特徴

本研究では、物体の局所的な動きの時間的変化を捉えた時空間特徴である Space-Time Patch (ST-patch) 特徴 [1] を用いる。この ST-patch 特徴は、画像を時間方向に重ねた 3 次元データである時空間画像における 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、局所領域の画素数 n でスタックした行列 G を算出する。

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

この時空間勾配値からなる $n \times 3$ の行列 G と行列 G^T を掛け合わせた行列を M とすると、行列 M は式 (2) で表すことができる。

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

式 (2) より求められる行列 M が 1 つの ST-patch から得られる特徴量となる。行列 M には、左上の 2×2 の行列にアピランスの情報、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 特徴は対角成分が同一値のベクトルとなっているため、本研究ではこの ST-patch 特徴を式 (3) の 6 次元ベクトル e として扱う。

$$e = \left(\sum P_x^2, \sum P_x P_y, \dots, \sum P_t^2 \right) \quad (3)$$

3. ST-patch 特徴を用いた人の異常行動検知

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のモデリングを行う。作成したモデルを用いることで動的背景を除去し、物体領域のみを検出する。次に、検出された物体領域から Real AdaBoost 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、異常行動の検知を行う。

3.1. 混合正規分布による時空間特徴モデリング

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。データ x の確率密度分布 $p(x)$ を C 個のクラスを持つ確率密度分布 $p(x|i)_{i=1, \dots, C}$ の重み付き線形結合とすると、 $p(x)$ は式 (4) で表される。

$$p(x) = \sum_{i=1}^C w_i p(x|i) \quad (4)$$

$$p(x|i) = \frac{1}{\sqrt{(2\pi)^M |\sigma_i|}} \exp \left\{ -\frac{1}{2} (x - \mu_i)^T \sigma_i^{-1} (x - \mu_i) \right\} \quad (5)$$

この時、 M は次元数、 w_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 w_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現した ST-patch 特徴 e に混合正規分布を当てはめることでモデリングを行う。この混合正規分布モデルから、入力特徴量の逸脱度を求めることによって物体/背景領域の判別を行う。図 1 に提案手法と正規化距離による物体検出法をエスカレータ監視映像に適用した例を示す。



図 1: エスカレータ監視映像における物体検出例

3.2. Real AdaBoost に基づく人領域検出

時空間特徴と統計的学習手法を組み合わせ、人領域検出を行う。提案手法では、統計的学習手法として Real AdaBoost を用いる。また、Real AdaBoost により学習する時空間特徴は、時空間特徴モデリング手法により物体領域として判別された特徴ベクトルを使用し、背景領域と判別された特徴ベクトルは 0 として扱う。この時、6 次元ベクトルとして表現した ST-patch 特徴 e を式 (6) により、ベクトルの総和で正規化する。

$$e_N = \frac{e}{\left(\sum_{i=1}^6 |e_i| + \epsilon \right)} \quad (6)$$

正規化特徴ベクトルを用いて学習を行うことにより、衣服の違いや照明変動の影響を受け難い検出器を構築することができる。構築された識別器を用いて人領域検出するには、検出ウィンドウのラスタ走査を行い、検出ウィンドウ内の特徴ベクトルが人領域/非人領域であるかの判別を行う。

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、検出された人領域を 1 つの ST-patch とし、ST-patch より求められる CRIM (Continues Rank-Increase Measure) により正常/異常行動の判定を行う。

CRIM の算出 ST-patch 特徴を表す 3×3 の行列 M と、式 (7) の行列 M^\diamond より求められる固有値を用いて $CRIM \Delta r$ を算出する。

$$M^\diamond = \begin{bmatrix} \sum P_x^2 & \sum P_x P_y \\ \sum P_y P_x & \sum P_y^2 \end{bmatrix} \quad (7)$$

$\lambda_1 \geq \lambda_2 \geq \lambda_3$ を行列 M の固有値、 $\lambda_1^\diamond \geq \lambda_2^\diamond$ を行列 M^\diamond の固有値とする。行列 M , 行列 M^\diamond は共に対称行列であり、行列 M , 行列 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、 $CRIM \Delta r$ を式 (8) のように定義する。

$$\Delta r = \frac{\lambda_2 \cdot \lambda_3}{\lambda_1^\diamond \cdot \lambda_2^\diamond} \quad (0 \leq \Delta r \leq 1) \quad (8)$$

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、動きベクトルが分散して

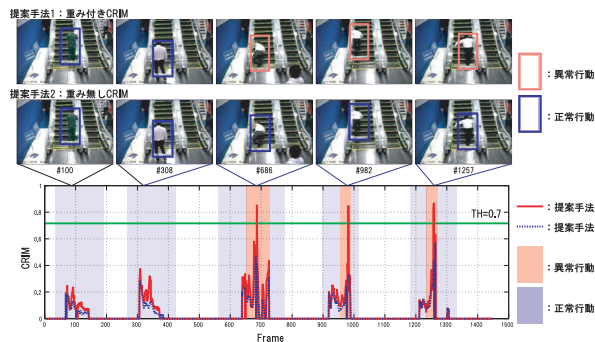


図 2: CRIM による異常行動検知結果

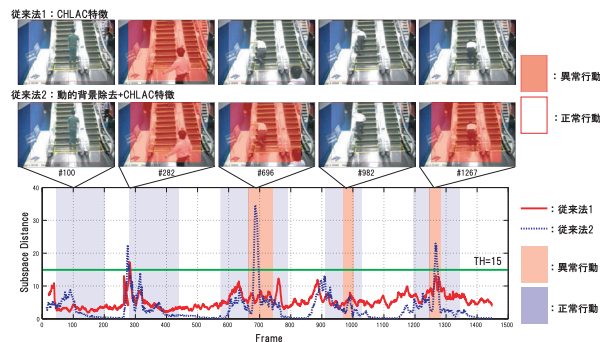


図 3: CHLAC による異常行動検知結果

いるとき $\Delta r \approx 1$ 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．

重み付き時空間勾配を用いた CRIM の算出 提案手法では，Real AdaBoost 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．Real AdaBoost に基づく人領域検出時には各弱識別器より，以下に示す弱識別器出力 $h(x)$ ，確率密度分布の分離度 α の 2 つの値を得ることができる．

$$h(x) = \frac{1}{2} \ln \frac{W_+^j + \epsilon}{W_-^j + \epsilon} \quad (9)$$

$$\alpha = 1 - \sum_j \sqrt{W_+^j W_-^j} \quad (10)$$

ここで， x が入力された ST-patch 特徴， W_+^j と W_-^j は Positive クラスと Negative クラスの確率密度分布， j が 1 次元ヒストグラムのピン番号を表す．

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．ここで，提案手法では評価値が $h(x) > 0$ 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，評価値を式 (11) として表現する．

$$E = \begin{cases} h(x)\alpha & \text{if } h(x) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

このような評価値によって，動的背景の影響を抑制し，さらに物体の局所領域の動きに注目す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，時空間勾配に掛合わせることで，式 (12) のような重み付き時空間勾配の行列 G_E を算出する．

$$G_E = \begin{bmatrix} E_1 P_{x_1} & E_1 P_{y_1} & E_1 P_{t_1} \\ E_2 P_{x_2} & E_2 P_{y_2} & E_2 P_{t_2} \\ \vdots & \vdots & \vdots \\ E_n P_{x_n} & E_n P_{y_n} & E_n P_{t_n} \end{bmatrix}_{n \times 3} \quad (12)$$

評価値によって重み付けされた時空間勾配の行列 G_E から，先に述べた手順により CRIM の算出を行う．

4. 評価実験

本研究の有効性を示すため，動的背景であるエスカレータ監視カメラ映像を用いて人の異常行動検知実験を行う．比較手法には，重み付き時空間勾配により算出した CRIM を用いた手法，重み無し時空間勾配により算出した CRIM を用いた手法，従来法である CHLAC 特徴を用いた手法，動的背景除去画像における CHLAC 特徴を用いた手法を使用する．

4.1. 異常行動検知結果

図 2，図 3 に示す異常行動検知結果より，提案手法 1 では正常/異常行動を正しく判別できているが，提案手法 2，従来法 1，従来法 2 では，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．

4.2. 定量的評価

図 4 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ので，異常行動検知の良い手法であると言える．さらに，提案手法 1 では従来法 1 と比較して誤検知率 10% の時の検知率を約 27% 向上させることができた．

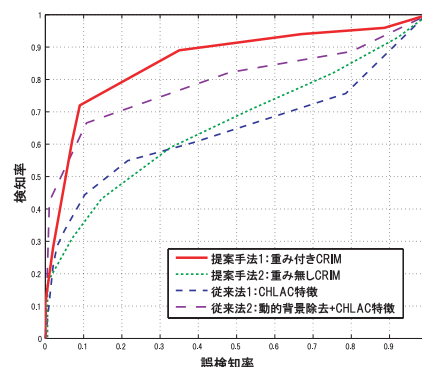


図 4: ROC カーブ

4.3. 複数人シーンの異常行動検知

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可能である．図 5 に提案手法による複数人の異常行動検知例を示す．

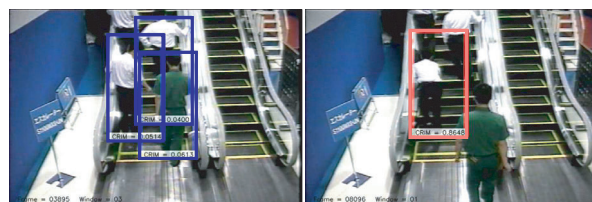


図 5: 提案手法による複数人の異常行動検知例

5. おわりに

時空間モデリング手法により動的背景除去を行い，統計的学習である Real AdaBoost に基づく人領域検出，及び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も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．

参考文献

- [1] 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 .

研究業績

- [1] 村井泰裕, 藤吉弘巨, 金出武雄. Space-Time Patch を用いた物体の移動方向識別とセグメンテーション. 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.1, No.2, pp.21-31, 2008.

(他 国際会議口頭発表 1 件, 学会口頭発表 1 件, 研究会技術報告 2 件)