

長時間の監視映像からの非定常シーケンスの検出

数藤 恭子[†] 若林 佳織[†] 荒川 賢一^{††} 安野 貴之[†]

[†] NTT サイバースペース研究所

横須賀市光の丘 1-1

^{††} NTT サイバーコミュニケーション総合研究所

E-mail: †{sudo.kyoko,wakabayashi.kaoru,arakawa.kenichi,yasuno.takayuki}@lab.ntt.co.jp

あらまし 映像中に含まれる動物体の軌跡から非定常なシーケンスを識別し、非定常性の定量的な指標を抽出する手法を提案する。本研究では、動物体領域の中心軌跡を特徴量として1クラスSVMによる学習を行い、高次元特徴空間における識別境界からの外れ具合を非定常性の指標とした。オフィスの入退室映像を入力として、動物体が出現してから消えるまでの各シーケンスの識別実験を行った結果、様々なパターンの人の滞留や複数人の同時通過などのシーケンスが非定常の度合いの高い映像として識別された。また、通常の入退室のシーケンスは定常的な映像として識別され、おおむね直感と近い結果が得られた。これらの結果を非定常性の度合いの高い順に並べて提示することで、監視映像のスクリーニングの効果が期待できる。

キーワード 監視映像, 非定常検出, 1クラスSVM

Detecting Anomalous Sequences in Long Duration Monitoring Videos

Kyoko SUDO[†], Kaoru WAKABAYASHI[†], Kenichi ARAKAWA^{††}, and Takayuki YASUNO[†]

[†] NTT Cyber Space Laboratories

1-1 Hikarinooka, Yokosuka, JAPAN

^{††} NTT Cyber Communications Laboratory Group

E-mail: †{sudo.kyoko,wakabayashi.kaoru,arakawa.kenichi,yasuno.takayuki}@lab.ntt.co.jp

Abstract A method to discriminate anomalous image sequences containing moving objects from monitoring videos is proposed. This method learns the feature of tracks of moving objects in each sequence by using the 1-class SVM algorithm to discriminate the anomalous outlier sequences. We define the degree of anomalousness as the distance from the decision boundary to the samples in a high dimensional feature space. As the input of the application, we use image sequences from monitoring videos taken at an office entrance. The results show that various anomalous sequences such as people hanging around the entrance or multiple people passing by were detected that evidenced high levels of anomalousness. The results suggest that this method can be effectively used as a screening method for long duration monitoring videos.

Key words monitoring video, anomaly detection, 1-class SVM

1. はじめに

近年、社会的不安の増加のためビデオカメラを用いた自動監視システムの開発が盛んに行われている。カメラの低価格化、ネットワークやデータ蓄積技術の発達により、多数の監視カメラ設置が可能になり、大量の映像データがセンターに送られ蓄積されている。しかし、それらのデータのほとんどは定常的な映像であり、実際に監視員によるチェックが必要な部分の時間的な割合はわずかである。このような現状を踏まえて、監視業務の効率化を目指した非定常的なシーケンスの自動認識技術が求められている。

非定常的なシーケンスを判断するための一つの重要な指標は、映像中の人物の有無とその行動である。このため、近年監視映像を対象とした顔検出や動領域抽出、人物追跡技術に力が注がれ、精度のよい人物検出や追跡のための有効なアルゴリズムも提案されている。これらの検出や追跡の技術を有効利用して行動パターンを解析し、非定常的なシーケンスから順番に映像を提示することができれば、長時間の監視映像のスクリーニングに有効である。

本研究では、監視映像のサバイランスの一手法として、非定常の度合いの高い順にシーケンスを並べ変えて提示する方法を提案する。そのために、非定常の度合い(以下、非定常性)の指標を定量的に抽出することを目的とする。

以下本稿では、2.章において映像からの非定常検出に関する従来手法を整理し、3.章において、本提案手法の詳細を述べる。特徴量として用いた動領域の重心軌跡画像の抽出方法、および1クラスSVMを用いた非定常性の指標の抽出方法を説明する。??章では、オフィスの出入口で撮影した監視映像を用いて行った実験の結果と考察を述べる。

2. 非定常検出に関する従来手法

映像中の人物や物体の行動パターンの学習に基づいて非定常検出を行う従来研究には、いくつかの異なるアプローチがみられる。ここではまず、教師ありの場合と教師なしの場合に大別して考える。本研究では監視映像を対象としているが、監視映像への適用を目的とする場合、学習に十分な数の非定常サンプルを得ること自体が困難な場合も多い。このため、現実的には教師あり学習が適用困難なケースがある。そこで以下では、教師なしの場合のみについて考える。

教師なし学習による非定常検出に関する従来研究は大きく分けて2つのアプローチがある。一つは、定常的な行動パターンをクラスタリングし、新たな入力シーケ

スをクラスタリングされたパターンと照合する方法である。もう一つは、サンプルの特徴空間での分布を学習し、分布からのはずれ具合によって非定常性を判定する方法である。

一つめの定常的な行動パターンをクラスタリングするアプローチの例として、青木らは家庭内の行動監視を目的として、玄関・トイレ・冷蔵庫といった行動の拠点にモニタやセンサを設置し、各センサの反応をシンボルとするHMMで人物の家の中の行動パターンを学習している[?]。Fuらは交通監視を目的として、道路を通る車の定常的な軌跡をクラスタリングし、新たなサンプルとクラスの照合を行って、車の動きの定常性を判定している[?]。小阪らは、未知の行動を行う人物の検出を目的として、全方位画像から歩行軌跡を抽出してモデル化している[?]。これらの方法では、学習されたクラスのいづれかに入力サンプルが属する場合は定常、いづれのクラスにも属さない場合は非定常となる。

もう一つの方法である、サンプルの特徴空間での分布を学習し、分布からのはずれ具合によって非定常性を判定する例として、横井らは、移動監視ロボットが捉えた映像からパターンの分布モデルを推定し、確率的に異常判定を行っている[?]。Nanriらは、歩行人物を含む映像から教師なし学習によって線形部分空間を構成し、部分空間との距離で非定常性の尺度を定義することで、転倒などの映像を自動抽出している[?]。McKennaらは、家の中の映像から、椅子、ベッド、出入口などの間の動線を確率分布モデルとしてEMアルゴリズムを用いて推定し、入力サンプルの定常性を確率的に判定している[?]。

前者のアプローチでは、出力は定常または非定常という2クラスである。後者のアプローチでは、非定常の度合いを学習サンプルの分布との距離などで定義することにより、定量的に表現している。本研究では、非定常性の尺度を抽出することを目的としており、後者のアプローチに属する。すなわち、学習サンプルの特徴空間における分布の教師無し学習を行い、はずれ値を非定常と判定し、学習サンプルの分布との距離から非定常性の度合いを定量化する。

3. 提案手法

監視映像における不審者検出を実現する場合、最終的な判断は人間の目視によって行われるようにするのが現実的な方策である。その場合、判断する人間の手間が省けるよう、非定常性の度合いの大きい順に提示することが望ましい。本研究では、データの分布からのはずれの程度を定量化するために適した識別手法として、1クラスSVMを用いる。1クラスSVMは主に外れ点検出に利用される手法だが、集合からのはずれ具合をも数値とし

て求めているので、これを非正常性の度合いの指標とみなして利用することにする。

1クラス SVM は、ガウシアンカーネルを用いて特徴量を写像した高次元特徴空間において、はずれ値が原点付近に寄る性質を利用した1クラスの識別器である。実際のデータへの適用を示した研究例では、Manevitz[?]がドキュメントの分類に1クラス SVM を用いている。nearest neighbour と比較して、良好な結果が得られるが、パラメータに敏感であると報告している。非正常検出に1クラス SVM を用いた従来研究としては、発電所のセンサ出力に適用した研究がある[?]。例外データ検出実験により、ベテラン運転員の操作をサポートする結果が得られたとの報告がある。監視映像からの非正常検出への1クラス SVM 適用については、特徴量やパラメータをどのように用いると効果的であるか知られていないため、今後検討の必要がある。

3.1 映像からの特徴量の抽出

1クラス学習の入力とするための特徴量を決定するにあたり、(1) 人が監視映像をチェックするときに注目する変化を反映した特徴量であること、(2) 時間的な伸縮にロバストな特徴量であること の2つを条件とする。(1)は、本研究では映像を人がチェックする前のスクリーニングを目的としていることによる。(2)は、時間的な伸縮による変動の大きさを吸収し、例えば、映像中の人物の歩行速度の違いは無視できるようにするためである。まず、(1)を考慮し、監視映像を人が見る時に注目していると考えられる、物体や人物の動き方や経路を反映する特徴量として動領域の重心軌跡を用いることにした。

入力画像の時系列から、フレームごとに動領域を抽出し、動領域の内部は1それ以外は0に二値化する。n番目のフレームの画像の座標 (x, y) における画素値を $I(x, y, n)$ とすると、動領域の重心座標は式(1)で求められる。

$$C_x(t) = \frac{\sum_{x=0}^X \sum_{y=0}^Y x I(x, y, t)}{\sum_{x=0}^X \sum_{y=0}^Y I(x, y, t)}$$

$$C_y(t) = \frac{\sum_{x=0}^X \sum_{y=0}^Y y I(x, y, t)}{\sum_{x=0}^X \sum_{y=0}^Y I(x, y, t)}$$

$$t = 1, 2, \dots, N$$
(1)

このようにして動領域の重心座標の時系列を求めた後、各時刻の重心座標の間を線分をつなぐ。以上によって、動領域の軌跡形状を連続線として求める。

次に、(2)を考慮して、上記のように求められた軌跡形状を時間の伸縮を吸収できるように特徴ベクトル化する。軌跡形状の座標値を1、それ以外を0とする画像を作成し、この画像の画素値の並びを特徴量とする。軌跡を画像として扱うことにより、時間の伸び縮みを吸収できる。例えば、通過する人物の経路が同じであれば、歩く速度

が異なっても同じ特徴量が得られる。時間の伸縮については、非正常として速度の違いが重要な場合もあり得るが、本研究では(2)の条件を重視して時間情報は用いないこととした。

次に、この特徴ベクトルについて、時刻0における座標が一致するように位置の正規化を行う。始点の位置を揃えることによって、動きの方向が区別できる。例えば、人が全く同じ道筋を歩いた場合、得られる軌跡の形状は同じであるが、特徴ベクトルとしては方向の違い(行きと帰り)を区別できる(図1(b), (c))。

3.2 非正常性の指標の抽出

本研究では、特徴空間におけるはずれ点を非正常とみなし、多くのサンプルを含む領域からはずれている度合いが大きいほど非正常性が高いとする。多くのサンプルを含む領域を求めるために、1クラス SVM を用いる。

3.2.1 1クラス SVM

1クラス SVM [?][?] は、ガウシアンカーネル $K(x_i, x) = \exp(-\frac{\|x_i - x\|^2}{\sigma^2})$ を用いて特徴空間への写像を行うと、入力空間で他から孤立しているはずれ点は高次元特徴空間の原点近くに写像されるという性質を利用している。式(2)を識別関数とすると、式(3)を解くことにより、あらかじめ決められた割合 ν のサンプル群が原点側に残るように原点とサンプル群を分けるような超平面が求められる。

$$\text{識別関数 } f(x) = \text{sign}(\omega\Phi(x) - \rho) \quad (2)$$

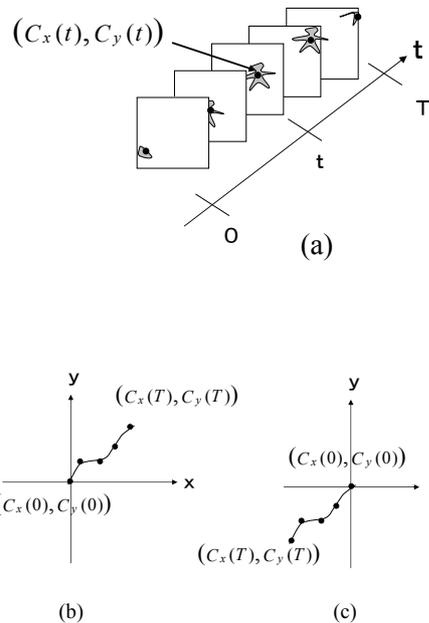


図1 特徴量の抽出方法。(a) 画像の時系列の各フレームごとに動領域を抽出し、重心を求める。(b) $t=0$ における重心を原点として重心座標の時系列を画像上にプロットし、各時刻の重心座標の間を線分をつなぎ重心軌跡画像とする。

$$\min_{w \in F, \xi \in R^n, \rho \in R} \frac{1}{2} \|\omega\|^2 + \frac{1}{\nu n} \sum_i \xi_i - \rho \quad (3)$$

制約条件 $\omega \Phi(x_i) \leq \rho - \xi_i, \xi_i \leq 0$

式 (2) と式 (3) のカーネルトリックによる非線形への拡張はそれぞれ式 (??), 式 (??) となる .

$$\text{識別関数 } f(x) = \text{sign}(\sum_i \alpha_i K(x_i, x) - \rho) \quad (4)$$

$$\text{最適化する関数 } \min_{\alpha} \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j) - \rho \quad (5)$$

$$\text{制約条件 } 0 \leq \alpha_i \leq \frac{1}{\nu n}, i = 1, \dots, n$$

$$\sum_{i=1}^n \alpha_i = 1$$

1 クラス SVM は, 式 (??) を最適化することにより, 原点とすべてのサンプルが最も離れるような識別軸を決定する . このとき, あらかじめ設定したパラメータ ν によって決まる割合のサンプルがはずれ点となる .

3.2.2 非正常性の定量化

特徴空間において, あるサンプル x と x 以外のすべての点の距離が離れているとき, 式 (??) の識別関数 $f(x)$ 中の $\sum_i \alpha_i K(x_i, x) - \rho$ の値は小さくなる . $\sum_i \alpha_i K(x_i, x) - \rho$ が負の時, そのサンプル x は非正常となる . そこで, 非正常と識別されたサンプルについて, $|\sum_i \alpha_i K(x_i, x) - \rho|$ (以下 = $|g(x)|$) の値を非正常性の指標として用いることにする . $|g(x)|$ が大きいほど非正常性が高く, 非正常データについて $|g(x)|$ の最大値は ρ である .

非線形変換のため, $g(x)$ の大小と元の特徴空間でのサンプル同士の距離の関係は不明であるが, Scholkopf は実験によって, ν の値を変化させたとき, 識別境界が分布のまとまりの中心から等高線状に変化する様子を 2 次元特徴量のデータを用いて示している [?] . 本研究においては, 2 次元データを用いた予備実験を行い, ν の値の大きさと分布の中心からの外れ具合について順序性がみられることを確認した . また, σ の値は今回は固定値とし, 予備実験の際に適切な値を設定した . 式における ν の値は, はずれ点の割合を示すため, 全データの何%を非正常とするかをユーザが設定することができる .

4. 実験と考察

4.1 入力データ

本研究は, 監視映像のスクリーニングを目的としているため, 実験用のデータとして, オフィスの入退室の監視映像を用いた . 撮影には, ネットワークカメラ^(注1)を使用した . カメラは, 部屋の入り口のドアから数 m 内側の壁の約 2m の高さに設置されている . 今回は, カメラの

動物検知機能を利用し映像中に動物体が出現してから消滅するまでを検知して自動的にカットされたシーケンスを用いた . 各シーケンスは毎秒 15 フレームの画像列としてネットワークを經由して PC に蓄積し, その後オフラインで処理を行った .

カメラから得られる 640×480 画素の画像を $1/4$ に縮小した 160×120 画素の画像を用いて処理を行った処理に用いた入力画像列の例を図 2 に示す . ある 1 日について, 24 時間の撮影で得られた 180 シーケンスを入力として用いた .

4.2 特徴抽出

各シーケンスから動領域の重心軌跡画像を特徴量として抽出する . 抽出した特徴量の例を図??に示す . 図??において, 軌跡の部分は 1 それ以外は 0 の値をとる . 画像のサイズは 160×120 である . 1 クラス SVM の入力とする特徴ベクトルの次元数はこれに等しい .

4.3 1 クラス SVM

本研究では, 1 クラス SVM には LIBSVM^(注2) のパッケージを用いた . パラメータの値は $\nu = 0.08$, $\sigma = 0.00001$ とした .

4.4 結果

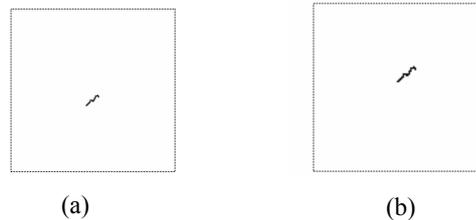
$g(x)$ の値の分布のヒストグラムを図??に示す . 入力



(a)

(b)

図 2 入力シーケンスの例 (代表フレームのみ表示) .



(a)

(b)

図 3 重心軌跡画像の例 (図 2(a)(b) それぞれの各シーケンスから抽出) .

(注1): カメラは AXIS 211A ネットワークカメラを使用した .

(注2): <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

シーケンスのうち、13シーケンスについて、 $g(x)$ の値が負となり、非定常状態と識別され、それ以外は定常状態と識別された。定常状態と識別されたシーケンスの例の代表フレームを図5に示す。非定常状態と識別されたシーケンスは、上位から順に(1)深夜の警備員の見回り、清掃業者の出入り(2)入り口付近での来訪者への対応や、複数人数での通過などであった。これらのシーケンスの例の代表フレームを図6に示す。また、それらの重心軌跡特徴を図7に示す。 $|g(x)|$ の値は、図中左のシーケンスから順に0.0042, 0.0013, 0.0010であった。

4.5 考察

本提案手法では、特徴量を決定するにあたり、人物の動き方と位置の違いを反映し、移動速度の違いを吸収するため、動領域の重心軌跡画像を用いた。非定常状態として識別されたシーケンスの重心軌跡画像は、警備員の見回りについては、暗闇の状態からドアが開く様子の変化として検知されるため、動領域の重心軌跡の形状が直線的で、軌跡の現れる位置が他と異なっていた。清掃業者のシーケンスについては、大きい台車やゴミ袋を持っているため、通常とは動領域の重心位置が異なっていた。来訪者対応のシーケンスの場合は、軌跡が複雑な形状になっていた。定常状態として識別されたシーケンスは、ほぼすべて入室と退室の類似したシーケンスであった。人によって歩く速度が異なり、1シーケンスが20フレームから50フレームと幅があったが、いずれも定常状態として識別された。以上の結果では、軌跡の形状と位置を反映した特徴量で、単純な通過とそれ以外のシーケンスを区別することができたといえる。今後他の監視映像に適用した場合に、

- 通常と異なる経路の移動
- 滞留(うろろう・覗き込み)
- 不審物の持ち込み

などが検出できる可能性がある。

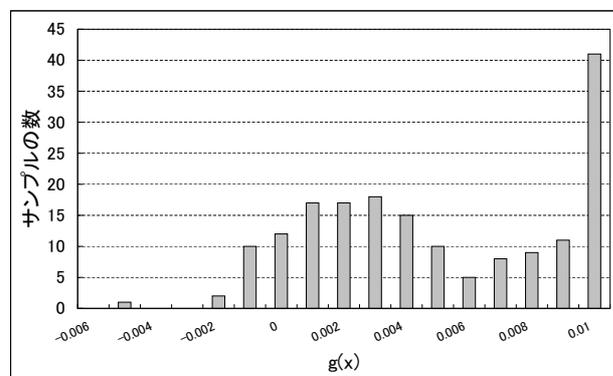


図4 $g(x)$ の分布のヒストグラム。 $g(x) < 0$ に分布しているサンプルが非定常に識別される。

今回非定常シーケンスについて得られた非定常性の数値については、現段階では定量的な評価は行っていないが、おおむね直感に近い結果が得られた。本提案手法により識別される非定常シーケンスが実際に人の主観に近いものであるかどうか、評価の手法も含めて今後の検討が必要である。

実験において検出された非定常シーケンスの全シーケンスに対する割合は、1クラスSVMのパラメータ ν の値によってあらかじめ設定した。このパラメータはユーザが設定しなくてはならないが、人が最終的に見てチェックする監視映像のスクリーニングに適用する場合、ユーザが非定常検出の割合をパラメータで指定することは妥当と思われる。

5. おわりに

長時間の監視映像のスクリーニングを目的として、映像中の動物体を含むシーケンスを非定常性の高い順に提

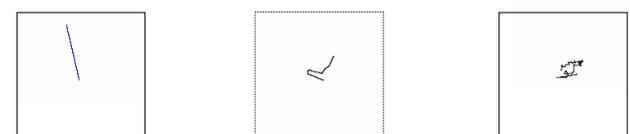


図5 通常に分類されたシーケンスの例(それぞれ代表フレームのみ表示)。



(a) (b) (c)

図6 非定常と識別されたシーケンスの例(それぞれ代表フレームのみ表示)。(a)(b)(c)の順に非定常性が高い。



(a) (b) (c)

図7 非定常と識別されたシーケンスの重心軌跡画像。図6(a)(b)(c)に対応。

示する手法を述べた．

まず，特徴量として始点を正規化した動領域の重心軌跡の画像を用いることを提案した．始点を揃えることで通過する人物の出入りの方向が区別できるようにし，画像にすることで，時間方向の変動を吸収できるようにした．次に，この特徴量を用いた特徴空間で他から孤立している外れ点を非正常データと考え，他のサンプルからの外れ具合を非正常性と捉えた．ガウシアンカーネルを用いた画像では入力空間における孤立点が高次元特徴空間の原点近くに写像されるという性質に注目し，1 クラス SVM の識別関数を利用して非正常性の指標の数値化を提案した．最後に上記の手法を適用し，オフィスの入退室の監視映像を入力とする実験を行った．その結果，非正常として検出されたシーケンス，および提示された非正常性の順序はおおむね直感に近いものであった．

今後は，本提案手法により得られる非正常判定結果が人が映像を見る場合の主観とどの程度一致するかの検証を進める．

文 献

- [1] 青木, 大西, 小島, 福永: “人物の行動パターンの学習・認識と独居高齢者の非日常状態検出への応用”, 第 18 回生体・生理工学シンポジウム論文集, No. 3B1-3, 2003.
- [2] 青木, 大西, 小島, 福永: “全方位センサを用いた行動パターンの学習・認識と非日常状態検出への応用”, 第 21 回日本ロボット学会学術講演会予稿集, No. 3K33, 2003.
- [3] Z. Fu, W. Hu and T. Tan: “Similarity Based Vehicle Trajectory Clustering and Anomaly Detection”, Proc. IEEE International Conference on Image Processing, ICIP2005.
- [4] 小阪, 平山, 岩井, 谷内田: “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-149, pp.163–170, 2005.
- [5] 横井: “画像に基づく監視ロボットのための異常検出手法”, 信学技報 PRMU2004-144, pp. 61–66, 2004.
- [6] T. Nanri and N. Otsu: “Unsupervised Abnormality Detection in Video Surveillance”, Proc. IAPR Conference on Machine Vision Application, pp. 574–577, 2005.
- [7] S. J. McKenna and H. Nait-Charif: “Learning Spatial Context from Tracking using Penalised Likelihoods”, Proc. IEEE International Conference on Pattern Recognition, ICPR2004.
- [8] L.M.Manevitz and M.Yousef: “One-Class SVMs for Document Classification”, Journal of Machine Learning Research 2, pp.139–154, 2001.
- [9] 小野田, 村田, 山名, 清水, 野邊, 村川: “One Class SVM に基づく水力発電所軸受異常振動の予兆発見 (1)”, 第 18 回人工知能学会全国大会, 2F2-01, 2004.
- [10] B. Scholkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola and R. C. Williamson: “Estimating the support of a high-dimensional distribution”, Neural Computation 13, pp.1443–1471, 2001.
- [11] 麻生, 津田, 村田: “パターン認識と学習の統計学”, 岩波書店, 2003.