

複数人動画像からの異常動作検出

南 里 卓 也[†] 大 津 展 之^{†,††}

セキュリティ分野における映像監視や老人介護のモニタリングシステムなどにおいて、異常動作の検出は非常に重要な課題である。そこで、本論文では、固定カメラによる複数人動画像からの異常動作検出のための教師なし手法を提案する。本手法では、画面内で頻繁に行われる動作を通常動作とし、異常動作をその通常動作の特徴分布から逸脱するものとして定義する。動作特徴として立体高次局所自己相関特徴を用いており、この特徴の加法性の性質と固有空間法の線形性がうまく組み合わせることによって、画面内に複数人いる場合でも、個々の人物の切り出しやトラッキングをすることなく容易に学習および検出が可能となる。しかも対象に関する先見知識もいっさい必要としない。実験では、複数人の歩行者の中での転ぶ動作を異常動作として検出し、手法の有効性を確認した。

Anomaly Detection in Motion Images Containing Multiple Persons

TAKUYA NANRI[†] and NOBUYUKI OTSU^{†,††}

The detection of anomaly (abnormal/unusual) movements is an important problem in video surveillance applications. We propose an unsupervised method for anomaly movement detection in scenes containing multiple persons. Our method uses cubic higher-order local autocorrelation (CHLAC) to extract movement features. We show that the additive property of CHLAC in combination with a linear eigenspace method is well suited to simplify the learning of usual movements and to detect anomaly movements even in scenes containing multiple persons. One particular advantage of this method is that it does not necessitate the object segmentation and tracking, and also any prior knowledge about objects. Some experimental results are shown to exhibit the validity of the method.

1. はじめに

近年、セキュリティ対策としてカメラによる監視システムが広く利用されているが、人による動画像からの異常検出は労力も多大であり、これをコンピュータで代替できるならば大幅な労力削減につながる。また老人介護においても、発作などを検出する自動通報システムにより、介護者の負担を減らすことができる。したがって、異常を検出し知らせるなどの知的機能を持ったカメラによるモニタリングシステムのニーズは高い。

監視カメラを想定した場合、まず第1に画面内に複数の人間がいることを考慮しなければならない。動作認識に対する通常のアプローチでは、前処理として

それぞれの人間をセグメンテーションし、トラッキングすることになる^{1),2)}。このアプローチでは、認識の精度がセグメンテーションやトラッキングの精度に強く依存してしまう。また、ビデオサーベイランスにおける動画像解析^{10),12)}では、抽出する特徴が対象の動作特徴だけでなく、画面内の対象の位置に依存したものであるため、複数人の動作の場合には非常に扱いにくい。

それに対して、本手法の特徴抽出では、複数人の動作特徴として立体高次局所自己相関(CHLAC)特徴⁴⁾を用いる。この特徴は、二次元 (x, y) データからの特徴である高次局所自己相関(HLAC)特徴⁵⁾を三次元 (x, y, z) の立体データへ拡張したものであり、動画像のような時間を含む三次元 (x, y, t) 立体データに対する特徴抽出となる。CHLAC特徴の利点としては、セグメンテーションやトラッキングを行うことなく複数人の動作特徴を抽出できる点であり、計算量は対象の数によらずつねに一定となる。また、動作特徴を用いた多くのアプローチでは、対象に対する先見知識が仮定されている^{8),9)}のに対し、CHLAC特徴は対象の先

[†] 東京大学大学院情報理工学系研究科

Graduate School of Information Science and Technology, The University of Tokyo

^{††} 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology (AIST)

見知識を必要としない基本的な統計的特徴であり、汎用性が高い。

次に、異常動作を検出する段階においては、異常動作を学習させることが必要となる。しかし、サンプルの少ない異常動作をすべて前もって学習させることはほぼ不可能であるため、異常動作を直接明示的に学習させることは難しい。そこで、辞書においても定義されているように、異常動作を「通常動作ではないもの」として定義する。ここで通常動作とは、カメラの前で頻繁に起こる動作を表す。通常動作は、前もって教師なしで統計的に学習することができるため、異常動作はそれからの逸脱として容易に検出することができる。

そのような統計的なアプローチを用いているものとして、SVM で確率分布を推定することによって外れ値を検出するもの⁷⁾や、時系列に分割した動画画像特徴の共分散行列から異常動作を検出するもの¹²⁾などがある。しかし、これらは、対象が複数あるような場合にはそのまま適用することができない。

そこで本論文では、異常動作の検出に、CHLAC 特徴の性質をうまく利用することのできる固有空間法を用いる。CHLAC 特徴は積分特徴であるため、画面内での対象に関する加法性があり、その加法性と固有空間法の線形性を組み合わせることによって、複数人の動作が同時に存在する場合でも、すべての通常動作は通常動作の部分空間内に閉じ、異常動作のみを検出することができる。さらに、通常動作は教師なしで学習されるため、どのような動作が存在する場面においても通常動作として適応的に学習することができ、汎用性を持つ。

本論文ではさらに、オンライン学習のため、通常動作部分空間の逐次的な学習についても考察する。そのための手法として、主成分分析 (PCA) の固有値問題をステップごとに解く手法と、固有値問題を解かずに逐次的に固有ベクトルを推定する手法³⁾の2つを提案し、実験により比較する。また、後者の手法を、より頑健にした手法も提案し、実験により比較する。

2. 提案手法

2.1 前処理：動きの検出

本手法では静止したカメラを想定し、動き部分の検出に、ピクセルごとの時間差分画像を用いる。画像は、解像度 352×240 のグレイスケール (256 階調) である。時間差分画像ではアナログ値はあまり意味をなさず、変化 (動き) あり (1) なし (0) の二値のみが重要となる。そのため二値化をする必要があるが、時間差分画像は値が 0 以外のピクセルすべてが対象物体の

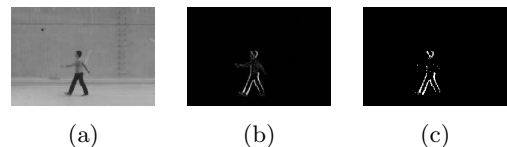


図 1 前処理。(a) 原画像、(b) 差分画像、(c) 二値画像
Fig. 1 Preprocessing, (a) original image, (b) difference image, (c) binarized image.

動きを表しているわけではなく、光の変化などによるノイズも含まれている。そこで、ノイズを除去するために適切な閾値を決める必要があり、その閾値決定にはたとえば判別二値化手法⁶⁾を用いて二値化を行う。二値化の結果を、図 1 に示す。

2.2 立体高次局所自己関連特徴

このように画素ごとの動き情報を与える時系列二値画像からの動作特徴の抽出には、立体高次局所自己関連 (CHLAC) 特徴⁴⁾を用いる。この特徴は、二次元データに対する特徴である高次局所自己関連 (HLAC) 特徴⁵⁾を時間方向に拡張し、三次元データへ適用可能としたものである。定式化としては HLAC 特徴とほぼ変わらず、各成分である N 次局所自己関連関数は次のように表される。

$$x_f^{(N)}(a_1, \dots, a_N) \triangleq \int f(r)f(r+a_1)\dots f(r+a_N)dr \quad (1)$$

ここで、 f は時系列画像を表し、変数 r (参照点) と N 個の局所変位 a_i ($i = 1, \dots, N$) は、画面内の座標 x, y と時間 t を成分として持つ三次元のベクトルである。さらに、積分範囲は $W \times H \times T$ の立体とし、 W, H は画像の幅と高さ、 T は特徴を抽出する時間方向の窓幅を表す。

本手法では、CHLAC 特徴の次数 N を 0, 1, 2 とする。 $N = 0$ のときは、 f の単純な三次元画像内の積分であり、 f が二値画像の場合は、画素値が 1 のピクセルの数を表す。 $N = 1$ のときには、変位ベクトルは a_1 のみとなるので、2 つのピクセル間の相関を表し、画像境界を無視して積分が同値とならないための変位ベクトル a_1 の数は、 $3 \times 3 \times 3$ 近傍の場合 13 個となる。図 2 の記号で表すと、参照点を r' として、たとえば $h, h', h'', a, a', a'', b, b', b'', g, g', g'', r$ の 13 個である。積分が同値とならないためというのは、たとえば、 r' と h の相関、 r' と d' の相関の積分は同値となるため、そのような冗長な変位ベクトルは省くという意味である。同様に、 $N = 2$ のときは 237 個の組合せを持つ。図 2 に、 $N = 2$ のときの変位ベ

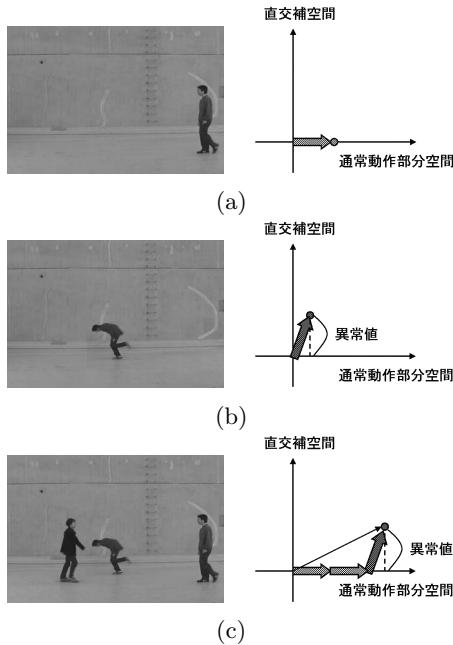


図4 CHLAC 特徴の加法性と部分空間の性質。(a) 通常動作(歩く), (b) 異常動作(転ぶ), (c) 通常動作と異常動作

Fig.4 Additive property of CHLAC and linear subspace, (a) usual movement (walk), (b) anomaly movement (tumble), (c) both movements.

であるかどうかの指標である異常値として定義する。射影子を用いることによって、複数人の場合でも異常動作のみが検出されることを数式で表すと、次のようになる。

$$\begin{aligned}
 \mathbf{x} &= \mathbf{x}_1^{(U)} + \dots + \mathbf{x}_n^{(U)} + \mathbf{x}^{(A)} \quad \text{とすると} \\
 (U : \text{usual}, A : \text{anomaly}) \\
 \|P_{\perp} \mathbf{x}\| &= \|P_{\perp}(\mathbf{x}_1^{(U)} + \dots + \mathbf{x}_n^{(U)} + \mathbf{x}^{(A)})\| \\
 &= \|P_{\perp}(\mathbf{x}_1^{(U)} + \dots + \mathbf{x}_n^{(U)}) + P_{\perp} \mathbf{x}^{(A)}\| \\
 &\simeq \|P_{\perp} \mathbf{x}^{(A)}\| > 0 \quad (4)
 \end{aligned}$$

また、図による直観的な説明を図4に示す。

2.4 オンライン学習

実世界への応用では大量のデータを扱うため、すべてのデータを保持しておくことは難しく、前節の定式化は適用できない。そこで、オンライン学習が必要となる。そのためには、逐次的に通常動作の部分空間を学習、更新しなければならない。本論文では、ステップごと(あるいは数ステップごと)に固有値問題を解く手法と、固有値問題を解くことなく固有ベクトルを逐次更新する手法(CCIPCA)³⁾の2つを適用し、比較する。

前者の手法は、前節の定式化に忠実ではあるが、ステップごとに固有値問題を解くため計算量が多いとい

う問題がある。固有値問題に必要な自己相関行列 R_X は、次のように更新される。

$$R_X^{(n)} = \frac{n-1}{n} R_X^{(n-1)} + \frac{1}{n} \mathbf{x}_n \mathbf{x}_n^T \quad (5)$$

ここで $R_X^{(n)}$ は n ステップ目の自己相関行列であり、 $\mathbf{x}_n$ は n ステップ目の入力ベクトルである。

後者の手法は、相関行列を求めない逐次的な主成分分析(CCIPCA)³⁾を用いた手法である。このアルゴリズムでは、ステップごとに固有値問題を解く必要がないため、非常に高速な手法となっている。また、この手法では、固有値の収束はあまり良くないが、固有ベクトルの収束は速いという性質を持つ。第1固有ベクトルと第1固有値は、次のように更新される。

$$\mathbf{v}_n = \frac{n-1}{n} \mathbf{v}_{n-1} + \frac{1}{n} \mathbf{x}_n \mathbf{x}_n^T \frac{\mathbf{v}_{n-1}}{\|\mathbf{v}_{n-1}\|} \quad (6)$$

ここで、固有ベクトルは $\mathbf{v}/\|\mathbf{v}\|$ であり、固有値は $\|\mathbf{v}\|$ である。この更新則では、 $n \rightarrow \infty$ のときに $\mathbf{v}_n \rightarrow \pm \lambda_1 \mathbf{e}_1$ となることが証明されている¹¹⁾。ここで λ_1 は、サンプルの相関行列 R の最大固有値であり、 $\mathbf{e}_1$ はそれに対応する固有ベクトルである。第 n 固有ベクトルと第 n 固有値については、第1固有ベクトルと第1固有値から Gram-Schmidt の直交化法に則って漸次的に更新され、真の固有値、固有ベクトルに収束することが示されている。

この後者の手法では、さらに、収束を速めるためのパラメータ w を導入している。これは、式(6)のようにすべてのサンプルを等価に学習するのではなく、学習の初期段階のノイズをより速く減少させることを目的としている。更新式は次のように表される。

$$\mathbf{v}_n = \frac{n-1-w}{n} \mathbf{v}_{n-1} + \frac{1+w}{n} \mathbf{x}_n \mathbf{x}_n^T \frac{\mathbf{v}_{n-1}}{\|\mathbf{v}_{n-1}\|} \quad (7)$$

ここで、第1項と第2項の係数の和は1となるように設定されており、 w の値としては、2から4程度が良いとされている。

しかし、この更新式では、ノルムの大きな例外値が入力された場合に、結果としてこの例外値に強く引っ張られた固有ベクトルになってしまう。また、式(6)においても、入力ベクトル $\mathbf{x}_n$ の二次の相関行列を用いて更新されているために、ノルムの大きな例外値には強く影響されることになる。そこで、頑健に通常動作の部分空間を推定するために、式(7)のパラメータ w を入力ベクトルに依存した変数に変更する。すなわち、入力ベクトルを一定の重みで学習するのではなく、通常動作の部分空間から遠い入力ベクトルに対しては重みが小さくなるように、次のようにパラメータを設定する。

$$w(x_n) = k \exp\left\{-\frac{\|P_{\perp}x_n\|^2}{2\tilde{\sigma}^2}\right\} - 1 \quad (8)$$

ここで、 k は例外サンプルを無視する割合（程度）を表すパラメータであり、 $\tilde{\sigma}$ は、通常動作部分空間からの距離がどの程度のデータまで学習に用いるか、を表すパラメータである。式 (8) が、通常動作部分空間からの距離 $\|P_{\perp}x\|$ 、すなわち異常値を用いた正規分布の形で表されているのは、入力ベクトルがすべて通常動作であると考えた場合、 $P_{\perp}x$ は、平均 0、分散 $\sigma^2 I$ のノイズであると考えられるためである。ここで σ は、 $P_{\perp}x_i \in V^M$ ($i = 1, \dots, n$) の共分散行列を Σ としたときに $\sigma = \text{trace}(\Sigma)/M$ として求め、また、式 (8) において $\tilde{\sigma} = 3\sigma$ とすることで、分布よりやや広い範囲の入力ベクトルを学習に用いることとなる。 $w(x)$ は、 -1 から $k-1$ の範囲の値をとる。

3. 実験

実験では、まず第 1 に、画面内に 1 人存在する場合、複数人存在する場合の 2 つの場合に対して、学習フェーズ、検出フェーズを設けて本手法の実験を行った。次に、画面内に複数人存在する場合に対して、フェーズを分けずにオンラインで学習する場合の実験を行った。

本実験に用いた特徴のパラメータとしては、Kobayashi ら⁴⁾ の認識結果を用いて、立体高次局所自己相関特徴の画面内変位幅を 10 ピクセル、時間方向の変位幅を 1 とし、積分範囲である時間幅は 20 とし与えた。図 2 の記号を用いると、画面内変位幅は r と a の距離を表し、時間方向の変位幅は r と r' の距離を表す。また、これらの変位幅の値は、対象の動きに対して相関のとれる範囲のものであり、時間積分範囲 T は、対象の動作の周期とおおよそ同等のものである。ただし、これらのパラメータは結果に対してあまり敏感なものではない。

3.1 画面内に 1 人存在する場合

ここでは、異常動作検出の過程を学習フェーズと検出フェーズに分ける。学習フェーズでは通常動作のみのデータを用いて通常動作部分空間を学習し、検出フェーズでは通常動作と異常動作のデータを与えて結果の比較を行った。学習フェーズには、通常動作として「歩く」動作を採用し、被験者 6 人が右方向、左方向にそれぞれ 4 回ずつ「歩く」動作を行ったデータ（計 48 回）を用いた。検出フェーズでは、学習フェーズの 6 人とは別の 1 人が「歩く (walk)」、「走る (run)」、「転ぶ (tumble)」の動作を行ったデータを用いた。動画の例を図 5、図 6 に示す。ここで、学習、検出に用いられる動画は、人が画面中にいない状態からフレーム



図 5 検出フェーズで用いたデータの例

Fig. 5 Examples of data used in a detection phase.



図 6 「転ぶ」動作の例 (20, 30, 40, 70, 80, 90 フレーム)
Fig. 6 Examples of tumble movement (20, 30, 40, 70, 80, 90th frame).

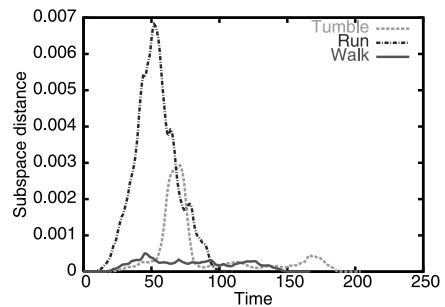


図 7 異常値 (1 人の場合)

Fig. 7 Anomaly value (one person).

インし、フレームアウトするまでのものとなっている。

実験の結果を図 7 に示す。この図は、ステップ（フレーム）ごとの部分空間からの距離（異常値）のグラフであり、通常動作である「歩く」は部分空間からの距離が小さく、異常動作である「走る」「転ぶ」は大きな値であり、異常値として正しく検出されていることが確認された。ここで、寄与率 0.99 となる固有空間（通常空間）の次元は、特徴ベクトル 251 次元のうち 12 次元であった。

3.2 画面内に複数人存在する場合

この場合も、学習フェーズと検出フェーズを設けて実験を行った。本実験では、画面内に最大 3 人存在するように設定し、3 人とも「歩く」動作を行ってもらった。学習フェーズで用いるデータには「歩く」のみのデータを用い、検出フェーズで用いるデータには「歩く」のみのデータと、1 人だけ「転ぶ」データを用いた（図 8）。

実験の結果を図 9 に示す。実験の結果を示すグラフは画面内に 1 人存在する場合の実験のときと同様であり、usual が通常動作の動画、anomaly が異常動作を



図 8 複数人の例．(a) 3 人とも歩く動画，(b) 1 人だけ転ぶ動画
Fig. 8 Examples of multiple persons. (a) three persons' walking, (b) one person's tumbling.

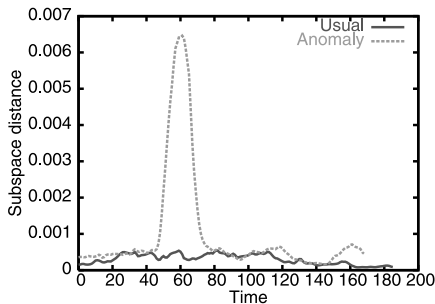


図 9 異常値 (複数人の場合)
Fig. 9 Anomaly value (multiple persons).

含んだ動画である．結果として「転ぶ」フレームにおいて正しく異常が検出された．

ここでは，体の一部が隠れた状態になるフレームインとフレームアウトのときや，人が重なりあう状態のときには，厳密には加法性は成り立たなくなる．しかし，結果としてはそのような条件下においても異常を正しく検出できたといえる．これに対する仮説として，体の一部が隠れた状態や重なりあった状態の特徴量もある程度学習され，固有空間に含まれたと考えることができる．しかし，体の一部が隠れたときや重なりあったときに小さな異常値の変化もあったことから，すべての重なりの特徴量が固有空間に含まれているわけではないことが分かる．そうならば，もし多くの動作を学習することによってすべての重なりの特徴量を学習した場合，そのような重なりの部分にも反応しなくなる可能性がある．

また，寄与率 0.99 となる固有空間の次元としては，特徴ベクトル 251 次元中 5 次元であった．これは，画面内に 1 人存在する場合よりも少ない結果となっており，自然な予想と反する．この理由としては，1 人の場合は人による歩く動作の微妙な違い（歩様）が部分空間に反映されるが，複数人の場合は，それらの歩様成分が平均化されて，共通的な歩く動作成分が加算されて支配的になるため，部分空間の次元がむしろ小さくなっていると考えられる．

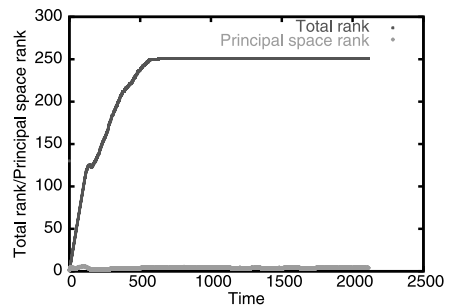


図 10 入力データの累積次元 (rank) と固有空間の次元
Fig. 10 Rank of input data and dimension of principle subspace.

3.3 オンライン学習

本節では，学習フェーズと検出フェーズを設けず，逐次，データから通常動作の部分空間を構成しつつ，現在の動作の評価を行うオンライン学習を行う．用いる手法としては，最初のフレームから現在のフレームまでの自己相関行列を逐次計算し，その固有値問題をステップごとに解くことによってその時点での部分空間を構成，その部分空間に対する距離を測るという手法と，固有値問題を解くことなく固有ベクトルを更新する手法 (CCIPCA) の 2 つを適用した．

データは，前節の複数人が行き交う動画を連続的な動画として用いた．異常動作である「転ぶ」動作は，全体 2,130 フレーム中で約 1,950 フレーム目にあり，それ以外は複数人の「歩く」通常動作のみの動画である．

まず第 1 に，前者の手法を適用した．図 10 は，次元の変化の結果である．Total rank は，そのタイムステップまでの入力データ全体の累積次元 (ランク) を示しており，フルランクである 251 次元に到達するのは，629 フレーム目となっている．120 フレーム目までは，1 つのサンプルが増えるにつれてランクも 1 つずつ増加しているが，その後はややなだらかな変化をしている．通常動作部分空間の次元に関しては，つねに 4 次元程度で変化しており，割合としては 1% 程度と非常に小さいものとなっている．これは，前節において考察した，歩様成分の平均化が原因であると考えられる．

また，図 11 は，通常動作部分空間からの距離の時間変化のグラフである．この図において，初めの方のフレームで大きな値が出ているが，動画では初めて 3 人が交差するというシーンに対応しており，部分空間の学習がまだ完全ではない状態であるといえる．中盤の変化は，画面内に人がいるか否かでやや値が変化しているが小さい値で安定しており，最後の「転ぶ」フレーム (1,950 フレーム) では正しく異常値を検出で

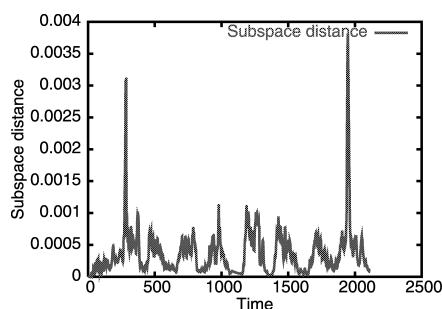


図 11 異常値の変化 (オンライン学習)

Fig. 11 Anomaly value (online learning).

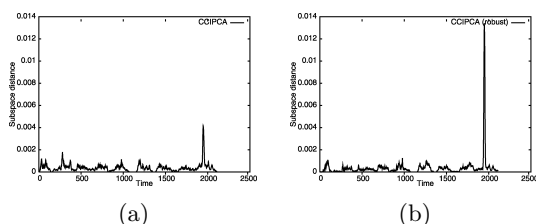


図 12 異常値の変化. (a) CCIPCA, (b) CCIPCA (robust)

Fig. 12 Anomaly value. (a) CCIPCA, (b) CCIPCA (robust).

きている。

次に、CCIPCA を用いて実験を行った。図 12 は、CCIPCA をそのまま適用した手法と、入力ベクトルにより重みを変える、より頑健な CCIPCA を適用した手法の、通常動作部分空間からの距離の時間変化のグラフである。後者の手法において、パラメータの重みを表す係数 k の値は 5 とした。これらの手法は、漸近的に通常動作の部分空間を推定する手法ではあるが、異常動作を正しく検出していることが分かる。また、例外値を学習しないように変更した CCIPCA の方が、異常動作の値をより大きく検出した。これは、例外値に影響されることなく、より適切に通常動作の部分空間を学習できたためといえる。

本実験においては、通常動作部分空間の次元は、ステップごとに寄与率から求めるのではなく、ある時間幅の入力ベクトルに対して一度固有値問題を解くことによって求めた。ここでは、次元は 4 とした。これは、さきほどの実験から部分空間の次元がほとんど変化していなかったため、初めに一度だけ次元を求めたとしてもあまり悪い近似ではないからであるが、他に、次のような 2 つの理由があげられる。まず第 1 に、計算量が大きいという点である。寄与率を求めるためには、すべての固有値を推定しなければならず、その計算時間は、特徴抽出の計算時間を含めない場合、Pentium 4 3.06 GHz において 21 sec/frame であった。逆に、次

元を定数の 4 として計算すると、0.0039 sec/frame となり、特徴抽出の計算時間である、0.0174 sec/frame と合わせても実時間で可能な処理時間となる。2 つ目の理由としては、この CCIPCA の手法において固有値の収束が遅いという点である。実験では 2,130 フレームのデータを用いたが、最終的に通常動作部分空間の次元は 211 程度であり、収束値である 4 にまったく収束していないことが分かる。これらの理由により部分空間の次元を一定として求めたが、結果はむしろ良いものとなっており、計算時間のことを考慮すると有効な手法であると考えられる。

4. おわりに

4.1 ま と め

本論文では、加法性を持つ汎用的な動き特徴である CHLAC 特徴と固有空間法を用いて、通常動作を統計的に部分空間として学習し、異常動作をそこからの逸脱として容易に検出することができることを示した。この手法は、複数人の場合にもそのまま適用可能であり、画面内で 1 人でも異常動作を行えば、検出することができることを示した。しかも、対象の切り出しは不要で、計算量は人数によらず一定であり、有効で実用性の高い手法となると思われる。また、この手法では、通常動作を陽に定義することなく統計的に（教示なし）学習しているため、設計の段階で通常動作とは何かについて定義する必要がなく、応用に則した自然な検出を行うことができる。さらに、対象についての仮定をまったくおいていない点で、人の動作に限らず、さまざまな環境および課題に対しても、通常・異常を判別できる汎用的な手法であるといえる。また、実時間処理への適用のため、学習フェーズを設定せずに、オンライン学習によっても同様に異常動作を検出することができることも示した。

4.2 今後の課題

本論文では、通常動作が単一モードの場合の実験を行ったが、2 つのモードがある場合、あまり正しく通常動作部分空間を構成することができていない。本手法では、1 つの部分空間で通常動作部分空間を近似しようとしているため、分布が多峰性の場合には、大きな分布の間の谷に異常動作が含まれる場合など、やや限界があると考えられる。そのため、通常動作部分空間内においてもクラスタリングを行った後、それらからの距離を測るなど、多峰性の分布にも対応できるよう改良することが必要である。

また、カメラの視点としては、現状では対象が画像平面に対して平行に動くシーンへの適用に限定して

いるが、対象がカメラに向かって動く場合にも原理的には適用可能である。今後、そのような状況設定も含め、複雑な背景や多様な動作など、さまざまな実環境状況下での本手法の有効性を評価していきたいと考えている。

最後に、本手法は、人の動作に限らず正常・異常を判別することができる汎用的な手法であると考えられることから、今後さらに、種々のアプリケーションを通して本手法の適用範囲の拡大を図っていく予定である。

参 考 文 献

- 1) Efros, A.A., Berg, A.C., Mori, G. and Malik, J.: Recognizing Action at a Distance, *Proc. IEEE Int. Conf. on Computer Vision*, Vol.2, pp.726–733 (2003).
- 2) Haritaoglu, I., Harwood, D. and Davis, L.S.: W4: Who? When? Where? What? A Real Time System for Detecting and Tracking People, *Proc. IEEE Int. Conf. on Face and Gesture Recognition*, pp.222–227 (1998).
- 3) Juyang, W., Yilu, Z. and Hwang, W.: Candid Covariance-Free Incremental Principal Component Analysis, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.25, No.8, pp.1034–1040 (2003).
- 4) Kobayashi, T. and Otsu, N.: Action and Simultaneous Multiple-Person Identification Using Cubic Higher Order Local Auto-Correlation, *Proc. IAPR Int. Conf. on Pattern Recognition*, Vol.4, pp.741–744 (2004).
- 5) Otsu, N. and Kurita, T.: A New Scheme for Practical Flexible and Intelligent Vision Systems, *Proc. IAPR Workshop on Computer Vision*, pp.431–435 (1988).
- 6) 大津展之：判別および最小2乗規準に基づく自動しきい値選定法，電子通信学会論文誌D，Vol.J63-D，No.4，pp.349–356 (1980).
- 7) Schölkopf, B., Platt, J., Shawe-Taylor, J., Smola, A. and Williamson, R.: Estimating the Support of a High-Dimensional distribution, *Neural Comput.*, Vol.13, No.7, pp.1443–1471 (2001).
- 8) Song, Y., Goncalves, L. and Perona, P.: Unsupervised Learning of Human Motion, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.25, No.7, pp.814–827 (2003).
- 9) Yacoob, Y. and Black, M.J.: Parameterized Modeling and Recognition of Activities, *Computer Vision and Image Understanding*, Vol.73, No.2, pp.232–247 (1999).
- 10) Zelnik-Manor, L. and Irani, M.: Event-Based Analysis of Video, *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Vol.2, pp.123–130 (2001).
- 11) Zhang, Y. and Weng, J.: Convergence Analysis Complementary Candid Incremental Principal Component Analysis, Technical Report MSU-CSE-01-23, Computer Science Department of Michigan State University (2001).
- 12) Zhong, H., Shi, J. and Visontai, M.: Detecting Unusual Activity in Video, *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Vol.2, pp.819–826 (2004).

(平成 17 年 1 月 10 日受付)

(平成 17 年 7 月 18 日採録)

(担当編集委員 日浦 慎作)



南里 卓也

平成 16 年東京大学工学部機械情報工学科卒業。現在、東京大学大学院情報理工学系研究科知能機械情報学専攻修士 2 年。平成 17 年電子情報通信学会パターン認識とメディア理解研究会研究奨励賞受賞。コンピュータビジョン、パターン認識の研究に従事。電子情報通信学会学生員。



大津 展之(正会員)

昭和 44 年東京大学工学部計数工学科卒業。昭和 46 年東京大学大学院計数工学数理コース修士課程修了。同年電子技術総合研究所入所。パターン認識の理論と応用研究に従事。カナダ NRC 客員研究員，情報科学部情報数理研究室長を経て平成 2 年首席研究官。平成 2 年知能情報部長に就任。平成 3 年筑波大学連係大学院教授併任。平成 13 年産業技術総合研究所フェロー，東京大学大学院情報理工学系研究科教授併任，現在に至る。工学博士。昭和 53 年電子通信学会学術奨励賞，昭和 63 年行動計量学会賞優秀賞，平成 10 年 IAPR/MVA'98 Most Influential Paper of the Decade Award，平成 14 年映像情報メディア学会丹羽高柳業績賞等を受賞。行動計量学会，分類学会等の会員。