

スケール変化に対応した人物動作判別

石井 健幸^{*1}, 村上 仁己^{*2}, 小池 淳^{*2}

Scale Invariants Identification Scheme of Human Behavior

Takeyuki ISHII^{*1}, Hitomi MURAKAMI^{*2}, Atsushi KOIKE^{*2}

ABSTRACT : This paper proposes a scale invariant identification scheme of human behavior, so that distinguishing abnormal behavior to normal behavior. This scheme makes use of CHLAC(Cubic Higher order local auto correlation), which was proposed by Otsu[5], as the feature quantities to describe moving behavior of human. In order to cope with change of human body scale in input moving images, our proposed scheme adopts an average processing for scale variation of human in feature quantities calculation. Computer simulation using moving images of human is performed with the proposed scheme. The experimental results showed that the proposed scheme has a better performance than these conventional schemes in detection of abnormal human behaviors.

Keywords : pattern recognition , abnormal detection, CHLAC(Cubic Higher order local auto correlation)

(Received March 22, 2012)

1. はじめに

近年、コンピュータの発達や高度情報化社会に伴い、動画からの物体や人物認識に関する研究が盛んに行われている[1-2]。その中でも人物の動作の解析に関しては、特に高い注目が注がれている。人物解析の研究は、老人介護における被介護者の異常検知や、犯罪防止のための監視システムを実現するための必要不可欠な技術である。

人物の動作を記述するための有効な特徴量の一つに、立体高次自己相関特徴量(CHLAC)[5]が提案されている。これは、時系列に並べられた動画を、マスクパターンに沿って251次元の特徴量ベクトルに特徴量化する手法である。

この特徴量を用いた人物動作解析の応用の一例に、異常動作(逸脱行為)の検出[2]がある。人物動画から得られたCHLAC特徴量が、通常動作(動画中に人物が頻繁に行う動作)から事前に求めたその特徴量との類似度を計算して判断する。具体的には、人物動画から得られたCHLAC特徴量が、通常動作を表現する部分空間の

補空間(つまり、異常動作を表す部分空間)への射影成分が閾値以上であるものを異常動作として判定する。

この手法では、人物の横方向への歩行動作など、対象人物の大きさが変化しない場合は、通常動作における特徴量変化は少ないため、判別は容易である。しかし、カメラに対して人物が前後に動くなど、画像中の人物の見かけの大きさが変化する動画においては特徴量そのものが大きく変化する傾向がある。そのため、通常動作と異常動作の判別が正しく行われぬ可能性がある。

これに対する改善手法の案として、松永ら[3]が提案するマルチスケールCHLACや画像ピラミッドの手法がある。前者は、画素間隔及び時系列間隔を人物の大きさの変化に合わせて適応的に変更することにより、後者は画像を階層的に表現することによって、人物の大きさが変化しても対応できるようにしている。

本研究では、画像中の見かけ上の人物の大きさの変化が大きい動画に対して、時系列上で特徴量の平均化処理を行うことによって、人物の大きさの変化による特徴量の急激な変化を最小限に抑える手法を提案する。これにより、人物の大きさが変化する場合においても、安定的に人物の異常動作の判別が可能となる。コンピュータによる実験を行い、横方向への人物移動に加えて、人物

*1 : 理工学研究科理工学専攻学生

*2 : 情報科学科教授 (koike@st.seikei.ac.jp)

が前後に動く動画像において、人物動作の判別が正しく認識できることを示す。

2. 従来手法

2.1 時系列動画像

本提案手法では、動領域部分の検出に画像のフレーム間差分を用いる。入力動画像中の時間的に前後する2枚の画像フレームに対して、画素毎に差分を取り、その絶対値を計算する。このとき、得られた差分画像に対して、閾値処理を行って静止領域と動領域からなる2値画像を得る。

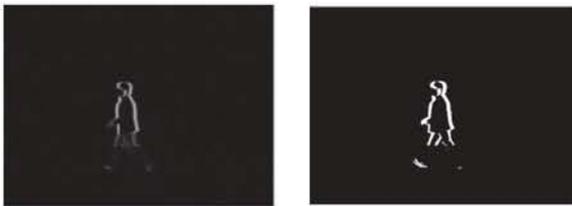


図1：背景差分画像（左）と2値化画像（右）の例

2値化処理のための閾値は、2値化後の領域間の分散が最大になる閾値決定手法[4]を用いた。図1に差分画像と2値化処理後（動領域が白点で表示）の結果を示す。

2.2 立体高次自己相関特徴

本研究では、動きを記述する特徴量として小林と大津によって提案された立体高次局所自己相関(CHLAC)特徴[5]を用いる。CHLAC特徴量は、対象の切り出しが不要でかつモデルを仮定しないため汎用的な手法である。

CHLAC特徴量は、静止画(2次元)に対して汎用的でかつ高速、高精度な認識手法を実現できる高次局所自己相関(HLAC)特徴[6]に、時間方向を加え3次元に拡張した特徴である。画像を時系列に並べたデータに対して、251種類の局所的な自己相関特徴を計算する。これは、画像系列全体にわたって、あるいは区間毎に積分することにより特徴量が得られる。ここで、時系列画像を f 、動画像中の位置をとすると、 N 次の立体高次局所自己相関関数は以下のように定義される。

$$r_f^N(a_1, \dots, a_N) = \int f(r)f(r+a_1) \dots f(r+a_N)dr \quad (1)$$

ここで、 r_f は時系列画像であり、変数 r 、 N 個の変位 $a_i(i=1, \dots, N)$ は画像内の二次元座標と、時間方向を成分として持つ3次元のベクトルである。

変位ベクトルの数は、HLAC特徴では、次数 N を高々2までとし、変位方向 a を 3×3 の局所領域に限定する。平行移動により等価となる特徴を除くと、変位方向 a の

組み合わせは2値画像の場合で0次が1個、変位方向 a の組み合わせは2値画像の場合、0次が1個、1次が4個、2次が20個、合計で25個となり、HLAC特徴は25次元のベクトルとして得られる。CHLAC特徴においても、次数 N を高々2までとし、変位方向 a を $3 \times 3 \times 3$ の局所領域に限定する。さらにHLAC特徴と同様に、平行移動により等価となる特徴を除くと、変位方向 a の組み合わせは2値画像の場合、0次が1個、1次が13個、2次が237個、合計で251個となり、CHLAC特徴は251次元のベクトルとして得られる。

図2はHLAC特徴量のマスクパターンであり、別表1[7]はCHLAC特徴量を用いたマスクパターンを示す。

CHLAC特徴量を用いた認識方式は、動画像からの異常動作の検出[2]において高い認識性能が報告されている。

2.3 部分空間法

CHLAC特徴量を用いた異常動作検出の手法として、部分空間法の手法が用いられている。動画像内で人物が異常動作を行った場合、方向の異なるベクトルが加わるため、ベクトルが部分空間から飛び出し、部分空間からの距離として異常を検出することができる。

CHLAC特徴量を用いた部分空間法では、まず得られたCHLAC特徴量のベクトルを基にして、通常動作の部分空間を構成するために、主成分分析により主成分ベクトルを求める。主成分ベクトル $U = [u_1, u_2, \dots, u_M]$ 、

$$u_i \in V_M (i=1, \dots, M) \text{は、共分散行列} \Sigma_x \text{を用いて、}$$

$$\Sigma_x U = U \Lambda \quad (2)$$

の固有値問題より求まる。

固有値行列 $\alpha = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ において、累積寄与率 η_k は、

$$\eta_k = \frac{\sum_{i=0}^k \lambda_i}{\sum_{i=0}^M \lambda_i} \quad (3)$$

と表され、累積寄与率 η_k が0.99となる次元までの固有ベクトル (u_1, u_2, \dots, u_k) により張られる空間を、通常動作の部分空間として適用している。このようにして得られた主成分直交基底 $u_k = [u_1, u_2, \dots, u_k]$ によって張られた部分空間への射影子は、 $P = u_k$ として表され、それに対する直交補空間への射影子は $P_{\perp} = I_M - P$ となる。直交補空間への射影成分の距離 d は、

$$\begin{aligned} d_{\perp}^2 &= \|P_{\perp} x\|^2 \\ &= x^T (I_M - U_k U_k^T) \end{aligned} \quad (4)$$

と表すことができる。

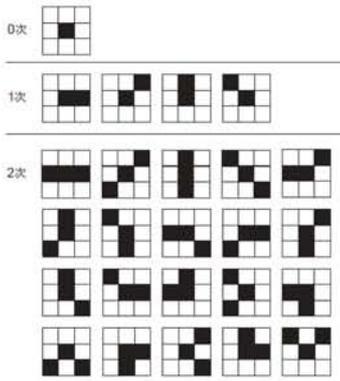


図2: HLAC 特徴のマスクパターン

2.4 マルチスケール CHLAC

CHLAC 特徴では、1画素毎に時間軸で1フレーム毎にマスクパターンを調べて計算を行っている。これに対して、マルチスケールCHLACでは、対象動画を複数のスケールで計算する。得られた特徴をすべて同時に評価することで、異常動作判別の精度をより高めることが可能となる。

一方、画像ピラミッドを用いたCHLACでは、高解像度の画像から低解像度の画像までのいくつかの異なる解像度の画像の集合として構成される画像群を用いる。画像ピラミッドを対象動画より構成し、それぞれの解像度の画像から得られる特徴を一度に評価する。

この手法では、低解像度の画像からは大域的な時間で見えた動きの特徴が得られ、高解像度の画像からは細部での細やかな動きが得られる。人物動作の動きをより細かく判別することが可能である。

3. 提案手法

今回我々が提案する手法は、CHLAC 特徴と部分空間法を基にした人物動作異常検出の手法をベースとして、スケール変化を考慮した特徴量処理と、特徴量の平均化処理を施したものである。

3.1 スケール変化を考慮した特徴量抽出

CHLAC 特徴は、時間軸における画像中の点（白の部分）の推移を特徴量として表すことができる。画像中の人物スケールが増大する場合、画像から得られる人物のスケールも増大するため、それにしただって点の数が増大する。その場合、本来は適合されないはずのマスクパターンが適用され、特徴量が大きく変化し、人物動作の判別が正しく行われな可能性ある。

そこで大きい点（ある点の周辺画素も白であること）

を一つの点としてみなしてマスクパターンの抽出を行う。これによって、CHLAC 特徴による不必要な情報を除外することができる。

この特徴量手法では、通常は3×3ピクセルで行われるマスクパターン領域を、9×9、15×15の3パターンに分割する。9×9、15×15のパターンの場合は、領域内を9つのブロックに分割する。したがって9×9のパターンの場合は1ブロックにつき3×3、15×15のパターンであれば1ブロックにつき5×5の領域となる。

ブロック内で、点がブロック中の半数（9×9の場合は1ブロック5点以上、15×15の場合は1ブロック8点以上）がある場合は、そのブロックは特徴量があるということになる。図3に拡大比率に応じた特徴量のパターン図を示す。

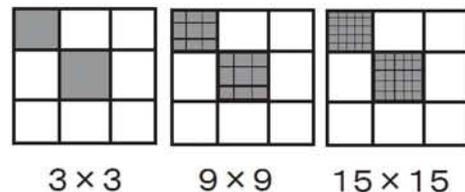


図3: 拡大特徴量の3パターン

例えば図5に示すようなマスクパターンを調べるとして、使用するパターンがそれぞれ9×9、3×3、3×3となったとする。この場合、画像2の3×3のピクセルの中心と、画像の位置の9×9の領域のうち、3×3の箇所が0であるかを調べればよい。

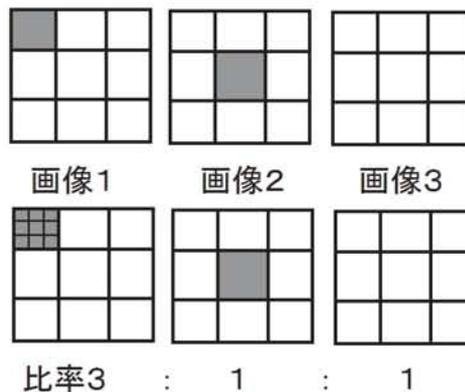


図4: マスクパターン例と(上), 比率3 : 1 : 1のパターンの場合のマスクパターンの手法

特徴量抽出の際は、入力画像で用いる3フレームそれぞれの画像に含まれる点（0次特徴量）の大きさで、3つの拡大パターンのうちどれを使うか決定する。特徴量が多い画像は画像中の人物スケールが大きくなっている（カメラに近づいている）と推測できるからである。

3.2. 特徴量平均化処理

CHLAC 特徴量の場合、画像中の点をそのまま特徴量としてとらえるため、物体のスケールが大きければ大きいほど多くの特徴量を得られる。入力画像が単純に前後に動いている場合、CHLAC 特徴量によって得られる特徴量ベクトルは、ベクトルの方向は変わらず、スケール分だけ増大、もしくは減少した値が得られると考えられる。

そこで、各画像の特徴量の増大傾向の推移を調べ、前後画像に対するベクトルの拡大率を推定し、各々の特徴量ベクトルの大きさを拡大率に従って平均化処理を行えば、画像中人物のスケール変化に関係なく人物行動を判別することが可能であると考えられる。

この特徴量平均化処理では、CHLAC 特徴量による特徴抽出によって得られた特徴ベクトルを基準として、

$$\|\alpha_i \mathbf{x}_i - \mathbf{x}_{i-1}\| \quad (\alpha_1 = 1, i = 1, 2 \dots N) \quad (5)$$

が最小をとる α_i の値を求める。ここで、 α_i は 2 つのベクトル間の拡大率を示す。

これにより、得られた α_i を基にして、各ベクトルを \mathbf{x}_1 に近似するようベクトルの平均化処理を行う。平均化処理されたベクトルを \mathbf{X}_i として、

$$\mathbf{X}_i = \frac{(\mathbf{x}_i + \alpha_{i+1} \mathbf{x}_{i+1})}{2} \quad (i=0, 1 \dots N-1) \quad (6)$$

としてベクトルを再構成する。

このスケール変化に伴う特徴量変化を最小限に抑える平均化手法を用いることにより、ベクトルのスケール変化に伴う特徴量の変化をなくし、異常動作の判別をより容易にすると考えられる。

4. 動作実験と評価

4.1 実験概要

シミュレーション実験では、画像フレーム中に人物が転倒する動作を含んだ人物 1 名が歩行動作を行っている動画シーンを使用した。

動画シーンは、横方向への人物が移動するシーン及び、人物が手前から奥に移動するシーンなどが含まれている。

今回の実験では、人物の横方向への移動における異常検出だけでなく、人物が前後に動くなど、画面中の人物の大きさ変化した場合においても、通常動作と、異常動作が正しく判別されているか、従来手法と提案手法の両方で検証を行った。

動画シーン中に含まれている、スケール変化を伴う人物歩行動作の例を図 5 に示す。また、動画像に含まれ

ている転倒動作のフレーム例を図 6 に示す。



図 5 人物動作中の通常歩行動作の例



図 6 動画像中の異常動作（転倒動作）の例（動画像中 50 フレーム目）

4.2 異常動作フレーム確認

1 つの動画像を用いて、動作が正しくなされているか検証するため、部分空間法による射影成分距離である d 値のフレームごとの推移を調べた。

d 値は、通常動作中はほぼ同じ値を示すが、動画像中に転倒などの異常動作が見られた場合は、他のフレームの値と比べて非常に高い値を示す。そのため、該当するフレームが異常動作であると認められる。

図 7 に提案手法におけるフレームごとの d 値の推移のグラフを、図 8 に従来手法におけるフレームごとの d 値の推移のグラフを示す。

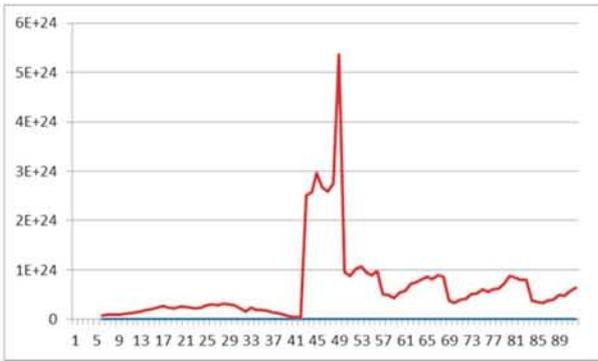


図7：提案手法によるd値推移

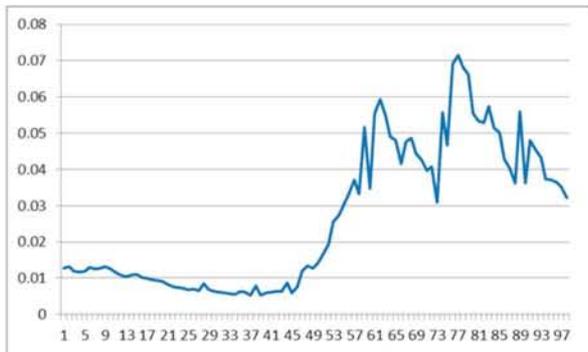


図8：従来手法によるd値推移

従来手法では、異常動作が正しく認識されなかったのに対して、提案手法では異常動作部分において大きな値を示していることがわかる。

4.3 認識率の測定

4.1 で示した例と同様のシーンが含まれる、スケール変化を伴う動画を50パターン、またスケール変化のない動画（人物の左右移動）を50パターン用意した。それぞれの動画において、正しく異常動作を判別できているかを調べた。

動画中の転倒動作のフレームにおいて、射影成分距離d値が特に高い値を示している場合、その動画では異常動作を正しく判別しているとみなす。

それぞれの手法における異常動作の判別率の結果を図9に示す。

	スケール変化あり	スケール変化なし
従来手法	24%	78%
画像ピラミッド	10%	12%
マルチスケールCHLAC	8%	10%
提案手法	60%	80%

図9：各手法における異常動作判別率

スケール変化のある動画においては、提案手法が高い判別率を示した。ただし、人物の前後移動（奥から手

前への移動）の動画については、正しく判別できない場合があった。

また、歩行速度が速い動画や、人物の移動方向の変化が激しい（曲線を描くように歩行する）動画などにおいては、人物の異常動作を正しく判別することができなかった。

単純な前後の歩行動画であれば、ベクトルの大きさは変化しても方向は変化しないため、提案手法で平均化処理を行うことで通常動作と異常動作の判別が可能となった。

一方で、今回判別できなかった例では、通常動作において人物の動きに一定の規則性がないため、そこから得られるベクトル方向の動きが一定でない。そのため、通常動作の判別が難しいと考えられる。

4.3 動作時間の測定

1800フレーム（1分）の動画に対して、開始から終了までに要した時間を、従来手法及び提案手法においてそれぞれ測定した。

その結果、従来手法による動作時間が55.27秒、提案手法による動作時間が62.02秒となった。

このことから、従来手法に付加した、人物スケールに対応するための処理による、特徴量抽出の時間の増大はほとんどない事が分かった。

5. むすび

本研究では、立体高次局所自己相関（CHLAC）を用いた人物動作の異常判別について、特徴量の平均化に基づく新たな人物動作の判別手法を提案した。その結果、これまでの手法では判別が難しかった人物の縦方向・斜め方向への動作判別においても、従来手法より高い判別率を得ることができた。

今後の課題は、今回の実験で判別がうまくいかなかった人物のスケール変化が特に大きい動画において、異常動作の判別を容易にできるよう、特徴量の手法を改善し、判別率の向上を試みる予定である。

付録

別表1に、立体高次自己相関特徴の251種類のマスクパターンを示す。

$n=0$ (1)	n					
$n=1$ (13)	mn an in	kn ln	jn fn	en cn	dn hn	bn gn
$n=2$ (237)	mno mnx knr knA ein jny enw dno dnx amn bnx ahn anr anx lur lnA ckn cln cnq cnz bny lms fnv cnp npA djn gno gux dln ejn hmv nry fjn npu gkn nru hjn nvA nsz muv	kln ben ehn jln dhn juz enx bln dny bnq bnz ain ans any fln knp bjn env cnr cnA nop lsv fny cns hin hnw eln gny dns inw hny nos fms hkn nqs hns ntz nsA muy	jkn abn dgn jno jnu jnA enz akn dnz bnr bnA aln ant anz egn ftn fnw eny cuv ghn hnx dkn guz dnt inx fut imn nqt hln gns ikn nrs ntA nuz nyA	mnr mnA ben jnp bfu dmn enA dur dnA bnt acn ano anu anA cfu fnx chn cnu bnp noy fgn mmu hnz gnr gnA dnu hnt gnt iln ijn nsu muA	efn jmn adn jnu aen bkn dfn dnv bhn bnu afn anv kmn bdn kny fnz cin cnw bns lmm cfn hna ent inA ens inv nps hnu gnu int ins nsx nty	den knq knz jnr jnx ajn din dnw bin bnw agn anq anw lnq luz emn fnA bmn cnx bnv lnp fnp cmn npx ekn gin gnw enu fkn nrv iny npt gjn nrt inu nvx nsy nxy

a	b	c	j	k	l	s	t	u
d	e	f	m	n	o	v	w	x
g	h	i	p	q	r	y	z	A

(3 × 3 × 3)

別表 1：立体高次自己相関特徴の 3 次元マスクパターン
(文字列表現)

参考文献

- [1] 森下 雄介, 大津 展之, 「時間重みと外的規準を用いた動作評価手法」 電子情報通信学会技術研究報告, 2008 年 3 月
- [2] 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, 2005 年
- [3] 松永 錦弥, 鈴木 雄太, 長井 歩, 太田 直哉 「階層化 CHLAC 特徴を用いた異常行動の検出」, 2009 年情報処理学会全国大会, No.5-136, 2009 年 3 月
- [4] 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, J63-D-4, 349-356, 1980
- [5] T. Kobayashi and N. Otsu, “Action and Simultaneous Multiple-Person Identification Using Cubic Higher-Order Local Auto-Correlation”, International Conference on Pattern Recognition, 2004
- [6] N. Otsu and T. Kurita, “A New Scheme for Practical Flexible and Intelligent Vision Systems”, in Proc. of IAPR Workshop on Computer Vision, 431-435p, 1988
- [7] 鈴木 一史, 「3 次元高次局所自己相関特徴マスクを用いたソリッドテキストの分類」, 情報処理学会論文誌 1524-1531p, 2007 年 3 月
- [8] 石井 健幸, 村上 仁己, 小池 淳, 「スケール変化に対応した人物動作の判別手法」, 2012 年電子情報通信学会総合大会, D11-75, 2012 年 3 月
- [9] 石井 健幸, 村上 仁己, 小池 淳, 「スケール変化に対応した人物行動の異常動作特定」, 2012 年画像工学研究会, 2012 年 4 月