

階層化CHLAC特徴を用いた異常行動の検出

松永 錦弥 鈴木 雄太 長井 歩 太田 直哉

群馬大学大学院工学研究科情報工学専攻

Abstract

立体高次局所自己相関CHLACは、動きと形の2つの情報を同時に評価できる特徴量で、異常動作の検出に優れる。しかし既存のCHLACの計算方法は、時空間に対して最適な間隔を見つけることは行わないもの、それぞれに対して単一スケールでの計算であった。異常動作をより良くモデル化するためには、時間的にも空間的にも複数のスケールでの変化を考慮すべきと考えられる。この考えに従い、時間的空間的に複数のスケールの画像を用いる手法として“マルチスケールCHLAC”と“画像ピラミッド”の2つを考案した。実験を行ったところ、既存のCHLACと比較すると大幅に検出精度が向上するという結果を得た。

1 はじめに

安心・安全な社会を目指し、現在様々な場所に監視カメラが設置されている。しかし、人が常時監視するにはコスト面で難しく、見落としも発生してしまうため、犯罪に対する迅速な対処に限界が生じる。この問題を解決するため、現在自動監視の研究がさかんに行われている。

立体高次局所自己相関CHLACは、動きと形の情報を同時に評価できる特徴量であり、動画像のパターン認識においてその効果が期待できる。しかし、既存のCHLACの計算方法は時空間に対して最適な間隔を見つける必要があり、単一スケールでの計算である。行動パターンが単純である動作の場合は単一スケールで十分であるが、複数のパターンが入り混じった複雑な動作の場合、単一スケールでは十分でない。

本研究ではこのCHLAC計算方法に対し、“マルチスケールCHLAC”と“画像ピラミッド”の2つの新手法を考案し、実験結果にて検出精度の向上を見た。

なお、顔の認識をする目的で2次元のCHLAC特徴に対して画像ピラミッドを適用した例は文献1で報告されている。ここではこの考え方を時間を含めた3次元に拡張したと見ることができる。

2 CHLAC (Cubic Higher-order Local Auto Correlation) による異常動作の検出

カメラ映像から異常を検出する手法として、高次局所自己相関(HLAC)特徴を用いた判別方法や、立体高次局所自己相関(CHLAC)特徴を用いた判別方法がある。CHLACはHLACを3次元に拡張した特徴量であり、複数人物の動画像より異常動作を検知できることがいくつかの例で示されている。

2.1 CHLACとは

高次局所自己相関特徴HLACは画像認識のための特徴であり、 3×3 の局所領域を用いる場合、2値画像に対し25次元の特徴ベクトルとなる。CHLACはHLACに時間軸を加えたものであり、 $3 \times 3 \times 3$ の局所領域を用いる場合、2値画像に対し251次元の特徴ベクトルとなる。

HLAC及びCHLACは、2つの優れた特性を持つ。1つは、画像中のどの位置にあっても同じ特徴が得られる、位置不変性である。この特性より、切り出しや位置合わせなどの手間がない。もう1つの特性は、画像中に複数の対象物が存在した場合の特徴は、個々の対象物の特徴ベクトルの和になる、加法性である。この特性より、複数の対象物の中で1人でも異常動作をとっていれば検知できることが期待される。

既存のCHLAC計算方法では、空間間隔 Δr 離れた点と時間間隔 Δt 離れた点を参照している。しかしこの計算方法では、検出対象が変わる度に最適な Δr と Δt を見つける必要があり、手間がかかる上、場合によっては良い精度が見られない(図4参照)。そこで考案したのが本論文の手法、マルチスケールCHLACと画像ピラミッドである。

2.2 マルチスケールCHLAC

既存の計算方法では単一のスケール(1つの Δr , Δt の組み合わせ)で計測するのに対し、複数のスケールを用いて計測する手法がマルチスケールCHLACである。対象動画像を複数のスケールで計測し、得られた特徴を全て同時

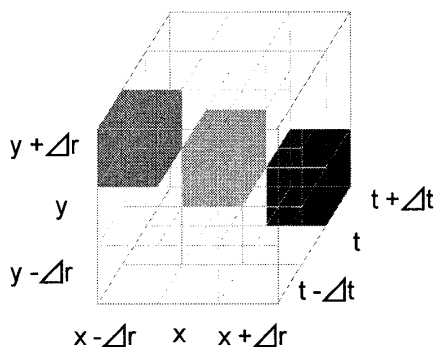


図1：CHLACのマスクパターンの例

に評価することで、より精度の高い行動パターンを抽出する。マルチスケールCHLACでは、 $3 \times 3 \times 3$ の局所領域を用いる場合、2値画像に対し $251 \times$ （用いるスケール数）次元の特徴ベクトルとなる。用いるスケール数を増やすことでより細かい特徴が得られるものの、計測時間は既存のCHLAC計算に要する時間の約（用いるスケール数）倍となる。本研究では、 Δr および Δt が1, 2, 4の3つのスケールでの特徴を用いる。

2.3 画像ピラミッド

画像ピラミッドとは、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である。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のが本手法“画像ピラミッド”である。高解像度の画像からは狭い時間で見た細かい動きの特徴が得られ、低解像度の画像からは幅広い時間で見た動きの特徴が得られるのが大きな利点である。 $3 \times 3 \times 3$ の局所領域を用いる場合、2値画像に対し $251 \times$ （用いる画像群の数）次元の特徴ベクトルとなる。本研究では $1/2$, $1/4$ の3レベルの画像群を用いる。

3 実験と結果

実験には図2, 3に示すような自動販売機前で撮影した映像データを使用する。商品を買っている場面と、人が横切ったりするような場面を通常動作として約3000フレームの動画を学習に用い、自動販売機を破壊しようとする行為を異常と判定することを目的とする。この実験結果を図4, 5, 6に示す。横軸が時間（0.1秒単位）を示し、縦軸が通常動作より作られる部分空間との距離、すなわち異常値を示す。図中の黒線が通常動作に対する異常値を表し、



図2：通常動作



図3：異常動作

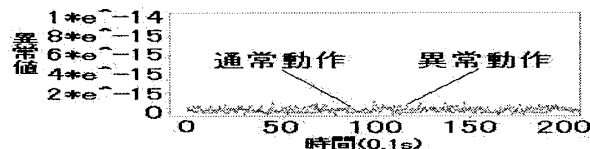


図4：従来CHLAC

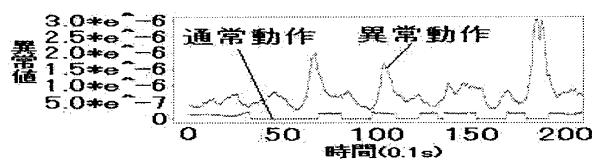


図5：マルチスケールCHLAC

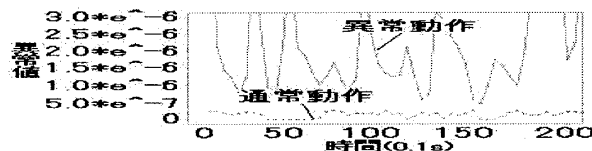


図6：画像ピラミッド

表1：1 CHLACあたりの演算量

マルチスケールCHLAC	0.0346 秒
画像ピラミッド	0.192 秒

青線が異常動作に対する異常値を表す。

従来のCHLAC計算方法と比較すると大きい差がみてとれる。マルチスケールCHLACと画像ピラミッドを用いた計算との結果はわずかに画像ピラミッドが良いが、ほぼ同等である。演算量に関してはマルチスケールCHLACが優れる。

4 まとめ

従来の単一スケールによるCHLACの計算をマルチスケールにすることで、これまで検出できなかった異常動作の検出に成功した。

参考文献

- (1) T. Kurita, N. Otsu, and T. Sato, "A face recognition method using higher order local autocorrelation and multivariate analysis," Proc. of Int. Conf. on Pattern Recognition, Vol. II, pp. 213-216, 1992.