

簡便でノイズ耐性の高いフレーム間差分による動体検知法 —抽出画素の時空間でのつながり情報を用いたランダム変動背景の除去—

絹川 亨† 石村 理知†

†神戸大学 大学教育推進機構 〒657-8501 神戸市灘区鶴甲1-2-1

E-mail: †kinugawa@kobe-u.ac.jp, masachik@kobe-u.ac.jp

あらまし 動物生態の野外での自動監視を目的として、WebカメラとノートPCを用いた簡易可搬システムを試みる中、ユニークでノイズ耐性の高い動体検知法を考案した。基本的にはプリミティブなフレーム間差分法であって、動体領域の候補として抽出された画素が空間的にも時間的にも適切につながっているかどうかを単純なアルゴリズムでリアルタイムに確認することによる。その圧倒的に高いノイズ耐性も放射線物理でのコインシデンス計測との類似性で定量的に説明できる。更には、2値化した差分画像どうしの論理積演算と組み合わせると、ランダム変動する背景の影響も除去できることを野外での実写例で示す。他の洗練された動体検知手法との併用も、簡便で有効な前処理法として可能であろう。

キーワード 動体検知、フレーム間差分、ランダム変動背景

Simple Noise-resistant Method of Motion Detection by Interframe Difference —Elimination of Randomly Fluctuating Backgrounds by Checking Space-Time Connectivity of Extracted Pixels—

Tohru KINUGAWA† and Masachika ISHIMURA‡

† Institute for Promotion of Higher Education, Kobe University, 1-2-1 Tsurukabuto, Nada-ku, Kobe, 657-8501 Japan

E-mail: †kinugawa@kobe-u.ac.jp, masachik@kobe-u.ac.jp

Abstract Developing a portable motion-detective system simply with a Web camera and a laptop PC for the aim of outdoor ethological surveys of animals, we devised a unique noise-resistant method. Basically the method is primitive interframe difference and checks in real-time the space-time connectivity of those pixels extracted from candidate regions of moving objects. The outstanding resistance against noises is explicable in a quantitative manner as the method so resembles to the coincidence technique in radiation physics. By applying the method to the logical AND of two or more binary difference images between neighboring frames, randomly fluctuating backgrounds are mostly eliminated, which is shown with some pictures captured outdoors. The method can work efficiently for selecting motions prior to the application of other sophisticated schemes.

Keyword motion detection, interframe difference, randomly fluctuating background

1. はじめに

近年、監視カメラの普及で動体検知技術が身近になりつつある中、その監視対象は人間・車両や火災などが専らであるが、実は動物の監視にも潜在的なニーズが存在する。たとえば、生態学では動物行動の自然環境下での記録に重きがおかれることがあり、学術専門誌[1,2]でも動画付き投稿が認められている。また、この分野では長時間撮影したビデオを人間の目で解析しているが、動体検知ソフトウェアを使った省力化[3]

も試みられている。応用面でも、「農作物を荒らす外来生物に対して、その天敵を野外撮影で探索する」[4,11]といった、自動監視が要望なテーマも出ている。

そこでわれわれは、WebカメラとノートPCを用いた簡易でポータブルな動体検知システムの試作に踏み切ったのだが、今回はその過程で必要に迫られて考案したユニークな手法について報告する。工夫を要した点は大別して3つであった。第1に、演算能力の制約からノートPCで連写画像の取得と動体検知を同時に

行う限り複雑な処理は無理と考えて、フレーム間差分法に何か単純な演算を付け加えるだけに限定した。第2に Web カメラのノイズである。動物行動の確実な記録という意図があるので、動物が視野を通過した時、漏らさず画像を残すことが絶対必要である。すると、動物検知の閾値を下げざるを得ず、Web カメラのノイズによる誤検知が無視できなくなった。第3が最も本質的で、野外撮影では(草木の揺れのように)ランダム変動する背景がふつうだが、簡単な処理で動物と区別できるのか、という問題である。本報告のポイントは、第1と第2の要件を満たすような動物検知法を工夫し、第3の要件にも有望な答えを得たことである。

われわれの拠り所は「動物に対して時空間方向に連続性を仮定する」[5]、という考え方である。さらに、文献[5]でのオプティカル・フロー解析を単純な演算で置き換えるため、以下のような指針を立てた。

- (1) 動物はある程度の大きさを持っており、動物領域も空間方向に連続なはず。
- (2) 動物はある程度滑らかに運動しており、動物領域の変化も時間方向に滑らかなはず。

(1) は動物領域の画素が空間的に連続かどうかを調べることである。これは単純だが意外にも新しいアイデアであって、本報告の第1の動機となった。われわれが知る限り、2003年に背景領域の雑音除去[6]で最初に提唱され、2年後に動物領域やその移動ベクトルの推定法[7,8]として新たに提起された。次節では(1)によってカメラノイズが徹底的に除去されることを示す。本報告では(2)のように時間方向への拡張も付け加え、(1)と(2)を総合して以下では(時空間での)連続性チェックと称することにする。

実は今回、連続性チェックという手法の新たな利点も見いだした。それは通常のカメラノイズだけではなく、ランダムに複雑変動する背景の影響も除去する可能性で、本報告の第2の動機となった。その端緒は、ランダム変動背景の差分画像を詳しく調べると、その時間変化がノイズに酷似している事実にあった。実際に、連続性チェックの処理を数多くの実写画像(降雨、草木のゆれ、噴水、川の水面のゆれ、滝、海の波など)に施すと、有効に働くことが経験的に判明して来た。

本報告ではまず、連続性チェックとはどういう処理かを説明し、カメラのノイズを徹底的に除去できることを示す。次に、ランダムに複雑変動する背景を除去するために、連続性チェックと差分画像どうしの論理積演算との組み合わせを提唱する。

2. 連続性チェック

連続性チェックの大きな特徴は、Web カメラのように性能の低い機器を用いてもノイズを除去できること

であって[8]、この節ではその理由を説明する。

2.1 最も単純なフレーム間差分法—閾値問題

まず、最も単純なフレーム間差分法を説明して、その問題点を明らかにする。フレーム間差分は時刻 t での入力画像を $I(t, x, y)$ 、一定時間前 $t - \Delta t$ での入力画像を $I(t - \Delta t, x, y)$ とするとき、

$$\Delta I(t, x, y) = I(t, x, y) - I(t - \Delta t, x, y) \quad (1)$$

で与えられる。この差分をもとにわれわれの試作機は連続性チェックの処理を省いた場合、以下のような最低限の手順で動物検知を行う。則ち、動き領域や動きベクトルの決定は全く省略する。

Step1 (2値化差分画像と動物・背景領域の判別)

閾値 $Th1$ を定めて2値化差分画像

$$A(t, x, y) = \begin{cases} 1 & \text{if } |\Delta I(t, x, y)| > Th1 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

を求め、動物領域(1)と背景領域(0)を判別。

Step2 (画素数で動物の存否を判定)

時間 $[t - \Delta t, t]$ で動物が存在するかどうかは動物領域の全画素数 Nt と閾値 $Th2$ を比較して $Nt > Th2$ ならば存在、それ以外は存在しない。

Step3 (画像の保存と消去)

動物が存在する場合 $I(t, x, y)$ と $I(t - \Delta t, x, y)$ を保存、それ以外では $I(t - \Delta t, x, y)$ を記録から消去。

Step3 は記憶容量の浪費を防ぐだけでなく、ヒューマンエラーを少なくするためでもある。保存画像は最終的に人間の目で確認する必要があるので、チェックする枚数が少ないと見落としも少ないと期待できる。

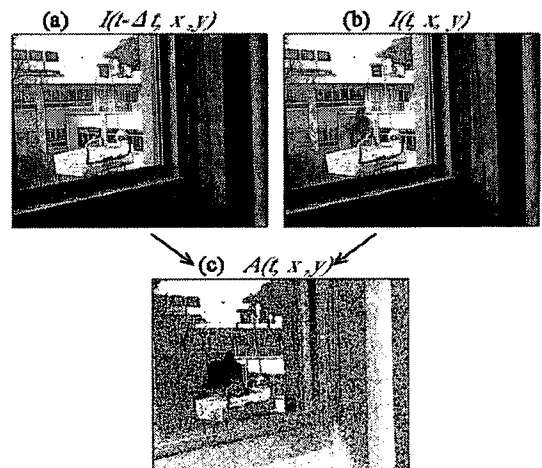


図1 : (a) (b) フレーム画像 と (c) 2値化差分画像。
(c)において■が動物領域で閾値は過小に設定。動物領域の画素数は鳥が~5000画素、それ以外は~63000画素。

そこで、簡易システムとしての必要最低条件は保存画像には必ず動物が写っていることであるが、実際に

はこの条件をクリアするだけでも難しい。動体の写った画像を必ず保存するには、自ずと $Th1$ や $Th2$ といった閾値も十分に下げざるを得ない。実は、Webカメラのようにノイズが比較的多い機器では閾値を下げると、図1の実写例ようになる。図1(a)(b)は鳥が窓辺に留まる前後の瞬間で 640×480 の30万画素カラー画像である。図1(c)で鳥は当然として動体領域と判別されているが、それ以外の領域もノイズのため抽出されている。この誤抽出された画素は鳥の画素よりも数が多く、たとえ動体が写っていないなくても **Step2** で動体が画面に存在すると誤った判断が出るのに十分な数 ($> Th2$) である。すると **Step3** を経ると、ノイズ画像が無駄に長時間保存され続けることになって、動体検知システムとして意味をなさなくなる。一方、図2(a)のように閾値 $Th1$ を上げてノイズによる誤抽出領域を~14000画素まで減らしても問題は解決しない。理由の一つは Webカメラのノイズは経時変化がかなり大きいことで、図1(b)の状態に再び戻ることもよく起きる。それも避けようと閾値 $Th1$ をさらに上げると、図2(b)のように今度は鳥の一部が動体領域として抽出されず、小さい動体を見逃してしまう恐れが出てくる。

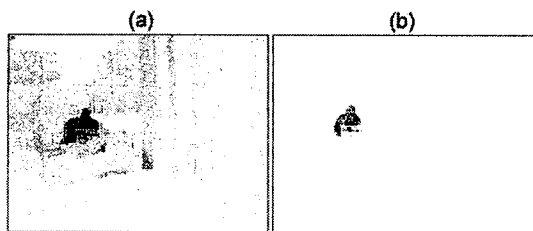


図2：2値化差分画像 (a) 閾値適正 (b) 閾値過大
(a)でのノイズによる誤抽出は~14000画素

このように単純なフレーム間差分法では

『ノイズ除去と動体の確実な検出が両立困難』という問題があり、変動閾値[9]が対策としてよく用いられるが、われわれは以下の全く異なる工夫を行った。

2.2 空間的な連続性チェック

上記の閾値問題に対する抜本的な解決策と考えたのが空間的な連続性チェックで、**Step1** と **Step2** の間に入る処理として2値化差分画像 $A(t,x,y)$ に対して次のように定める。

Step1A (2値化差分画像の水平方向へラン解析)
 動体領域の画素が x 方向へ N_x 個以上連続する箇所はそのまま動体領域として保持、
 そうでない箇所は全て背景領域と再判定。
Step1B (2値化差分画像の垂直方向へラン解析)
 動体領域の画素が y 方向へ N_y 個以上連続する箇所はそのまま動体領域として保持、
 そうでない箇所は全て背景領域と再判定。

図3で $N_x=N_y=5$ の具体例を考えることにする。図3(a)で②行目に水平ラン解析を行うと、5画素以上連続する箇所は全くないので、斜体で書いた動体領域(1)は全て背景領域(0)に転換されてしまう。同様にVII列目に垂直ラン解析を行うとこの列は全て背景領域(0)になる。最終的に図3(a)に **Step1A** と **Step1B** を施すと、中央の 5×5 のウィンドウ部分だけが残る。もう一つの極端な例が図3(b)で、中央にある背景領域(0)の1画素分のため、全体が全て背景領域(0)になってしまう。これらの例をまとめると空間的な連続性チェックとは傾向として、

『 $N_x \times N_y$ 以上で集まった動体領域(1)はそのまま保持、小さいものは背景領域(0)に転換』
 といった All or Nothing 的な処理で、基本的な演算と少ないメモリで高速処理が可能になっている。

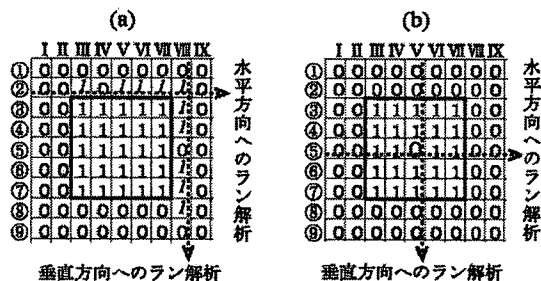


図3：水平および垂直方向へのラン解析

図1(b)と図2(a)の2値化画像に対して $N_x=N_y=5$ の条件で適用した結果を図4に示す。予想通りノイズ領域は将棋倒し的に背景領域へと変化し、一方で鳥の領域は確かに動体領域として残っている。さらにノイズ領域を削減したい場合には **Step1A** を再実行してから

Step1C (画素数の少ない部分を除去)
 動体領域の画素を縦横のつながりごとでまとめて閾値 $Th3$ よりも多い画素数の部分は動体領域、それ以外は背景領域と再判定。

を施すこともある。

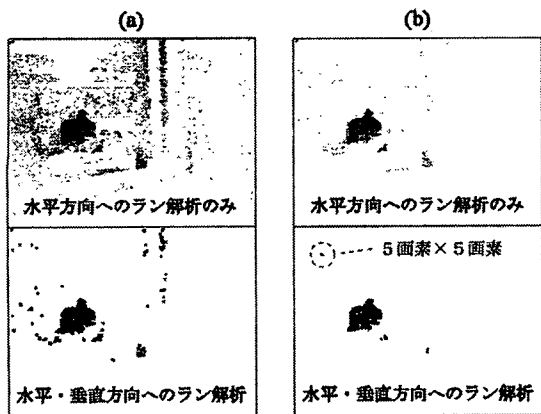


図4：空間的な連続性チェックの適用。
(a) 図1(b)に適用 (b) 図2(a)に適用

2.3 同時計測との類推によるノイズ耐性の定量的評価

図4のような動体とノイズとの峻別は非常に強力であって、実は物理実験での同時計測 (Coincidence Technique) との類推で考えると通常の空間平均 (ローパスフィルター) では実現不可能なレベルまでノイズを除去できることを次に示す。同時計測[9]とは放射線物理の手法で、稀に起こる現象を確実に検出したいときに用いる。1台の検出器からの信号だけでは本当の現象発生かノイズか区別できない状況でも、複数の検出器が現象発生を同時に示すならば本当である確率は高い、という考えである。「検出器」を「画素」、「稀に起こる現象」を「動体の侵入」と読み換えれば、

『空間的な連続性チェックとは
 $N_x \times N_y$ 個の画素を束ねる同時計測』

といえる。その定量的な説明に先立って、背景領域の1画素がノイズによって動体として誤検出される確率を p 、 $N_x \times N_y$ のウィンドウが1画面中に C 個取れるものとする。各々の画素でノイズが独立に発生し、画面中にはもともと動体は侵入していない場合、空間的な連続性チェック後に動体領域が残っているには、 $N_x \times N_y$ のウィンドウが少なくとも1個以上、ノイズによって動体領域(1)で全て満たされなければならない。したがって、その起こる確率は

$$\Pi \doteq C p^{N_x N_y} \quad (3)$$

となる。図1(b)の例では $N_x = N_y = 5$ 、 $C \doteq$ 全画素数 $\div 3 \times 10^5$ 、 p はおよそ6万画素 $\div 30$ 万画素 $= 0.2$ 程度であるから、最終的に $\Pi \sim 1 \times 10^{-12}$ と極端に小さい値になって、特異的に高いノイズ耐性を説明できる。近傍画素どうしのノイズは完全に独立ではないので実際の値は落ちるが、動体検知の閾値をかなり下げてもカメラ由来のノイズが確実に除去できるのである。一方で、

動体を検知から見逃すこともほとんどない。というのは、閾値を十分下げられるからである。逆に、5画素 \times 5画素よりもサイズが小さく、背景と違わない動体を見逃しても、情報損失としてあまり重要ではない。

2.4 時間的な連続性チェック

前節では同時計測の考えを「検出器=画素」との対応で展開したが、「検出器=フレーム」とすれば時間方向への連続性チェックが自ずと得られる。最も単純に『複数の連続フレームに動体が存在する場合に限り、動体の検知を確定する』

とすれば、同時計測の考え方と完全に符合する。則ち、1枚の画像にしか写っていないものは動体と見なさず、画像を削除するのである。実はこれだけでも動体検知の確度は上がり、急な日照変化や光の乱反射など、急激な変化をかなり排除できるようになった。

実用的には連続性チェックを時間と空間で統合すると、晴天では動体以外にはあまり反応しなくなった。具体的として図1の例を取り上げる。これはカメラを窓から屋外に向けて日中の12時間に13万枚の連続画像 (~3枚/秒 計~40GB) を撮影し続け、最終的に保存された528枚 (計132MB) のうちの1枚である。この528枚を調べると、動体として鳥が6回に渡って写っており、全く動体が写っていないのは41枚であった。その誤検知の内容を調べると、激しい日照変化と風による樹木や影の揺れであった。そのようなランダム変動する背景への対策として、時間的な連続性チェックを次節でさらに拡張する。

3 ランダム変動する背景の除去

3.1 連続性チェックと論理積演算との組み合わせ

連続性チェックと2値化差分画像どうしの論理積演算を組み合わせることで、実はランダム変動する背景を効率良く除去できる。図5にその例を示す。

図5の最上段はフレーム画像で、時間の経過は $F1 \rightarrow F2 \rightarrow F3$ の順番である。ホースからの水を写真右外から中央へ向かって広い範囲で散らしつつ、 $F2$ の瞬間に中央へボールを動体として投げ入れた。第2段は2値化差分画像で、第3段は第2段の画像に5画素 \times 5画素の空間的な連続性チェックを施したものである (以下、図中の記号*は空間的な連続性チェックを施したことを示す)。動体領域として抽出されたのはボールだけでなく、上部や右部の樹木、中央部の散水それに下部の水しぶきが見られる。最下段では第3段の2値化差分画像どうしで論理積を取り、その後にもう一度、連続性チェックが施されている。論理積を取ると動体領域は明らかにボール周辺に集中し、ランダム変動する樹木や水しぶきは除去されている。

このようにランダム変動する背景だけが選択的に除去できるのは偶然ではなく、その理由として

- (1) 空間的な連続性に乏しい。
- (2) 時間的な連続性に乏しい。

といった2つの性質に起因することが分かってきた。以下の副節ではこれらの性質を個別に説明する。

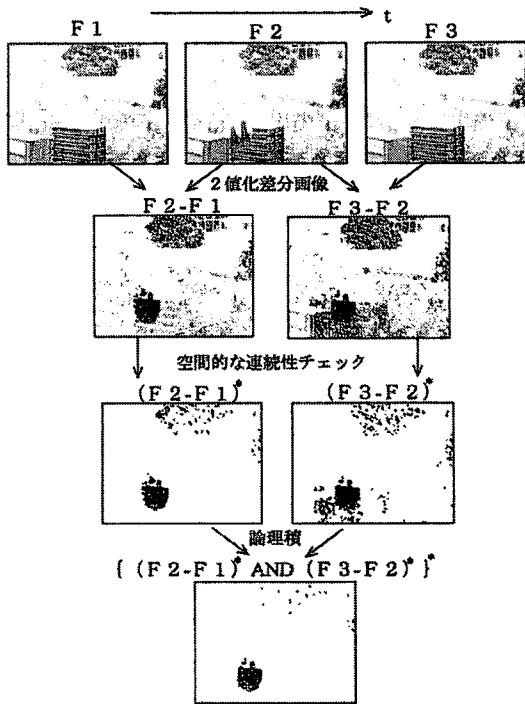


図5：連続性チェックと論理積演算との組み合わせ

3.2 空間的な連続性の欠如

空間的な連続性が欠如しているとは文字通りで、ランダム変動する背景の箇所を見ると、その中の動体領域が画素単位で縦横へあまりつながっていないことを意味する。その原因は2つあって、図5では水しぶきと樹木がそれぞれをよく表している。第一に、水しぶきはもともと形状が細かい断片であって縦横のつながりが少ない。したがって、図5では空間的な連続性チェックを施した第3段ですでにかなり除去されている。一方で、樹木も葉や枝といった細かい構造を持っているが、実は水しぶきとは違ったもう一つの理由が働いている。それは樹木全体が同じような色合いであることに起因し、ここでは説明のため樹木が風に揺れる状況を考える。この場合、フレーム間差分を求めると、『違う物体が写っていても、差分がたまたま小さい』

ということが実際に起こるのである。具体的には、樹木が揺れ動く前後で異なる葉が同一位置に来る場合に相当し、差分を取ると2つの葉の色合いがあまり違わないため、その位置の画素はたまたま背景領域(0)と判別されてしまう。これは図3(b)と似た状況で、偶発的に背景領域(0)とされてしまう画素が一つ存在して、その周囲を動体領域(1)で囲まれているのである。すると、周囲の動体領域(1)は空間的な連続性チェックによって次々と背景領域(0)に転換されていく。それで、図5で樹木の位置を見ると、第3段から白い背景領域(0)が急に増えるのである。以上のような理由で、水しぶきと樹木のいずれも図5の第3段までの処理で、除去されるか細かい破片のような動体領域となる。一方で対照的に、ボールは大ききまとまった動体領域(1)として残っている。

3.3 時間的な連続性の欠如

時間的な連続性が乏しいとは、ランダム変動する背景の場合、動体領域(1)の画素が時間的に不規則に背景領域(0)へと変化する、ということである。これは空間的な連続性が乏しい理由の2番目を考えた時点で、すでに予想されるものである。樹木の例を再び取り上げると、葉の重なりで偶発的に生じる背景領域(0)は、その位置だけでなく時間変化もランダムであろうということである。この予想を図5に沿ってさらに具体的に説明する。第3段の2つの図では樹木の位置に動体領域が散らばって見られるが、この2つの動体領域は空間的に重なりが少ないのではないかと予想されるのである。というのは、葉の動きの時間的なランダム性から、偶発的な背景領域(0)の位置も時間的に不規則に変化するはずだからである。この予想を確かめたのが図5の最下段である。論理積を取ることで重なる部分を重点的に残すようになっている。樹木の位置を見ると予想通り、第3段では動体領域であった箇所が最下段ではかなり背景領域(0)へと転換されている。一方でボールを見ると、背景が少々変化しても動体領域(1)と背景領域(0)が入れ替わるようなことが少ないので、最後の処理まで大きなまとまった動体領域(1)として残っている。

3.4 野外での実写への適用例

通常、野外での背景はしばしばランダムに変動する。実は、連続性チェックと論理積演算を組み合わせると、たいいてい背景が効率良く除去できる。背景が変動しても前の2節で説明したように、空間的・時間的連続性が乏しいからである。図6で滝の実写例を示す。

図6のように実写画像で動体領域を順次調べると、空間的にも時間的にもつながりが少ないことが分かる。

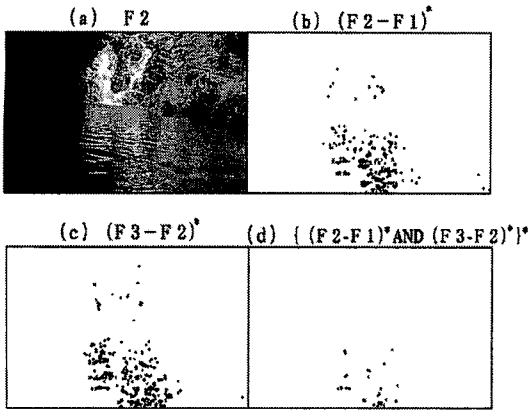


図6 野外での実写画像（滝）への適用例
(a)(b)(c)(d)は図5での同じ記号のものに対応

滝の実写画像における動体候補として、滝つぼに流れ込む水流および水面のゆれの2つがあげられる（図6(a)）。まず、水流の領域はほとんど同じ色合いをしている。そのため、差分をとると背景領域（0）に判定される画素が多い。したがって、差分画像に空間的な連続性チェックを施すと次々と背景領域（0）に転換されてしまい、水流に相当する位置では動体領域（1）はほとんど検出されていない（図6(b)、(c)中央上部）。一方、水面のゆれに相当する位置では、動体領域（1）が散らばって抽出されている（図6(b)、(c)中央下部）。この水面のゆれは、図5の樹木と同様に、時間的にランダム変動していると予想される。つまり、水面のゆれの位置に散らばって存在する動体領域は図6(b)、(c)で空間的な重複に乏しいと推測される。実際、図6(b)、(c)の論理積をとり連続性チェックを施した結果、画面のほとんどが白い背景領域（0）に転換されている（図6(d)）。水流の位置は完全に背景領域（0）となり、水面のゆれの位置がわずかに抽出されているに過ぎない。このように、滝の実写画像におけるランダム変動はほとんど背景領域（0）として除去される。

実は、連続性チェックと論理積演算の組み合わせは背景除去のフィルターとして普遍性がかなり高い。野外における背景の変動は空間的・時間的な連続性のどちらか、または両方に乏しいことが多いからである。実際、さまざまな実写画像（降雨、草のゆれ、噴水、川の水面のゆれ、海の波）において、ランダム変動する動体が存在しても本手法を施すと、ほとんど背景領域（0）と判定されることが分かってきた。

4 まとめ

本報告では時空間での連続性チェックという非常に単純な手法を提示し、ランダム変動する背景の除去に有効であることを実写画像も交えて説明した。

動体検知に関しては高度な手法がすでに様々な開発されている状況で「屋上に屋を架す」の感もあるが、本手法はその簡便さから撮像機器や録画機器といったハードウェアに直接組み込めるのではないかと期待している。そして、「何か動くものが写っている画像」だけがひとたびピックアップされれば、他の洗練された手法を抽出された画像に適用すれば、動き領域や動きベクトルの確度も改善され、ひいては動体の識別も可能性が出てくるものと思われる。則ち、連続性チェックという手法は演算負荷が非常に少ない割に有効性の高い前処理法として、他の動体検知法との併用も可能であることを最後に記してまとめとする。

本研究の一部は助成[11]を受けたもので、ここに謝意を表します。

文 献

- [1] Journal of Ethology, Video Article, <http://www.springerlink.com/content/105357/>
- [2] The Biological Bulletin, Video and Data Supplements, <http://www.biolbull.org/supplemental/>
- [3] 植田睦之, 田中啓太, “鳥の巣のビデオ録画の動体監視ソフトウェアによる自動解析”, Bird Research, vol.21, T1-T2, 2006.
- [4] Y.Yusa, N.Sugiura and K.Ichinose, Predation on the apple snail by the Norway rat in the field, Veliger, vol.43, pp.349-353, 2000.
- [5] 長井敦, 久野義徳, 白井良明, “複雑変動背景下における移動物体の検出”, 電子情報通信学会論文誌 (D-II), vol.J80-D-II, no.5, pp.1086-1095, May 1979.
- [6] 特開 2005-10840, 日本電信電話株式会社, 「背景領域検出装置、方法および該方法を実行するためのコンピュータプログラム」
- [7] 特開 2006-309510, 日本電信電話株式会社, 「動き検出装置およびそのプログラム」.
- [8] 特開 2006-311207, 日本電信電話株式会社, 「モニタリング映像配信システム、モニタリング映像配信サーバー、及びモニタリング映像生成方法」.
- [9] 藤吉弘亘, 小村剛史, 矢入郁子, 香山健太郎, 吉永宏, “歩行者 ITS のためのフレーム間差分による移動体検出法とその評価”, 情報処理学会論文誌, vol.45, No.SIG13(CVIM 10) pp.11-20, Dec. 2004.
- [10] W. ボーテ, “同時放電法”, ノーベル賞講演物理学 (7), ノーベル財団 (編) 中村誠太郎 小沼通二訳, pp.153-159, 講談社, 1979.
- [11] 遊佐陽一, 「生態特性を利用したスクミリンゴガイ制御技術開発—在来天敵による水域個体群の制御—」, 農林水産会議, 作物対応研究課題「担い手の育成に資する IT 等を活用した新しい生産システムの開発」, 平成19年度.