

## Second-order Neural Network を用いた移動物体領域の抽出

茶志川 孝和<sup>†,††</sup> 武藤 佳恭<sup>†††</sup>

動画画像から移動物体を抽出する方法として、従来から差分法が使われている。しかし、差分法による抽出では、移動物体の全体が得られないという問題がある。そのため、モルフォロジ演算によるノイズ除去と欠損個所の補完を後処理として行う方法が知られている。ところが、撮影環境によっては後処理が有効に機能しない場合も多く、また演算コストの高さも指摘されている。本論文では、Second-order Neural Network をフレーム間差分法の後処理として用いる移動物体領域の抽出手法を提案する。提案手法とモルフォロジフィルタを用いた従来手法を用いて、人工動画画像による実験と実動画画像によるシミュレーションを行った。それらの結果、提案手法はモルフォロジフィルタを用いた従来手法に比べノイズに影響されにくく、また抽出領域の欠損が少ないことが示された。また、オプティカルフローを用いた従来手法との比較実験においても、提案手法の優位性を確認した。

### Extracting Moving Object Areas Based on Second-order Neural Network

TAKAKAZU CHASHIKAWA<sup>†,††</sup> and YOSHIYASU TAKEFUJI<sup>†††</sup>

The frame difference technique has been widely used for extracting moving objects. However the technique can not extract the whole shape of the moving object. A morphology filtering is used for reducing noises and complementing lacks of extracted objects as an post-processing to deal with the problem. The morphology filtering is not always available under various conditions and its calculation cost is high. This paper proposes a new method for extracting moving object areas, based on the frame difference technique and the Second-order Neural Networks. We experimented our proposed method and a morphology-based conventional method using artificial movies and real movies. The experimental results show that the proposed method has a better performance for a noise reduction and for an extraction of whole moving objects than the morphology-based conventional method. Moreover, the proposed method has better performance than the Opticalflow-based method.

#### 1. はじめに

コンピュータビジョン研究において、動画画像からの移動物体領域の抽出は基本的かつ非常に困難な問題の1つである。

移動物体領域を抽出する方法は、これまでいろいろ提案されてきたが、オプティカルフローを用いる方法と差分画像を用いる方法に大別することができる。オプティカルフローを用いる方法では、まずオプティカルフローと呼ばれる動画画像中の運動物体の見かけの速度場を推定し、次にそのオプティカルフローを分析して移動物体の領域を決定する。オプティカルフローの

推定方法としては、少ない動画画像枚数から瞬間的なオプティカルフローを推定するのに有効で、測定時間も他の手法に比べて高速に行えるグラディエント法の局所最適化法が、移動物体追跡システムなどでよく用いられている<sup>1)</sup>。ただし、物体の境界など画像中のエッジにおける推定精度が他の領域と比べて大幅に低下してしまうという問題がある。また、雑音や環境変化などにより、1つの動物体に属する画素が不均一なオプティカルフローの推定値を持つという問題を解決するために、時空間輪郭積分法<sup>2)</sup>が提案されている。しかし、画素単位でのオプティカルフローが得られない、計算量が多いという問題がある。オプティカルフローから移動物体領域を決定する方法としては、オプティカルフロー場からハフ変換を用いて、一様な三次元の動きパラメータを持つ領域ごとに分割する方法<sup>3)</sup>やオプティカルフローに信頼性を与える方法<sup>4)</sup>がある。しかし、これらのオプティカルフローに基づいて移動物体領域を抽出する方法の問題は、物体の境界の近傍で

† ニッタン株式会社

Nittan Co., LTD.

†† 慶應義塾大学 SFC 研究所

Keio Research Institute at SFC

††† 慶應義塾大学環境情報学部

Faculty of Environmental Information, Keio University

は、動きが一樣でないために抽出されるフローの誤差が大きくなることである。また、オプティカルフローだけではなく色や明るさのエッジを用いる方法も提案されているが、柔物体や見かけ上形の変化する物体の領域はうまく抽出することができない<sup>5)</sup>。一方、差分画像を用いた方法<sup>6),7)</sup>は移動物体領域を抽出する最も直接的な手法であり、計算が容易でハードウェア化がしやすいなどの利点がある。しかし、1. カメラが移動すると移動物体の抽出が困難となる、2. 照明環境の変動がノイズとして出力される、3. 移動物体の全体が抽出されないことがある、といった問題もある<sup>8)</sup>。これらの問題のうち2と3を解決するために、後処理としてフィルタ処理によるノイズの除去や欠損個所の補完が一般的に行われている。特にモルフォロジ演算によるフィルタ処理を適用する方法<sup>9),10)</sup>が広く知られている。しかし、撮影環境によってはうまく機能しない、計算コストが高いなどの問題が指摘されている。

一方、ニューラルネットワークを応用して移動物体領域を抽出する研究報告もあり、SecondOrder Neural Network(以下,SONN)を用いた例として、Feed Forward Shunting(以下,FFS)が提案されている<sup>11)</sup>。FFSは、空間的に近傍との結合と入力に関係付けることで、ノイズの影響を受けることなく時系列に動く矩形波の進行方向の先端エッジを抽出することを可能としている。しかし、扱っている空間が一次元であることや、検出できる方向が決められているなど、一般の動画処理に応用するには問題がある。また、画像処理への応用研究が多いニューラルネットワークとしてPulse Coupled Neural Network(以下,PCNN)が知られている。PCNNは、Eckhornらが猫の視覚野におけるニューロン間の同期現象を証明するために開発したニューラルモデル<sup>12)</sup>を基に、Johnsonらが計算機上でのシミュレーションが可能となるように拡張を施したものである<sup>13)</sup>。PCNNが適用されている画像処理は、静止画を対象とした例がほとんどであるが、画像中からエッジやセグメントなどの特徴を抽出する機能をさまざまな形で応用した研究が多数報告されている<sup>14)~16)</sup>。これらの先行研究にみられるように、PCNNは二次元空間上の画像を容易に扱うことが可能であり、またFFSも容易に導入できる。

そこで、本論文ではPCNNに基づいた新たなSONNモデルを提案し、差分画像を用いた移動物体領域抽出手法の後処理に適用することで、差分画像を用いた方法がかかえる上記問題2,3に対応できることを示す。提案手法は、移動物体領域は差分画像を時系列にみると空間的にあるまとまりをもって連続的に変化すると

仮定する。そのため、差分画像に抽出されたランダムに現れるノイズは、時系列でみると離散的に変化しているため、移動物体と区別することができる。また、提案手法は差分画像において移動物体領域に属する画素は、移動物体が通過する間ある一定の差分値を連続して持つと仮定する。そのため、差分画像に欠損として現れる領域を画素の瞬間的な変化としてとらえることで、補完することができる。移動物体領域の正確な抽出は、動画を扱う多くのシステムの最初の処理にあたるため、この処理での結果がシステム全体の性能に大きな影響を与える。そのため、本研究の成果はすでに提案されている動画処理システムの性能向上にも貢献できるであろう。

以下、2章において、関連研究であるFFSとPCNNの基本モデルについての説明を示す。3章では、提案する移動物体領域の抽出手法に関する説明を行う。4章では、人工動画をを用いた実験を行い、さまざまな条件下における提案手法の挙動を明らかにする。5章では、移動物体領域を抽出する手法として代表的なオプティカルフローを用いた手法と比較実験を行い、提案手法の優位性を示す。6章では、提案手法を実動画に適用した結果を示し、実画像においても提案手法が有効に機能することを示す。7章では、提案手法の諸問題を取り上げ、8章で結論を示す。

## 2. 関連研究

この章では、提案モデルの基礎となっているFFSとPCNNの基本モデルの説明を行う。

### 2.1 Feed Forward Shunting (FFS)

近傍ニューロン間の高次元相関は、High Order Neural Networkとしてモデル化できる<sup>17)</sup>。特に近傍ニューロンと二次元の結合を持つニューラルネットワークはSONNと呼ばれ、ニューロンの活動電位は以下のように表すことができる。

$$X_i = \sum_{j(1)} W_{ij(1)} X_{j(1)} + \sum_{j(1)} \sum_{j(2)} W_{ij(1)j(2)} X_{j(1)} X_{j(2)} \quad (1)$$

for  $1 < j(d) < N$  and  $1 < d < 2$

ここで、 $X_i$ はニューロン*i*の活動電位で、 $N$ は近傍ニューロンの数、 $W_{ij}$ はニューロン*i*と*j*の結合重みである。そして、 $j(d)$ は*d*次元で結合されている最近傍ニューロンへの指標関数である。式(1)の右辺第2項は近傍ニューロンの活動の相関、すなわち $X_{j(1)}$ と $X_{j(2)}$ の積を計算している。そのため、近傍ニューロンの活動の同期性を測定することができる。

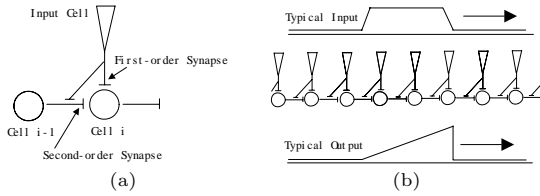


図1 FFSの構造と一次元ネットワーク

Fig. 1 The structure and the one-dimensional network of FFS.

FFSはSONNの一種であるが、その構造を図1(a)に、動作式を下式に示す。

$$X_i(t+1) = AI_i(t+1) + BI_i(t+1)X_{i-1}(t) \quad (2)$$

ここで、 $I_i(t+1)$ は時刻 $t+1$ における入力であり、 $X_{i-1}$ は左隣にあるニューロンの出力である。また、 $A, B$ は固定された結合重みである。図1(b)のように一次元にネットワークを組んだ場合の動作は以下のとおりである。もし、入力 $I_i(t+1)$ が十分大きければ、 $X_{i-1}(t)$ は増幅されて $X_i$ へ伝播する。しかし、大きくない場合は減衰して伝播する。増幅が生じるのは、 $BI_i(t+1)$ が1以上になる場合だけである。つまり、 $X_i(t)$ と相関のある入力 $I_{i+1}(t+1)$ がある場合のみ活動電位の伝播が生じる。このため、図1(b)の入力のようにかたまりとなって移動する外部刺激はニューロン間を伝播するが、ランダムに変動する外部刺激が伝播されることはない。

## 2.2 Pulse Coupled Neural Network (PCNN)

本節では、二次元に配列されたニューロン( $i, j$ )に着目して説明を行う。PCNNのニューロンモデルを図2に示す。ニューロンはDendritic tree, Linking modulator, Pulse generatorと呼ばれる3つの部分から構成される。Dendritic treeは、Feeding部、Linking部と呼ばれる2つの部分に分かれており、それぞれLinking modulatorに対して独立した出力を生成する。Feeding部とLinking部は、シナプス結合 $M, W$ を介して近傍のニューロンと接続されている。なお、外部刺激 $S$ はFeeding部だけが受ける。また、Feeding部とLinking部はそれぞれ前の時刻の値を保持しており、その値は時間とともに減衰する。Feeding部の値 $F[t]$ とLinking部の値 $L[t]$ は次式よりそれぞれ求まる。

$$\begin{aligned} F_{ij}[t] &= \exp(-\alpha_F \delta_t) F_{ij}[t-1] + S_{ij} \\ &\quad + V_F \sum_{kl} M_{ijkl} Y_{kl}[t-1] \\ L_{ij}[t] &= \exp(-\alpha_L \delta_t) L_{ij}[t-1] \end{aligned} \quad (3)$$

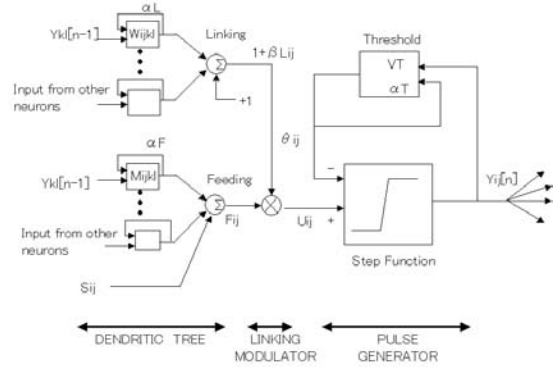


図2 PCNNのニューロンモデル

Fig. 2 The neuron model of PCNN.

$$+ V_L \sum_{kl} W_{ijkl} Y_{kl}[t-1] \quad (4)$$

ここで、 $S_{ij}$ は外部刺激、 $Y_{kl}[t-1]$ は $(k, l)$ に位置する近傍ニューロンの前の時刻での出力を示す。 $\alpha_F, \alpha_L$ は時定数を、 $\delta_t$ は単位時間を表す。 $V_F$ および $V_L$ はシナプスの重み行列を正規化するための定数である。

Linking modulatorは、Dendritic treeの2つの部分の出力を結合強度 $\beta$ に従って結合して内部状態を決定する。ニューロンの内部状態は次式によって計算される。

$$U_{ij}[t] = F_{ij}[t](1 + \beta L_{ij}[t]) \quad (5)$$

Pulse generatorは、内部状態と動的閾値を比較して出力を決定する。出力 $Y$ は以下の式より導き出される。

$$Y_{ij}[t] = \text{step}[U_{ij}[t] - \theta_{ij}[t]] \quad (6)$$

ここで、 $\text{step}(u)$ はステップ関数であり以下のように定義される。

$$\text{step}(u) \begin{cases} 1 & (u \geq 0) \\ 0 & (u < 0) \end{cases} \quad (7)$$

また、 $\theta$ は動的閾値であり以下のように定義される。

$$\theta_{ij}[t] = \exp(-\alpha_\theta \delta_t) \theta_{ij}[t-1] + V_\theta Y_{ij}[t] \quad (8)$$

ここで、 $V_\theta$ は正規化定数、 $\alpha_\theta$ は時定数を、 $\delta_t$ は単位時間を表す。なお、閾値 $\theta$ は次のように更新される。まず、ニューロンがパルスを発生すると閾値にフィードバックされ、パルスの生成が停止するまで閾値は急激に増加する。その後、閾値は再び内部電位の値より小さくなるまで時定数に従って減少する。

ネットワークの1回の動作は、ニューロンを構成する3つの部分が下記のように動作することで完了する。

- (1) それぞれの受容野の入力加重和を計算する (Dendritic tree)。

- (2) Linking 部の出力により Feeding 部の出力を調整 (Liking modulator) .
- (3) 出力パルスを生成 (Pulse generator) .
- (4) 発生したパルスが、受容野を介して他のニューロンに入力される .

また、PCNN を画像処理に適用するには、画像中の 1 画素に対して PCNN の 1 ニューロンを割り当ててネットワークを動作させればよい。具体的には、入力画素と同じ個数のニューロンを画素と同様に二次元に配置し、正規化した画素の輝度値を対応するニューロンの外部刺激  $S_{ij}$  として入力する。なお、出力としては、正規化したそれぞれのニューロンの出力を画素の輝度値と見なせば、ネットワークの動作 1 回につき 1 枚の画像を得ることとなる。

### 3. 移動物体抽出手法の提案

提案手法は、フレーム間差分法による差分画像の抽出と PCNN を変更した SONN による後処理からなる。本章では、PCNN の変更方法について説明を行った後、提案モデルを用いた移動物体抽出手法を示す。

#### 3.1 提案モデル

フレーム間差分法を用いて動画像から移動物体の全体を抽出するには、フレーム間差分法の出力に含まれるノイズを除去し、抽出領域の欠損補完を行う後処理が必要となる。このような後処理を行ううえで、提案手法は差分画像に表れる移動物体領域の変化に関して以下の 2 点を仮定した。

- 移動物体領域は差分画像を時系列にみると空間的にあるまとまりをもって連続的に変化する。
- 差分画像において移動物体領域に属する画素はある一定の差分値を移動物体が通過する間連続して持つ。

提案モデルでは、この仮定に基づいて PCNN の Feeding 部、Linking 部、そして閾値に対して以下に述べるような変更を行った。

##### 3.1.1 Feeding 部に行った変更

PCNN 基本モデルの Feeding 部は外部刺激と近傍ニューロンの出力を受けているが、提案モデルでは外部刺激と近傍ニューロンの出力の相関を受けるように変更を行った。これにより、FFS にみられるように移動する外部刺激とノイズを区別することが期待される。また、PCNN の基本モデルの外部刺激はニューロンが対応している画素の値となるが、提案モデルでは対応する画素の近傍画素の値も考慮し、さらに外部刺激に対して過渡応答性を持つように変更を行った。これにより、かたまりとして与えられる外部刺激の欠損を

補うことが期待される。式 (3) を式 (9)、(10) と置き換える。

$$F_{ij}[t] = P_{ij}[t] + \exp(-\alpha_F \delta_t) F_{ij}[t-1] + V_F \sum_{kl} M_{ijkl} Y_{kl}[t-1] S_{ij}[t] \quad (9)$$

$$P_{ij}[t] = V_p \sum_{kl} R_{ijkl} S_{ij}[t] + \exp(-\alpha_p \delta_t) P_{ij}[t-1] \quad (10)$$

ここで、 $R$  はシナプス結合重み、 $\exp(-\alpha_p)$  は減衰項、 $V_p$  は正規化定数を示す。

##### 3.1.2 Linking 部に行った変更

基本 PCNN において、Linking 部の値は Feeding 部の値を変調する役割を担っているが、近傍ニューロンが発火していない場合に負の値を出力するように変更を行った。これにより効果的にノイズが除去されることが期待できる。具体的には近傍ニューロンからの出力から 1 を引くことで実現させている。基本 PCNN の式 (4) を次式と置き換える。

$$L_{ij}[t] = \exp(-\alpha_L \delta_t) L_{ij}[t-1] + V_L \sum_{kl} W_{ijkl} (Y_{kl}[t-1] - 1) \quad (11)$$

##### 3.1.3 閾値に行った変更

PCNN 基本モデルの動的閾値はニューロンに不応期を与えているが、本研究の目的においてはつねに移動物体を抽出する必要があるため、閾値は定数とした。なお、この変更によりニューロンモデルとしての特徴が大きく変化するが、詳細に関しては次項で述べることとする。式 (8) を次式と置き換える。

$$\theta_{ij}[t] = \text{const.} \quad (12)$$

#### 3.2 提案モデルのニューラルネットワークとしての性質

提案モデルの基となっている PCNN には 3 つの基本的なメカニズムがある<sup>18)</sup>。

第 1 のメカニズムは動的閾値である。ニューロンの閾値はニューロンが発火すると急激に大きな値をとり、その後時間とともに減衰する。閾値があるニューロンの内部エネルギーよりも小さくなると、そのニューロンは再び発火し、閾値は増大する。このような振舞いによって、それぞれのニューロンは連続パルスを生成し続ける。

第 2 のメカニズムは、セグメンテーション機能である。これは、ニューロン間の局所的な相互結合によって引き起こされる。発火したニューロンは近傍のニューロンの発火を促す。そのため、あるニューロン群が発火しそうなときには、1 つのニューロンがトリガとなってニューロン群全体を発火させることができる。この

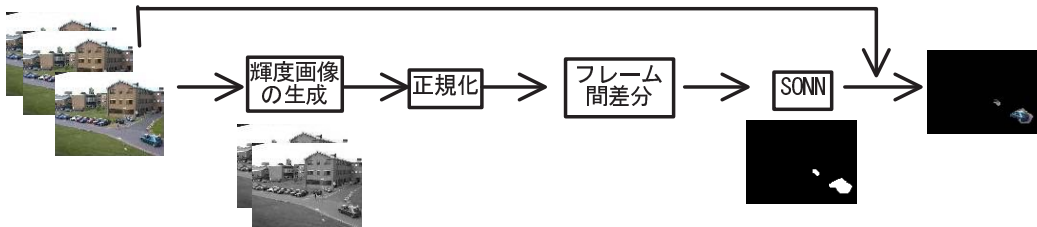


図3 提案モデルを用いた移動物体領域抽出方法

Fig. 3 Proposed method for extracting moving object areas.

結果、画像中の輝度値が似通ったセグメントは同期して発火する。

第3のメカニズムは、数回のイタレーション後に引き起こされる、発火の同期が時間とともに崩壊する現象である。この崩壊はセグメントの小さな輝度の違い(テクスチャ)が最終的には、ニューロンの内部エネルギーを変化させることによって引き起こされる。

提案モデルはPCNNの動的閾値を固定値に変更したことによって、第1のメカニズムに起因する性質を失ったことに注意しなければならない。たとえば、PCNNの有するセグメンテーション機能は、ニューロン間の局所的な相互結合によって近傍ニューロンの発火が互いに促されることに基づいている<sup>19)</sup>。もし、ニューロンが減衰してゆく閾値付近で同じような内部エネルギーを持っていたとすると、同じような外部刺激を持つニューロンは同期して発火し、その結果セグメントが形成される。これは、上記、第1と第2のメカニズムによってもたらされる性質といえる。一方、提案手法が有する移動物体の全体を抽出する機能は、外部刺激と近傍ニューロンの相関によって生じる発火の伝播作用とニューロンの過渡応答性によって一度発火したニューロンの発火が促されることに基づいている。外部刺激がかたまりとして移動しながら与えられる場合、そのかたまりの欠損個所にあたるニューロンの外部刺激は小さいが、移動方向に生じるドミノ効果とニューロンの過渡応答性によって内部エネルギーが高くなるので発火する。その結果、欠損個所が補完され移動物体全体が抽出される。これは、PCNNのメカニズムに関連していえば、上記第2のメカニズムに対して外部刺激との相関関係を設けたことによってもたらされる性質といえる。

このように、提案モデルはPCNNをベースとしているが性質は異なったものとなっている。そのため、提案モデルは広義の意味においてSONNモデルの一種と位置付けられる。

### 3.3 提案モデルを用いた移動物体領域の抽出法 フレーム間差分法の後処理として提案モデルを用い

た移動物体抽出手順を示す(図3)。

- (1) カラー動画から連続フレーム画像を生成。
- (2) フレーム画像の輝度値を256階調で計算し、輝度画像を生成。
- (3) 生成した輝度画像を0から1に正規化。
- (4) 正規化した輝度画像をフレーム間差分法で処理し、輝度変化の画像を生成。
- (5) 輝度変化の画像をSONNの外部刺激として入力しネットワークを動作。
- (6) SONNの各ニューロンの出力(1or0)と対応するカラーフレーム画像の画素の積を計算。
- (7) (2)~(6)の処理をフレームごとに繰り返す。

提案手法では、輝度変化の画像生成にフレーム間差分法を適用しているが、同じ差分法である背景差分法を適用する方法も考えられる。しかし、背景差分法は、背景モデルを事前に用意し、刻々変化する撮影条件に応じた更新を行う必要がある。そのため、環境の変化が激しい場所で撮影された画像を処理するには不向きといわざるをえない。一方、フレーム間差分法は、時系列上のどの部分においても一様な正確さで差分が得られるため、環境の変化に強い方法であるといえる。そこで、本研究では移動領域の抽出にフレーム間差分法を適用した。

## 4. 人工動画をを用いた実験と考察

提案手法を実動画像に適用する前に、本章で人工動画をを用いた実験を行う。まず、モデルパラメータ値が結果に与える影響を確認するために、モデルパラメータ値をさまざまに変化させて処理した結果を示す。次に、提案手法の基本性質を確認するために、背景に含まれるノイズと対象の輝度値を変化させた際の処理結果を従来手法と比較して示す。そして、速度変化に関する提案手法の挙動を明らかにするために、対象物の移動速度を変化させた際の処理結果を示す。

この実験に用いた入出力画像は、ノイズが分かりやすいようにガンマ補正後の画像を示した。

表1 パラメータセット一覧  
Table 1 The list of parameter sets.

パラメータセット名	$\beta$	$r_M$	$r_W$	$r_R$	$\log_2/\alpha_L$	$\log_2/\alpha_F$	$\log_2/\alpha_P$	$V_L$	$V_F$	$V_P$	$\theta$
PSet1	0	0	0	0	0.001	0.001	0.001	0	0	5	0.9
PSet2	0	1	0	0	0.001	0.001	0.001	0	5	5	0.9
PSet3	0	0	0	1	0.001	0.001	0.001	0	0	5	0.9
PSet4	0.25	1	3	1	0.4	1.2	0.3	4	100	50	0.9
PSet5	0.25	1	3	1	0.4	1.2	1.2	4	100	50	0.9
PSet6	0.25	1	3	1	0.4	2	2	4	100	50	0.9
PSet7	0.40	1	3	2	0.2	0.2	0.1	1.9	5	15	0.6
PSet8	0.25	1	1	1	0.4	1.2	0.3	3	100	50	0.9



図4 人工動画(1, 15, 30フレーム)

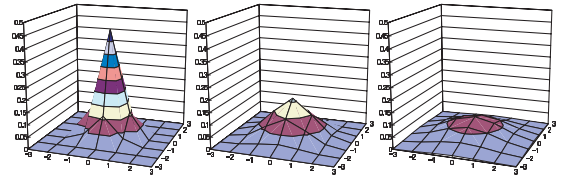
Fig. 4 Examples of the artificial movie (1, 15, 30 frame).

人工動画は、ランダムノイズを付加した黒い背景画像に、屋外で車を撮影した画像から車の部分のみ手動で切り出した領域を、1ピクセルずつ位置を右にずらしながら重ね合わせて作ったフレーム画像からなる。なお、ランダムノイズとは全体の画素のうち「発生率」に相当する割合の画素をランダムに選択し、それらの画素に対してレベル0～「最大輝度」の一樣乱数をフレームごとに加えた。そのため動画として見た場合、ノイズはランダムに出現する。ここで扱うノイズは、発生率と最大輝度によって決まるため、以降、NOISE(発生率(%), 最大輝度(0~255))と表現する。

なお、このようなノイズ生成法を選択した理由としては、本実験が提案手法の振舞いを調査することを目的としており、その振舞いは近傍ニューロンから伝わるエネルギー値と入力値によって変化する個々のニューロンの振舞いに基づくため、ニューロンにどのような値が入力されているか明確になるよう、輝度レベルは0から最大値までの一樣乱数もちいて決定した。また、差分画像に現れる照明環境の変動などによって生じるノイズには、時系列でみた場合空間的にランダムに発生する場合がある。たとえば、雨上がりのアスファルトに日が差すと光の反射がランダムに発生するであろう。本実験ではそのようにランダムに発生するノイズに対する振舞いを調査するため、ノイズの加わる画素を発生率に応じてランダムに選択した。

#### 4.1 モデルパラメータ値を変化させる実験

本節では、モデルパラメータ値が結果に与える影響を確認するため、パラメータ値をさまざまに変化させて人工動画を処理する実験について述べる。



(a)  $r_M, r_W, r_R=1$  (b)  $r_M, r_W, r_R=2$  (c)  $r_M, r_W, r_R=3$

図5 ガウシアンカーネルの例

Fig. 5 Examples of Gaussian kernel.

#### 4.1.1 実験方法と結果

入力データは輝度値に変化があるハッチバック車とNOISE(5%, max50)を付加した背景からなる連続画像30フレームである(図4)。使用したパラメータセットを表1にまとめる。なお、シナプス結合重み $M, W, R$ には要素の総和を1に正規化したガウシアンカーネルを用いた。そこで、表中ではガウシアンカーネルのそれぞれの半径( $r_M, r_W, r_R$ )、つまりいくつ迄離れたニューロンと結合しているかを示した。図5に半径1, 2, 3のカーネルを示す。

各パラメータセットは、次の5つの点を確認する目的で設定している。なお、以下の実験結果はすべて30フレーム目の外部刺激(図6(a))を入力した際のものである。

##### (1) 外部刺激の様子

外部刺激に関する正規化定数  $V_P = 5$ 、閾値  $\theta = 0.9$  とし、他項の影響を排除したパラメータ値(PSet1)に設定して、動作させた際の各ニューロンのFeeding部のエネルギー値  $F[t]$  を図6(b)示す。なお、結果は画像水平軸 × 画像垂直軸 × エネルギー軸の三次元で表したグラフを画像水平軸 × エネルギー軸平面に射影した二次元グラフで示した(以下同じ)。

##### (2) 近傍からの出力に入力を関連付けたことによる効果

PSet2は、PSet1に対してFeeding部の結合重み $M$ の半径を1とし、正規化定数 $V_F$ を5としたものである。各ニューロンのFeeding部の

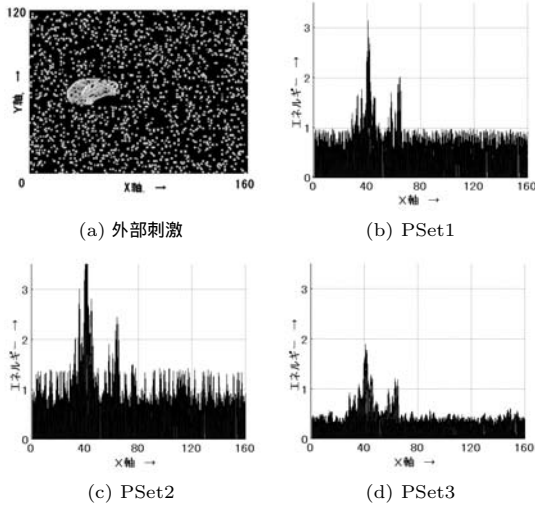


図 6 各ニューロンの Feeding 部の値 (30 フレーム目)

Fig. 6 Value of Feeding part in each neurons.

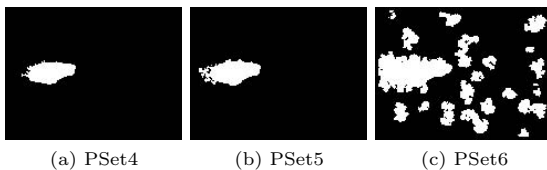


図 7 過渡応答性が出力に与える影響

Fig. 7 Effect of the transient response of the neuron.

エネルギー値  $F[t]$  を図 6 (c) に示す .

- (3) 近傍からの外部刺激による効果  
PSet1 に対し, 入力の場合重み  $R$  の半径を 1 とした (PSet3) 際の各ニューロンの Feeding 部のエネルギー値  $F[t]$  を図 6 (d) に示す .
- (4) 過渡応答特性の影響  
対象を適切に抽出できるように調整を行った PSet4, それより過渡応答特性を強めた PSet5, さらに強めた PSet6 に設定して処理した出力結果をそれぞれ図 7 に示す .
- (5) Linking 部における改良点の効果  
PSet4 を設定し, 3 フレーム目と 5 フレーム目を処理した際の各ニューロンの Linking 部の値  $L[t]$  を図 6 (b) ~ (d) の  $F[t]$  と同様に図 8 に示す .

#### 4.1.2 考察

モデルパラメータ値を変化させる実験結果より, 以下の 5 つの点が確認できる .

- (1) 外部刺激の様子  
Feeding 部の式 (9), (10) に PSet1 の値を代入

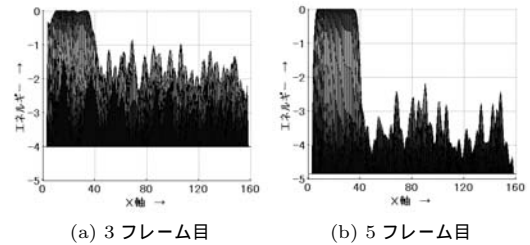


図 8 各ニューロンの Linking 部の値

Fig. 8 Value of Linking part in each neurons.

すると

$$F_{ij}[n] = 5 \cdot 1 \cdot S_{ij}[n]$$

となるから, 与えられたノイズによる外部刺激の最大値は 50 であるため, Feeding の値に与えるノイズの影響は最大で  $50/255 \times 5 \approx 1$  となるはずである . 図 6 (b) において, ノイズ位置にあたるニューロンのエネルギー値は確かに最大 1 程度である . 一方, 移動物体の画素から外部刺激を受けているニューロンの値は最大 3 程度となっている .

- (2) 近傍からの出力結合に入力を関連付けたことによる効果

Feeding 部の式 (9), (10) に PSet2 の値を代入すると

$$F_{ij}[n] = 5 \cdot S_{ij}[n] + 5 \cdot \sum_{kl} M_{ijkl} Y_{kl}[n-1] S_{ij}[n] \quad (13)$$

となる . ここで  $\frac{S_{ij}}{255} \times 5 > 1$  より, 外部刺激として  $S_{ij} \geq 50$  が与えられたニューロンが出力が増幅されて伝播されるニューロンである . しかし, 与えられたノイズによる  $S_{ij}$  の最大値は 50 であるので, 近傍ニューロンの出力値 1 はすべて 1 以下に抑制されて伝播される . また, 空間的にランダムに出現するノイズが前の時間にあった場所の隣に出現する率は低いのでさらに抑制されて伝播されるはずである . 図 6 (c) より, Feeding 部に現れるノイズの最大値は 1.5 程度になっており, 図 6 (b) におけるノイズの最大値 1.0 からの増加分, つまり出力結合に入力を関連付けたことによる増分は 0.5 であるから, 確かに近傍ニューロンの出力値 1 は減衰して伝播されている . つまり, ノイズの影響が抑制されていることが確認できる . 一方, 移動物体から外部刺激を受けているニューロンの  $F_{ij}$  の最大値はグラフには示されていないが 4.5 程

度であり、出力結合に入力を関連付けたことによる増分は 1.5 となるから、近傍ニューロンの出力値 1 が増幅されて伝播されている。つまり、移動物体の影響は促進されていることが確認できる。

- (3) 近傍からの外部刺激による効果  
PSet3 を設定して処理した結果である図 6(d) を (1) の実験結果である図 6(b) と比較すると、Feeding の値が全体的に小さくなっているが、特にノイズ部分の減少が著しい。したがって、ノイズの影響を低減できたといえる。
- (4) 過渡応答特性の影響  
移動物体が適切に抽出されるように調整された時定数パラメータ (PSet4) による処理結果 (図 7(a)) と時定数を少し大きくした PSet5 による処理結果 (図 7(b)) を比べると、図 7(b) に進行方向と逆側に尾ひれのようなものが抽出されている。さらに時定数を大きくした PSet6 による処理結果にはその尾ひれがより大きくなる現象がみられた (図 7(c))。この結果より、時定数を適切に設定すれば導入した過渡応答特性により抽出領域の欠損を補完する効果が期待できる。
- (5) Linking 部に行った改良の効果  
PSet6 に設定した際の処理結果 (図 8) より、Linking 部の値は、自分自身および近傍が前の時刻に発火しているところだけ 0 となっており、それ以外は負の値を出力することが確認できる。つまり、発火している領域の周辺部分は、発火していないニューロンの影響により負の値を持つことになる。Linking 部の値が負である場合、式 (5) の変調作用によりニューロンの発火を抑える方向に働く。この作用によりノイズを減少させる効果が期待できる。なお、図 8(a)、図 8(b) と処理が進むとノイズの影響で発火していたニューロンの Linking 部のエネルギーが負の値に変化していることが確認できる。

以上のことから基本 PCNN に対して行った変更によるそれぞれの効果の相互作用により、提案モデルは外部刺激からノイズを除去し、まとめて移動する外部刺激の欠損の補完を行う機能を有することが期待できる。次節において本手法を人工動画に適用する実験を行い、それらの機能の有効性を確認する。

#### 4.2 背景ノイズと移動物体の輝度値を変化させる実験

本節では背景ノイズと移動物体の輝度値が結果にど

のような影響を与えるかを確認するために、背景ノイズと移動物体の輝度値の条件を変化させて処理する実験について述べる。

実験に使用するデータは、ノイズ発生率が 1, 3, 5, 7, 10%, ノイズ最大輝度が 10, 20, 50 であるノイズを付加して生成した背景と、輝度値に差がみられるハッチバック車と輝度値にほとんど差がないワンボックス車を移動物体として用意し、それぞれを合成した人工動画からなる計 30 シーンである。なお、処理結果の評価値としては、入力画像中の移動物体に相当するピクセルの抽出率 (対象抽出率:  $O_e$ ) と背景に相当するピクセルを移動物体として抽出した率 (ノイズ抽出率:  $N_e$ ) に関して、シーンごとの平均値を求めた。評価値がそれぞれ、移動物体抽出率 100%, 背景抽出率 0% であれば理想的に移動物体を抽出できたことを表す

なお、本実験では従来手法として、フレーム間差分処理を行った後にモルフォロジフィルタ<sup>9)</sup>処理を行う方法を用いた (以下、従来手法 I)。モルフォロジフィルタを以下に示す<sup>20)</sup>。

$$S_k = [(R_k < c > C_3) < o > C_3] < c > C_5 \quad (14)$$

ここで、 $R_k$  は原画像、 $S_k$  は処理結果、 $C_3$ 、 $C_5$  は直径がそれぞれ 3, 5 の円形オペレータを表す。

##### 4.2.1 実験方法と結果

本実験では提案手法のパラメータを PSet4 に設定して処理を行った。結果としてノイズ抽出率・対象抽出率を提案手法の処理が安定する 10 フレーム (ハッチバック車)、20 フレーム (ワンボックス車) 目以降の平均値を表 2 にまとめた。また従来手法 I は、フレーム間差分の閾値を 0.000 ~ 0.080 まで 0.004 ごとに变化させながら処理を行った。結果としてノイズ抽出率 ( $N_e$ ) の基準値を満たすことができる閾値の最小値と対象物抽出率 ( $O_e$ ) の基準値を満たすことができる閾値の最大値をそれぞれ表 3 にまとめた。また、提案手法・従来手法 I それぞれにおいて抽出が行えたか否かを表 4 にまとめた。ここで、抽出が行えたか否かの判定方法としては、移動物体抽出率が 85% 以上かつノイズ抽出率が 1% 以下であるものを「可」それ以外を「不可」とした。

なお、出力画像の代表的例として次の 3 つの入力データの結果を示す。まず、ハッチバック車に背景 NOISE (10%, max10) を付加したデータの 25 フレーム目の処理結果を図 9 にまとめた。提案手法による処理結果を (a) に、従来手法 I に関しては最適閾値 0.044 を設定した処理結果を (b) に示した。なお、ここで最適閾値とは、ノイズ抽出率の基準を満たす最小値と対象物



表 2 提案手法による処理結果  
Table 2 Experimental results by the proposed method.

(a) ハッチバック車

ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	Ne	Oe	Ne	Oe	Ne	Oe
1	0.57	99.20	0.58	99.20	0.61	99.20
3	0.59	99.20	0.62	99.20	0.66	98.84
5	0.62	99.20	0.65	99.20	0.84	99.28
7	0.63	99.20	0.74	99.20	0.96	99.00
10	0.68	99.20	0.81	99.20	<b>2.59</b>	99.32

(b) ワンボックス車

ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	Ne	Oe	Ne	Oe	Ne	Oe
1	<b>1.45</b>	85.66	<b>1.46</b>	85.66	<b>1.51</b>	85.66
3	<b>1.47</b>	85.66	<b>1.51</b>	85.66	<b>1.59</b>	85.66
5	<b>1.52</b>	85.66	<b>1.63</b>	85.66	<b>1.94</b>	85.84
7	<b>1.56</b>	85.66	<b>1.56</b>	85.66	<b>2.20</b>	85.84
10	<b>1.68</b>	85.66	<b>1.90</b>	85.66	<b>3.52</b>	85.44

\* 基準を満たさないものは太字

表 3 従来手法 I による判定基準を満たす閾値一覧  
Table 3 Experimental results by the conventional method I.

(a) ハッチバック車

ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	Ne 基準値内	Oe 基準値内	Ne 基準値内	Oe 基準値内	Ne 基準値内	Oe 基準値内
1	0.000 以上	0.052 以下	0.000 以上	0.052 以下	0.000 以上	0.052 以下
3	0.020 以上	0.052 以下	0.040 以上	0.052 以下	不可	<b>0.052</b> 以下
5	0.028 以上	0.052 以下	<b>0.056</b> 以上	<b>0.052</b> 以下	不可	<b>0.072</b> 以下
7	0.032 以上	0.052 以下	<b>0.060</b> 以上	<b>0.052</b> 以下	不可	<b>0.072</b> 以下
10	0.032 以上	0.052 以下	<b>0.068</b> 以上	<b>0.056</b> 以下	不可	<b>0.080</b> 以下

(b) ワンボックス車

ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	Ne 基準値内	Oe 基準値内	Ne 基準値内	Oe 基準値内	Ne 基準値内	Oe 基準値内
1	0.000 以上	0.008 以下	0.000 以上	0.008 以下	0.000 以上	0.008 以下
3	<b>0.024</b> 以上	<b>0.008</b> 以下	0.044 以上	<b>0.008</b> 以下	不可	<b>0.008</b> 以下
5	<b>0.028</b> 以上	<b>0.008</b> 以下	<b>0.056</b> 以上	<b>0.012</b> 以下	不可	<b>0.012</b> 以下
7	<b>0.032</b> 以上	<b>0.012</b> 以下	<b>0.064</b> 以上	<b>0.012</b> 以下	不可	<b>0.012</b> 以下
10	<b>0.032</b> 以上	<b>0.012</b> 以下	<b>0.068</b> 以上	<b>0.012</b> 以下	不可	<b>0.012</b> 以下

\* 基準を同時に満たせない場合は太字

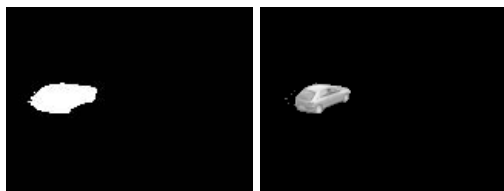
表 4 抽出判定結果一覧  
Table 4 Determination results of extracting moving object areas.

(a) ハッチバック車

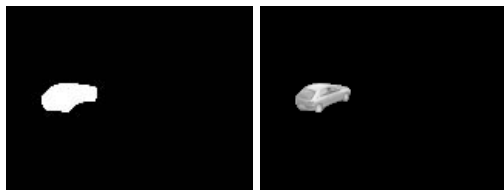
ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	提案手法	従来手法 I	提案手法	従来手法 I	提案手法	従来手法 I
1	可	可	可	可	可	可
3	可	可	可	可	可	不可
5	可	可	可	不可	可	不可
7	可	可	可	不可	可	不可
10	可	可	可	不可	不可	不可

(b) ワンボックス車

ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	提案手法	従来手法 I	提案手法	従来手法 I	提案手法	従来手法 I
1	不可	可	不可	可	不可	可
3	不可	不可	不可	不可	不可	不可
5	不可	不可	不可	不可	不可	不可
7	不可	不可	不可	不可	不可	不可
10	不可	不可	不可	不可	不可	不可

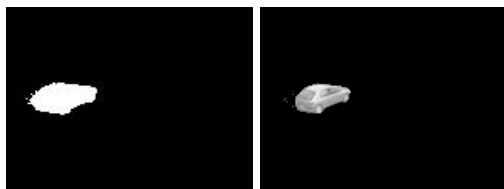


(a) 提案手法

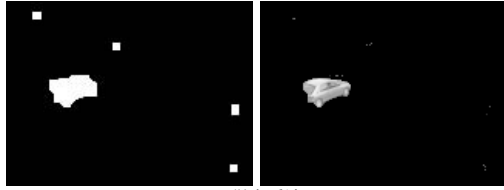
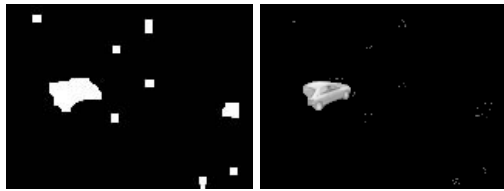
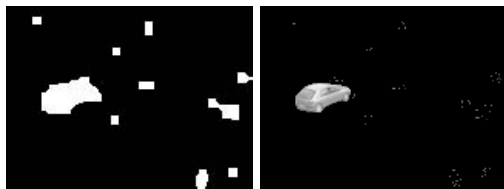


(b) 従来手法 I (閾値=0.044)

図 9 ハッチバック車の抽出結果: NOISE ( 10% , max10 )  
Fig.9 Experimental results of the hatchback car.



(a) 提案手法

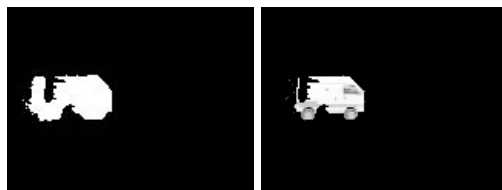


(b) 従来手法 I

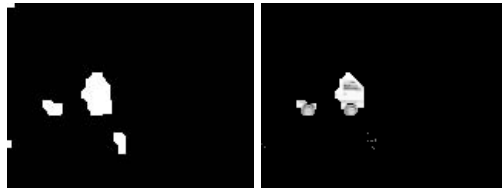
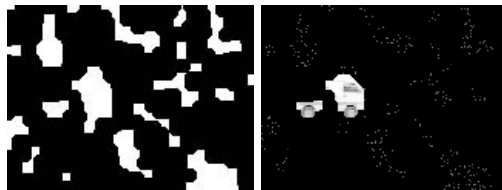
( 上: 閾値=0.052, 中: 0.056, 下: 0.060 )

図 10 ハッチバック車の抽出結果: NOISE ( 7% , max20 )  
Fig. 10 Experimental results of the hatchback car.

出率の基準を満たす最大値のおおよそ中間の値とした。次に、ハッチバック車に NOISE ( 7% , max20 ) を付加したデータの 25 フレーム目の結果を図 10 に示した。提案手法による処理結果を (a) に、従来手法 I



(a) 提案手法



(b) 従来手法 I

( 上: 閾値=0.012, 中: 0.040, 下: 0.064 )

図 11 ワンボックス車の抽出結果: NOISE ( 7% , max20 )  
Fig. 11 Experimental results of the minivan.

関してはノイズ抽出率の基準を満たす最小値 0.060 と対象抽出率の基準を満たす最大値 0.052, そして最適閾値 0.056 を設定した処理結果をそれぞれ (b) に示した。そして、ワンボックス車に NOISE ( 7% , max20 ) を付加したデータの 24 フレーム目の結果を図 11 に示した。提案手法による処理結果は (a) に、従来手法 I に関してはノイズ抽出率の基準を満たす最小値 0.064 と対象抽出率の基準を満たす最大値 0.012, そして最適閾値 0.040 を設定した処理結果をそれぞれ (b) に示した。

#### 4.2.2 考 察

まず、ノイズの変化による影響について考える。ハッチバック車の場合の提案手法の処理結果 ( 表 2 (a) ) において、NOISE ( 10% , max50 ) を付加した場合以外のノイズ抽出率は判定基準内の 0.57 から 0.96 の値に収まっている。同一のパラメータを使った処理結果において、入力ノイズ量を変化させたにもかかわらず、出力結果のノイズ抽出率に大きな変化がなかったということから、提案手法はノイズの変化の影響を受けに

くといえる。ただし、NOISE(10%, max50)を付加した場合には、SONNのノイズ除去能力を超えてしまうため影響があったと考えられる。一方、従来手法Iの処理結果(表3(a))から、ノイズ抽出率を基準値内にする最適閾値の最小値が、ノイズの発生率が1%の場合を除いては、ノイズ量の増加にともなって増加しているので、ノイズの変化の影響を受けやすいといえる。

次に、対象物の輝度変化による影響について考える。提案手法による結果(表2(b))においては、対象物によって対象抽出率に差があるものの、ともに基準値を満たす範囲であるので、階調値の変化量に結果はそれほど影響されないといえる。一方、従来手法Iによる結果(表3(b))から対象物の全体を抽出するには、階調に変化があるハッチバック車では閾値が0.052、階調値にほとんど変化がないワンボックス車では0.008に設定する必要がある。このことから、階調値の変化量に結果が強く影響されるといえる。

次に、提案手法と従来手法Iの抽出結果を比較すると、表4(a)に示すように、ハッチバック車を抽出不可能なNOISE(3%, max50) (5%, max20) (7%, max20), (7%, max50) (10%, max20) (10%, max50)を付加した背景データからも提案手法を用いれば抽出することが可能であった。なお、ワンボックス車のノイズ発生率1%のデータに対して、提案手法が抽出を行っていない点に関しては、以下に代表的な出力例を比較することで説明する。まず、NOISE(10%, max10)を付加した背景からハッチバック車の抽出を行った出力結果(図9)から、提案手法・従来手法Iの出力画像ともハッチバック車全体を抽出しており、ノイズもないことが確認できる。次に、NOISE(7%, max20)を付加した場合の出力結果(図10)を比較する。提案手法による出力画像には、ハッチバック車全体が抽出されており、ノイズもないことが確認できる。しかし、従来手法Iによる出力画像においては、対象抽出率の基準値を満たす最小閾値0.052で処理した結果には、ノイズが多く抽出され、また、ノイズ抽出率の基準値を満たす最大閾値0.060で処理した結果では、対象に欠損がある。閾値を両者の中間0.056に設定した場合においても、ノイズ除去と対象の全体を抽出することを両立させることはできていない。そして、ワンボックス車の抽出を行った出力結果(図11)を比較すると、提案手法では、ノイズの除去は行われているが、対象の一部に欠損がみられる。一方、従来手法Iでは、対象抽出率の基準値を満たす最小閾値0.012に設定すると、ほぼ画面全体が抽出されてしまい、また

ノイズ抽出率の基準値を満たす最大閾値0.060に設定すると、対象のほとんどが欠損してしまっている。両者の間である、0.040に設定した場合においても、ノイズ除去と対象の全体を抽出することはできなかった。

以上のことから提案手法は従来手法Iに比べ、ノイズ除去と対象の欠損を補う能力が高いことが分かった。なお、ワンボックス車の場合は提案手法を用いた場合でも多少欠損箇所が見受けられるが、従来手法Iに比べれば補完能力は高いといえる。このような能力の違いは、モルフォロジフィルタによるノイズ除去・抽出領域の連結が各ピクセルの近傍関係に基づいているのに対し、提案モデルは空間的要素だけではなく、各ニューロンの出力と外部刺激との関係や過渡応答特性といった時間的要素に基づいているためだと考える。

ここで、表2(b)において、提案手法の結果が出力結果からはノイズの除去が確認できるにもかかわらず、ノイズ抽出率が大きな値となっている点に関する検証を行う。ワンボックス車をノイズを加えない背景に合成したデータを用意し処理したところ、ノイズがない場合でもノイズ抽出率が1.43%となることが分かった。このことから、ノイズ抽出率が高くなった要因は付加したノイズ部分を抽出したことによるものではなく、対象の輪郭が不明瞭になっているためだと考える。特に移動物体の進行方向と逆側に抽出されている尾ひれのようなものがノイズ抽出率を増加させていると考える。この現象は、ハッチバック車を抽出した場合にもいえることだが、移動物体の大きさが小さいため、基準値を超えるほどの影響を与えなかったと考えられる。

#### 4.3 対象物の移動速度を変化させる実験

提案手法は時間軸情報の利用を特徴としているが、これはフレーム間差分の結果における近傍画素の時間軸方向の相関値を反映したものである。ところが、ある部位との相関計算の対象となる次フレームでの画像領域は、物体の形状・移動速度によって異なるため、一般には物体の同じ部分間の計算とはならない。そこで本節では、移動物体の移動速度を変化させ、提案手法の物体の移動速度に対する振舞いを実験により明らかにし、提案手法がどのような相関に基づいて動作しているかに関する考察を行った。

##### 4.3.1 実験方法と結果

本実験は4.2節の実験方法に準じて行った。実験に使用するデータは、ノイズ発生率が1, 3, 5, 7, 10%, ノイズ最大輝度がすべて20のノイズを背景にそれぞれ付加し、対象物体であるハッチバック車、ワンボックス車の移動速度を1, 2, 3 [pixel/frame]と変化させて生成した計30シーンである。なお、実験に用いた

表 5 対象物体の移動速度を変化させた際の処理結果  
Table 5 Experimental results for changing speed of the target.

(a) ハッチバック車

NOISE	移動速度 [pixel/frame]					
	1		2		3	
	Ne	Oe	Ne	Oe	Ne	Oe
1%, max20	0.70	100.00	1.44	100.00	2.33	100.00
3%, max20	0.78	100.00	1.58	100.00	2.46	100.00
5%, max20	0.87	100.00	1.75	100.00	2.75	100.00
7%, max20	1.02	100.00	2.10	100.00	3.03	100.00
10%, max20	1.50	100.00	2.75	100.00	4.00	100.00

(b) ワンボックス車

NOISE	移動速度 [pixel/frame]					
	1		2		3	
	Ne	Oe	Ne	Oe	Ne	Oe
1%, max20	1.79	91.03	3.37	99.68	5.00	100.00
3%, max20	1.95	91.06	3.65	99.68	5.37	100.00
5%, max20	2.10	91.06	3.86	99.68	5.73	100.00
7%, max20	2.34	91.13	4.17	99.68	6.21	100.00
10%, max20	3.11	91.14	5.29	99.68	7.55	100.00

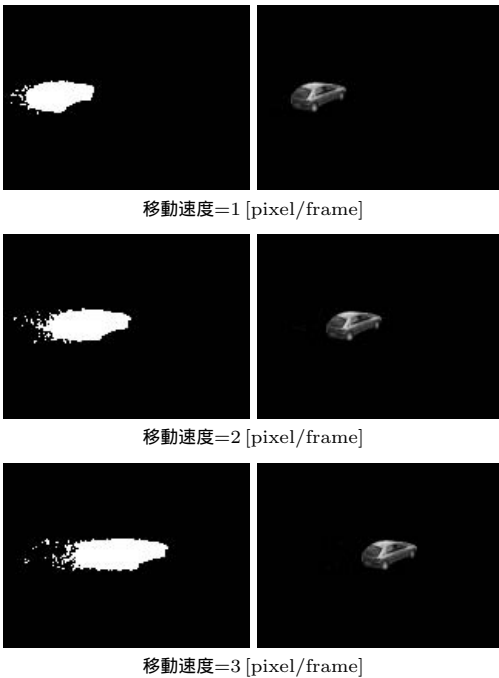


図 12 ハッチバック車の抽出結果：NOISE ( 7% , max20 )  
Fig. 12 Experimental results of the hatchback car.

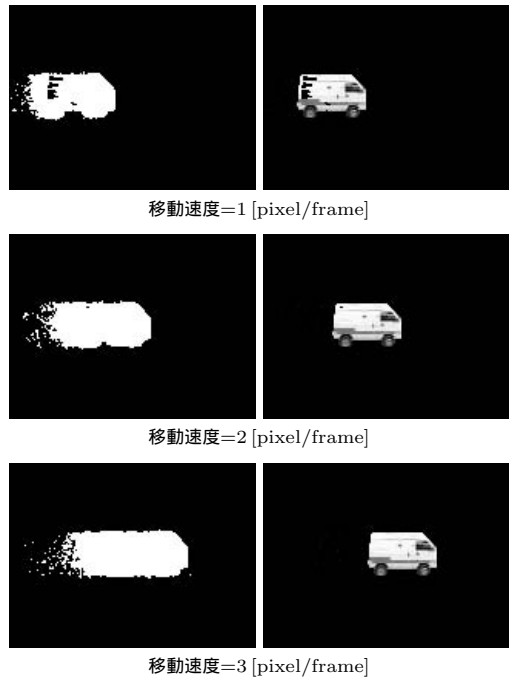


図 13 ワンボックス車の抽出結果：NOISE ( 7% , max20 )  
Fig. 13 Experimental results of the minivan.

パラメータは表 1 の PSet8 である .

処理結果の評価値として、対象抽出率：Oe とノイズ抽出率：Ne を表 5 に示し、代表的な処理結果として、NOISE ( 7% , max20 ) のときの出力画像を図 12、図 13 に示した .

4.3.2 考 察

処理結果の評価値 ( 表 5 ) から対象の移動速度の変化に対する提案手法の振舞いについて考える . まず、ハッチバック車の結果について、対象抽出率がすべて 100% となっているので、いずれの移動速度において

も提案手法は対象の移動物体領域の全体を抽出できたことが分かる . なお、対象物の抽出の様子は代表的な出力を示した図 12 から確認できる . ノイズ抽出率は移動速度が早くなるにつれて大きくなっている . しかし、図 12 から分かるように、ノイズの領域を抽出しているのではなく、対象の輪郭が移動方向の前後に長く抽出されていることが原因である . このような抽出が行われるのは提案手法がフレーム間差分法に基づいているため対象の移動速度が増すと対象の前後部分に対象の移動距離に応じた差分が検出され、これが提

案手法の後処理部に入力されるためだと考えられる。次にワンボックス車の結果について、対象抽出率は移動速度が 2, 3 [pixel/frame] の方が 1 [pixel/frame] のときより大きな値となっている。これは、図 13 から分かるように対象の欠損が補完されていることによる。速度が速くなると提案手法の対象の欠損の補完機能がより効果的に働く理由としては、提案手法が、差分画像において移動物体領域に属する画素はある一定の差分値を移動物体が通過する間連続して持つという仮定に基づいて動作しているため、速度が速ければ欠損部分にあたる領域に次フレームにおいて対象の領域がくる可能性が高くなるので、欠損領域を瞬間的な変化として補完するためだと考えられる。また、ノイズ抽出率は移動速度が早くなるにつれて大きな値となっているが、この理由としてはハッチバック車のときと同じことがいえる。

以上、提案手法の欠損領域補完機能が有効に働くには、差分画像において欠損になっている領域と抽出されている領域がフレーム間において重なればよいことが分かる。

つまり、提案手法の欠損補完機能は対象の同じ部分間の相関のみではなく、対象の内であれば別々の部分の相関であっても動作するといえる。そのため、欠損領域が大きい場合には物体の移動速度が速い方が有利に働くが、相関の対象となる差分画像領域が対象物体領域の一部によるものなのか、ノイズによるものなのかの区別はなされないで、移動方向の逆がわりに尾ひれのようなものを抽出する領域も大きくなるという問題がある。

## 5. オプティカルフローを用いた手法との比較

本章では、移動物体領域抽出手法の従来手法として代表的なオプティカルフローを用いた手法(以下、従来手法 II)と提案手法を比較することにより、提案手法の優位性を明らかにする。

### 5.1 実験方法と結果

従来手法 II としては、文献 21) による方法を用いた。ここで用いられているオプティカルフローの推定方法はグラディエント法の局所最適化法に分類される。また、フローの信頼性の指標としてフローの共分散行列の最大固有値を用いている。

移動物体領域抽出手順は以下のとおりである。

(1) 原画像を  $\sigma = 1.0$  画素のガウシアンフィルタで

平滑化する。

- (2) オプティカルフローを計算する。
- (3) フローの共分散行列の最大固有値を計算し、これを信頼性の指標とする。
- (4) 信頼性の指標が閾値 (0.05) 以下の部分を移動物体領域として抽出する。

オプティカルフロー推定のパラメータは、ウィンドウのタイプ:  $5 \times 5$  の矩形領域、ウィンドウをシフトするフレーム: 2, 反復計算の回数: 2 で、微分カーネルのタイプは以下のようなものとした。

$$E_x = (E[x+1, y, t] - E[x, y, t] + E[x+1, y+1, t] - E[x, y+1, t] + E[x+1, y, t+1] - E[x, y, t+1] + E[x+1, y+1, t+1] - E[x, y+1, t+1])/4$$

$$E_y = (E[x, y+1, t] - E[x, y, t] + E[x+1, y+1, t] - E[x+1, y, t] + E[x, y+1, t+1] - E[x, y, t+1] + E[x+1, y+1, t+1] - E[x+1, y, t+1])/4$$

$$E_t = (E[x, y, t+1] - E[x, y, t] + E[x+1, y, t+1] - E[x+1, y, t] + E[x, y+1, t+1] - E[x, y+1, t] + E[x+1, y+1, t+1] - E[x+1, y+1, t])/4$$

実験データは、4.2 節と同じものを用いた。処理結果の評価値を表 6 に、提案手法による処理結果を表 7 に再掲する。また、出力の代表例として NOISE (7%, max20) を付加したデータを処理した結果を図 14、図 15 に示す。

### 5.2 考察

まず、結果の評価値を比較する。表 6 と表 7 から、NOISE (10%, max50) 以外のすべてのデータにおいて、提案手法による処理結果のノイズ抽出率は小さく、また、対象物抽出率が大きい。つまり、提案手法はオプティカルフローを用いた手法に比べ結果の評価値が良かった。なお、ここには示していないが出力画像を調査したところ、ノイズを加えた部分を抽出しているのは、提案手法で NOISE (10%, max50), オプティカルフローを用いた手法においては、NOISE (10%, max50) (10%, max20) を加えた場合であった。

次に、対象の抽出具合を代表的な出力例である図 14、15 を評価する。オプティカルフローを用いた方法の結果は、ハッチバック車、ワンボックス車ともに、移動物体領域を一回り大きく抽出している。また、ハッチバック車の前輪の後ろあたり、ワンボックス車の後ろ上部の一部の領域が欠けている。一方、提案手法は、ハッチバック車、ワンボックス車ともに進行方向の逆側に尾ひれのようなものを抽出している。また、ワンボックス車の後ろ下部に欠けている領域がある。しかし、全体的な傾向としては、オプティカルフローを用いた手法に比べ、提案手法は対象移動物の輪郭をよく

<http://www.ail.cs.gunma-u.ac.jp/ohata/>で公開されているプログラムを基に実装を行った。太田氏に謹んで感謝の意を表す。

表 6 従来手法 II による処理結果  
Table 6 Experimental results of the conventional method II.

(a) ハッチバック車

ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	Ne	Oe	Ne	Oe	Ne	Oe
1	1.24	95.64	1.24	95.63	1.24	95.67
3	1.24	95.63	1.26	95.67	1.25	95.69
5	1.25	95.63	1.27	95.67	1.27	95.68
7	1.24	95.70	1.29	95.69	1.30	95.67
10	1.25	95.68	1.39	95.72	1.39	95.69

(b) ワンボックス車

ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	Ne	Oe	Ne	Oe	Ne	Oe
1	2.05	84.68	2.07	84.63	2.06	84.62
3	2.07	84.63	2.09	84.60	2.09	84.63
5	2.07	84.61	2.12	84.69	2.11	84.55
7	2.07	84.61	2.13	84.58	2.14	84.68
10	2.09	84.70	2.25	84.79	2.25	84.75

表 7 提案手法による処理結果 (再掲)  
Table 7 Experimental results of the proposed method.

(a) ハッチバック車

ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	Ne	Oe	Ne	Oe	Ne	Oe
1	0.57	99.20	0.58	99.20	0.61	99.20
3	0.59	99.20	0.62	99.20	0.66	98.84
5	0.62	99.20	0.65	99.20	0.84	99.28
7	0.63	99.20	0.74	99.20	0.96	99.00
10	0.68	99.20	0.81	99.20	2.59	99.32

(b) ワンボックス車

ノイズ発生率 [%]	ノイズ最大輝度					
	10		20		50	
	Ne	Oe	Ne	Oe	Ne	Oe
1	1.45	85.66	1.46	85.66	1.51	85.66
3	1.47	85.66	1.51	85.66	1.59	85.66
5	1.52	85.66	1.63	85.66	1.94	85.84
7	1.56	85.66	1.56	85.66	2.20	85.84
10	1.68	85.66	1.90	85.66	3.52	85.44



入力画像 従来手法 II 提案手法  
図 14 ハッチバック車の例 (20-21 フレーム)

Fig. 14 Experimental results of the hatchback car (20-21 frame).



入力画像 従来手法 II 提案手法  
図 15 ワンボックス車の例 (20-21 フレーム)

Fig. 15 Experimental results of the minivan (20-21 frame).

とらえている。

オプティカルフローを用いた手法が領域を大きめに抽出する理由として、オプティカルフローの推定は、画像中の輝度勾配が運動中に変化しないことを仮定しているため、物体の境界など画像中のエッジにおける推定精度が他の領域と比べて大幅に低下することが考えられる。なお、この領域はオプティカルフローの推定パラメータの1つであるウィンドウサイズを小さくすると、ある程度小さくなるが、ウィンドウサイズはオプティカルフロー推定精度に以下に示すように影響

するので注意する必要がある。

- 領域を小さく取った場合：用いる画素が少ないので計算時間が早く、オプティカルフローが一定という仮定が満たされやすい。しかし、ノイズや動きが大きくなると対応ができにくくなる。
  - 領域を大きくとった場合：用いる画素が多くなるので計算時間がかかる。オプティカルフローが一定という仮定を満たさない画素も存在しやすくなるため、推定精度低下の原因になる。
- 対象物の一部が欠ける問題に関しては、オプティカ



図 16 SCENE1 の例 (1, 15, 30 フレーム)

Fig. 16 Example frames of the SCENE1 (1, 15, 30 frame).



図 17 SCENE2 の例 (1, 15, 30 フレーム)

Fig. 17 Example frames of the SCENE2 (1, 15, 30 frame).

ルフローの推定において、輝度一様領域は背景と見なされるためであると考えられる。

一方、提案手法は、エッジ部分の精度が落ちることがなく、また、対象物の抽出欠損は小さい。なお、移動方向の逆がわりに尾ひれのようなものを抽出する問題もあるが、その領域はオプティカルフローが余分に抽出する領域に比べ小さいといえる。以上、本実験により、提案手法はオプティカルフローを用いた方法よりも精度良く移動物体領域を抽出することが確認できた。

## 6. 実動画像を用いたシミュレーションと考察

本章では、人工動画像による実験で示した点が、実動画像に関しても有効であるか確認するために行ったシミュレーションに関して述べる。

### 6.1 実験方法と結果

実動画像データとして、Dataset for 2nd IEEE International Workshop on Performance Evaluation of Tracking and Surveillance を用いたシミュレーションを行った。なお、実験は Pentium III processor 650 MHz を搭載した DOS-V 互換機上で、MATLAB6.5 を用いて行った。

入力としては、データセット名: dataset1, 分類: test, カメラ: camera1 の JPEG データを  $365 \times 274$  のサイズに変換したものから、以下の 2 シーンを抜粋して用いた。SCENE1 (フレーム番号: 450–600 [5 フレームごとに入力]) は、画面に向かって左の歩行者は右方向へ歩いており、画面右の車両は左方向に進行している様子を撮影した 30 フレームである (図 16)。

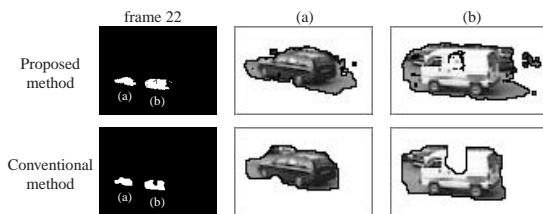


図 18 SCENE1 の結果 (上: 提案手法, 下: 従来手法 I)

Fig. 18 Experimental results of the SCENE1.

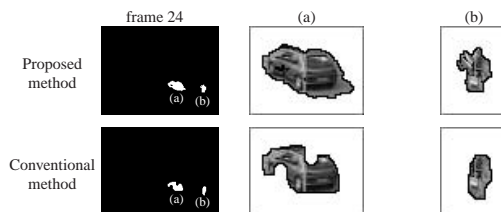


図 19 SCENE2 の結果 (上: 提案手法, 下: 従来手法 I)

Fig. 19 Experimental results of the SCENE2.

SCENE2 (フレーム番号: 2490–2640 [5 フレームごとに入力]) は、歩行者が画面の左方向へ歩いており、2 台の車両が T 字路の奥から歩行者の方向へ曲がったのち直進する様子を撮影した 30 フレームである (図 17)。

シミュレーションの方法としては、提案手法に関しては、パラメータを表 1 の PSet7 に設定し処理を行った。また、従来手法 I としては、4.2 節と同じ手法 ( $\theta=0.1$ ) を用いて処理を行った。SCENE1 の 22 フレーム目と SCENE2 の 24 フレーム目の結果をそれぞれ図 18、図 19 に示す。

### 6.2 考察

まず、SCENE1 の結果である図 18 より、従来手法 I による抽出物体 (b) の車体が大きく欠損してしまっているが、提案手法の出力ではその欠損がほとんどみられないことが分かる。また、本論文の実験結果としては示していないが、従来手法 I の結果には欠損しているフレームと欠損なく抽出されているフレームが混在していた。これは、移動中に車両の向きが変わっているため光のあたり具合が変化した影響を受けたものと考えられる。一方提案手法の結果は安定して移動物体を抽出していた。次に SCENE2 の結果である図 19 より、従来手法 I の抽出物体 (a) の車体には欠損している部分があるが、提案手法の出力には、ほとんど欠損がみられないことが分かる。

以上の結果より、提案手法は従来手法 I に比べノイズ除去能力に優れ、さらに抽出物体の欠損を補完する能力にも優れていることが理解できる。これは、人工動画像を用いた実験から示されたように、従来手法 I

ftp://pets2001.cs.rdg.ac.uk/PETS2001/からダウンロードできる。これは、英国 Reading 大学が作成したもので、ここに謹んで感謝の意を表する。

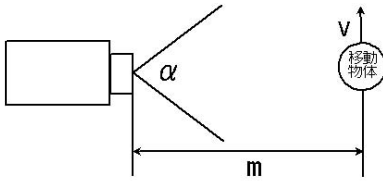


図 20 上から見たカメラと移動物体  
Fig. 20 The camera and the moving object.

がピクセルの空間的な関係に基づいてノイズ除去と欠損の補完を行っているため、対象の輝度変化に処理結果が影響されやすいのに対し、提案手法はピクセルの時空間的な関係に基づいて処理を行うため、対象の輝度変化に影響されにくいと考えられる。

## 7. 提案手法の諸問題

本手法には、以下の点について問題がある。

- (1) 処理の初期段階では、出力が安定しない。
- (2) 移動物体の後ろに尾ひれのようなものが抽出されやすい。
- (3) 対象物の画像における見かけ上の移動速度に対して、十分短い時間間隔で処理されないと、抽出物の輪郭が不明瞭になる。

(3)の関連事項として、複数の移動物体のスピードが異なる場合や、スピードが変化する場合は抽出が不安定になる。これは、フレーム間差分法の特性によるものであるが、フレーム間差分法は、特性をシステム側で調整することができる<sup>22)</sup>。以下に調整法を簡単に説明する。図 20 のような環境を想定する。カメラの画角を  $\alpha$ 、移動物体の速度を  $v$ 、移動物体までの距離を  $m$  とし、移動方向の取り込み解像度を  $b$  とすると、1 ピクセルが表す実世界の長さ  $d$  は以下の式で求まる。

$$d = \frac{2m \tan \frac{\alpha}{2}}{b} \quad (15)$$

一方、画像上での物体の移動ピクセル数を  $p$ 、フレームレートを  $n$  とすると、以下のような関係式が求まる。

$$v = ndp \quad (16)$$

したがって、輪郭部を明瞭に検出するためには、 $p$  の値が 1 以上かつ画像上の移動物体のピクセルサイズに対して十分小さな値になるよう  $n$ 、 $b$  を調整すればよい。また、フレーム間差分法のこのような問題の解決策として 3 フレーム間の差分を利用した方法も提案されている<sup>23)</sup>。複数の移動物体間の速度が大きく異なるなど、システムの側の調整が困難となる場合には、3 フレーム差分法をフレーム間差分法の代わりとして利用する方法も考えられる。

## 8. おわりに

本論文では、PCNN を基にした SONN をフレーム間差分法の後処理として用いる移動物体領域の抽出手法を提案した。提案手法とモルフォロジーフィルタを用いた従来手法を用いて、人工動画による実験と実動画によるシミュレーションを行った。それらの結果、提案手法はモルフォロジーフィルタを用いた従来手法に比べノイズに影響されにくく、また抽出領域の欠損が少ないことが示された。なお、オプティカルフローを用いた手法との比較においても、提案手法は精度良く移動物体領域を抽出することが確認できた。

本提案手法は、移動物体の全体を得ることができないというフレーム間差分法の問題を改善している。この問題は差分法全般にいえることだが、背景差分法に比べフレーム間差分法ではその影響が大きい。そのため、抽出領域の色情報などに基づいてモデル化を行うような手法を用いたシステムにおいては背景差分法が使われている。しかし、背景差分法には背景モデルの更新が難しいという問題がある。そこで、本提案手法を適用すれば背景差分法より環境の変化に強いとされるフレーム間差分法の利点を生かすことができ、システム全体の環境の変化に対する頑健性を向上させることが期待できる。今後は、動画像認識システムの前処理に本手法を適用し、背景差分法を用いた場合と比較しながら、システム全体に与える提案手法の有効性を検証していきたい。

## 参考文献

- 1) 三池秀敏ほか：パソコンによる動画像処理，森北出版 (1993)。
- 2) 鈴木 智，中根一成，ジャック・スクランスキー：時空間輪郭積分法による動柔物体の抽出，情報処理学会論文誌，Vol.34, No.10, pp.2184-2192 (1993)。
- 3) Adiv, G.: Determining 3-D motion and structure from optical flow generalized by several moving objects, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.7, pp.384-401 (1985)。
- 4) Chen, H.-J., Shirai, Y. and Asada, M.: Detecting Multiple Rigid Image Motions from an Optical Flow Obtained with Multi-Scale, Multi-Orientation Filters, *IEICE Trans. Inf. Syst.*, Vol.E76-D, No.10, pp.1253-1262 (1993)。
- 5) 前 泰志，白井良明，三浦 純，久野義徳：オプティカルフローとエッジを用いた複雑背景下での移動物体の追跡，日本ロボット学会誌，Vol.15, No.1, pp.103-108 (1997)。
- 6) Elgammal, A., Harwood, D. and Davis, L.S.:



- Non-parametric Model for Background Subtraction, *Proc. IEEE ICCV'99 FRAME-RATE Workshop* (1999).
- 7) Jain, R., Martin, W.H. and Aggarwal, J.K.: Segmentation through the detection of changes due to motion, *Computer Vision, Graphics and Image Processing* 11, pp.13-34 (1979).
  - 8) 安居院猛, 長尾智晴: 画像の処理と認識, 昭晃堂 (1992).
  - 9) Haralick, R.M. and Shapiro, L.G.: *Computer and Robot Vision*, Addison-Wesley, Reading, MA (1992).
  - 10) Chanda, B.: Application of binary mathematical morphology to separate overlapped objects, *Pattern Recognition Letters* (1992).
  - 11) Lopez, L.R.: Feedforward shunting: a simple second-order neural network motion sensor, *Proc. SPIE*, Vol.1297, pp.350-358 (1990).
  - 12) Eckhorn, R., Reitboeck, H.J., Arndt, M. and Dicke, P.: Feature linking via synchronization among distributed assemblies: simulations of results from cat cortex, *Neural Comput.*, 2, pp.293-307 (1990).
  - 13) Johnson, J.L.: Pulse-coupled neural nets: translation, rotation, scale, distortion and intensity signal invariance for images, *APPLIED OPTICS*, Vol.33, No.26, pp.6239-6253 (1994).
  - 14) Kuntimad, G. and Ranganath, H.S.: Perfect Image Segmentation Using Pulse Coupled Neural Networks, *IEEE Trans. Neural Networks*, Vol.10. No.3. pp.591-598.
  - 15) Kinser, J.M.: Pulse-Coupled Image Fusion, *Optical Eng.*, Vol.36, No.3, pp.737-742 (1997).
  - 16) Kinser, J.M. and Lindblad, T.: Detection of Microcalcifications by Cortical Stimulation, *Neural Networks in Engineering Systems*, Bulsari, A.B. and Kalli, S. (Eds), pp.203-206. ISBN 952-90-8669-9.
  - 17) Maxwell, T., Giles, C.L., Lee, Y.C. and Chen, H.H.: *Nonlinear dynamics of artificial neural systems*, *Am. Institute of Physics*, 0094-243x/86/1510299-17, pp.299-304 (1986).
  - 18) Kinser, J.M. and Johnson, J.L.: Stabilized input with a feedback pulse-coupled neural network, *Opt. Eng.*, Vol.35, pp.2158-2161 (1996).
  - 19) Kinser, J.M. and Lindblad, T.: Implementation of Pulse-Coupled Neural Networks in a CNAPS Environment, *IEEE Trans. Neural Networks*, Vol.10, No.3, pp.584-590 (1999).
  - 20) Heikkila, J. and Silven, O.: A Real-Time Tracker for Visual Surveillance Applications, *Proc. 1st Int. Workshop on PETS*, pp.8-13 (2000).
  - 21) Ohta, N.: Image movement detection with reliability indices, *IEICE Trans. Communications Electronics Information and Systems*, Vol.E74, No.10 (1991).
  - 22) 馬場功淳, 大橋 健, 乃万 司, 松尾英明, 江島俊朗: HeadFinder: フレーム間差分をベースにした人物追跡, 画像センシングシンポジウム 2000, pp.329-334 (2000).
  - 23) 安居院, 横前, 中島: 動画像からの通行車両の抽出, 信学春季大全 D-556 (1989).

(平成 15 年 2 月 3 日受付)

(平成 15 年 4 月 18 日採録)



茶志川孝和

武蔵工業大学工学部電気電子工学科卒業(1995), 同年ニッタン(株)入社. 慶應義塾大学大学院政策・メディア研究科修士課程修了(1999), 同博士課程単位取得退学(2001). 慶應義塾大学 SFC 研究所所員. ニューラルネットワークを用いた画像処理技術の研究に従事. Best Paper Award, Visualization and Imaging Session, 7th International Symposium on Artificial Intelligence in Real-Time Control (1998).



武藤 佳恭

慶應義塾大学工学部電気工学科卒業(1978), 同大学院修士課程修了(1980), 同博士課程修了(1983). 工学博士(1983). 南フロリダ大学コンピュータ学科助教授(1983~1985), 南カロライナ大学コンピュータ工学科助教授(1985~1988), ケースウエスタンリザーブ大学電気工学科准教授(1988~1996). tenured 受賞(1992). 慶應義塾大学環境情報学部助教授(1992~1997), 同大学教授(1997~現在). 研究: ニューラルコンピューティング, セキュリティ. NSF-RIA 賞(1989), IEEE Trans. on NN 功労賞(1992), IPSJ 論文賞(1980), TEPCO 賞(1993), KAST 賞(1993), 高柳賞(1995), KDD 賞(1997), NTT tele-education courseware 賞(1999). 政府顧問: NCC(フィリピン), VITTI(ベトナム), CTTISC(ジョルダン), タイ, スリランカ, マルチメディア大学(マレーシア). 22 冊の本と 200 編以上の論文.