

Pulse Coupled Neural Network の複数移動物体抽出への応用

茶志川孝和(*,**), 武藤 佳恭(***)

Pulse Coupled Neural Network は, 静止画に含まれるセグメントやエッジ抽出を学習やパラメータの調整なしに行なえる有効な手段であるが, 動画処理に適応した場合には PCNN が持つこのような性質を有効に活用することは困難であった. 本論文では, 外部刺激に対し応答が一過性で直ちに順応するような受容器を PCNN に適応する方法を提案した. 提案手法では, 小さなノイズやランダムに出現するノイズにネットワークが順応し, 移動物体のみを抽出することが出来る. 提案手法の有効性を示すため, 屋内外で撮影した動画画像に適応し, 複数の移動物体の抽出を行なった結果を示す.

Modification of PCNN for moving objects segmentation

Takakazu Chashikawa(*,**) and Yoshiyasu Takefuji(***)

The Pulse-Coupled Neural Networks (PCNN) is a powerful image-processing tool, because training or parameter adjustments are not required. The PCNN property doesn't work well when we use PCNN for image sequence processing. In this paper, we proposed a method to install a phasic receptor to the PCNN. The dynamics of the proposed PCNN yield to adaptation to noises which are small or random. The PCNN outputs present the segmentations of moving objects. We tested the proposed method in two types of image sequences, one is captured in the room and the other is captured in the outside parking. In both data the proposed method produced good moving object segmentations.

はじめに

コンピュータビジョン研究において, 動画からの移動物体の抽出は基本的かつ非常に困難な問題の一つである. これまで主にとられてきた, フレーム間差分, 背景画像との差分といった方法は, 照明環境の変動などのノイズが差分値として現れてしまうため, 実画像を入力とした場合, 移動物体とノイズに起因する変化領域を見分けることが困難であった. そのため多くの場合に, 利用する場所が制限されたり, 複雑な前処理や後処理が必要となっていた.

一方, Pulse Coupled Neural Network(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[1]を基に, J. L. 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[2]. PCNN はエッジやセグメントなどの特徴を画像から抽出する機能が先天的に備わったニューラルネットワークモデル[3]であり, 有効な画像処理エンジンとして応用されている[4]が, 動画を単純に入力できないため動画処理への応用は少なかった.

本論文では, 複数物体の自動抽出を目的として, 本来 PCNN に備わっているセグメント抽出機能に着目, 画素の変化に一過性に応答する性質をニューロンに導入することで, PCNN を動画入力に適応させる方法を提案する.

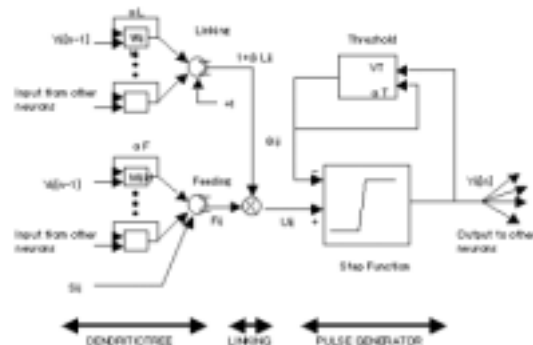


図 1 ニューロンモデル

提案手法では, 小さなノイズや不連続な動きをするノイズを含んだ動画においても, PCNN のダイナミクスにより自動的に抑制されるため, それらの影響を受けることなく複数移動物体の抽出が可能となる. 提案手法の有効性を示すため, 屋内外で撮影した動画からの複数移動物体抽出結果を示した.

Pulse Coupled Neural Network

本章では PCNN の基本モデルを説明する. ニューロンモデルを図 1 に示す. ニューロンは Dendritic Tree, Linking modulation, Pulse generator と呼ばれる三つの部分から構成される. Dendritic tree は, Linking modulation に対し, それぞれの独立した入力を生じさせる Feeding 部, Linking 部と呼ばれる二つの部分に分かれる. それらは, 近隣のニューロンとそれぞれシナプス結合 M , W を介して接続されている. なお, これらはそれぞれ受容野を持っているため, 二

(*) 慶義塾大学 SFC 研究所 (Keio Research Institute at SFC)

(**) ニッタン株式会社 (Nittan Co., LTD)

(***) 慶義塾大学環境情報学部 (Faculty of Environmental Information Keio University)

ニューロンは内部的に二つの受容野が存在する(図2)。

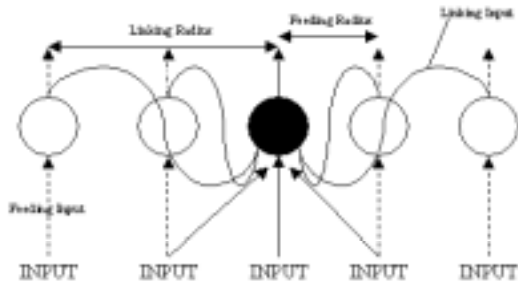


図2 ニューロンの相互作用

外部刺激 S は Feeding 部だけが受ける。それぞれの部分は前の状態を保持しており、その状態は時間とともに減衰項によって衰退してゆく。Feeding 部と Linking 部の値は次式よりそれぞれ求まる。

$$F_{ij}[n] = e^{-\alpha_F \delta_n} F_{ij}[n-1] + S_{ij} + V_F \sum_{kl} M_{ijkl} Y_{lk}[n-1] \quad (1)$$

$$L_{ij}[n] = e^{-\alpha_L \delta_n} L_{ij}[n-1] + V_L \sum_{kl} M_{ijkl} Y_{lk}[n-1] \quad (2)$$

ここで、 S は外部刺激、 Y はニューロンの前の状態を示す。 α は時定数であり、 δ は単位時間を表す。 V_F 及び V_L はシナプスの重み行列を正規化する為の定数である。

Linking Modulation は、Dendritic tree の二つの部分の状態からニューロンの内部状態を結合強度 β にしたがって結合し、ニューロンの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。

$$U_{ij} = F_{ij}[n](1 + \beta L_{ij}[n]) \quad (3)$$

ここで、 β は結合強度である。Pulse generator は、内部状態と動的閾値を比較してニューロンの出力を決定する。出力 Y は以下の式より導き出される。

$$Y_j[n] = \text{step}[U_{ij}[n] - \theta_{ij}[n-1]] \quad (4)$$

ここで、 Step はステップ関数である。ニューロンがパルスを発生すると、それは閾値にフィードバックされ、閾値の値はパルスの生成が停止する迄急激に増加する。その後、閾値の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。これにより、生体のニューロンが一度発火すると一定の期間発火しない現象(不応期)を実現している。このような閾値の変化の過程は時式によって表される。

$$\theta_{ij}[n] = e^{-\alpha_\theta \delta_n} \theta_{ij}[n-1] + V_\theta Y_{ij}[n] \quad (5)$$

ここで、 V_θ は正規化定数、 $e^{-\alpha_\theta \delta_n}$ は時定数である。ニューロンを構成するこれら三つの部分が下記の一

連の動作を行なうことで、ネットワークの一回の処理が完了する。

1. それぞれの受容野の入力過重和を取る。(Dendritic tree)
2. Linking 入力により Feeding 入力を調整する。(Linking)
3. 出力パルスを生成する。(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。

画像処理への適応

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る。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、入力画像の画素毎の輝度値を外部刺激 S_{ij} とする。PCNN は Linking 及び Feeding の作用により周りのニューロンからの入力を受けながら、そのダイナミクスを変化させ出力を決定する。それぞれのニューロンの出力を画素の輝度値とみなせば、ネットワークの動作一回につき一枚の画像を出力として得ることになる。

動画入力への適応

我々は PCNN を動画入力へ適応させるために、PCNN の基本モデルに対し、2 点改良を行なった。一つは、Feeding 部の外部刺激入力に関する改良。もう一つは閾値に関する改良である。本章では、これらの改良点に関して説明する。

PCNN の改良

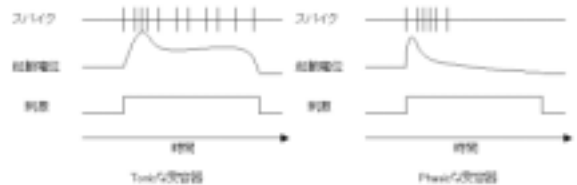


図3 Tonic な受容器と Phasic な受容器

基本モデルにおける外部刺激にはニューロンに対応する画素の輝度値が用いられている。PCNN に静止画を入力として与えて動作させた場合、外部刺激はネットワークが動作しているあいだ一定となる。そのため、PCNN のダイナミクスにより出力は入力画像に含まれる物体のセグメント エッジ エッジの拡大 セグメント . . . となる[3]。そこで、我々は PCNN の持つこのセグメント抽出機能に着目し移動物体の抽出を試みたが、入力を動画像に変えても移動物体の抽出は行なえなかった。

この問題を解決するためにニューロンの刺激の受け取り方法に注目した。基本モデルのニューロンは画素の輝度値を刺激として捕らえる持続性受容器を備えていると考えることが出来る。受容器とは、ニューロンの一部分をさし、刺激として加わるエネルギーを電気的变化に変える特別な細胞のことである。受容器の電氣的応答はほとんどの場合時間の経過とともに減弱す

る。この変化を順応と呼ぶが、その程度で受容器の性質を分類することが出来る。刺激が続く限り電氣的応答が続く受容器を持続性受容器、一方、刺激に対し応答が一過性で直ちに順応する受容器を相動性受容器と呼ぶ[5] (図3)。つまり、相動性受容器は入力の変化に反応すると考えることが出来る。ここで、移動物体の抽出を考えた場合、物体の移動は画像の輝度の変化として捉えることが出来るので、PCNN に刺激に対し応答が一過性で直ちに順応するような性質を導入すれば、PCNN をもちいて移動物体の抽出が行なえと考えた。

そこで、本論文では、PCNN に相動性をもたせる方法として以下の方法を提案する。基本モデルにおいて、外部刺激を受け取っている Feeding 部の式 1 に対し、刺激に対する一過性の応答 P_{ij} を導入、また、 R により複数の外部刺激を受けるように改良を行なった。

$$F_{ij}[n] = R_{jmn} P_{ij}[n] + e^{-\alpha_f \delta n} F_{ij}[n-1] + V_F \sum_{kl} M_{ijkl} Y_{kl}[n-1] \quad (1')$$

ここで、 P_{ij} は相動項、 R は外部刺激の結合重みである。なお、相動性を示す項を P_{ij} は次式で算出される

$$P_{ij}[n] = V_p |S_{ij}[n] - S_{ij}[n-1]| + e^{-\alpha_p} P_{ij}[n-1] \quad (6)$$

ここで、 $e^{-\alpha_p}$ は減衰項である。

PCNN の基本モデルにおける閾値(式5)はパルスの発生に伴って動的に変化し、これにより生体の不応期が実現されている。これにより PCNN はバースト発火現象のような生体ニューロン現象を示すことが出来るのだが、連続画像を入力とした場合には抽出物体の断片化を引き起こす原因となってしまうため、提案手法では閾値を定数にすることにより解決を図った。また、PCNN への入力は正規化した値を用いるものとする。

複数移動物体の抽出手順

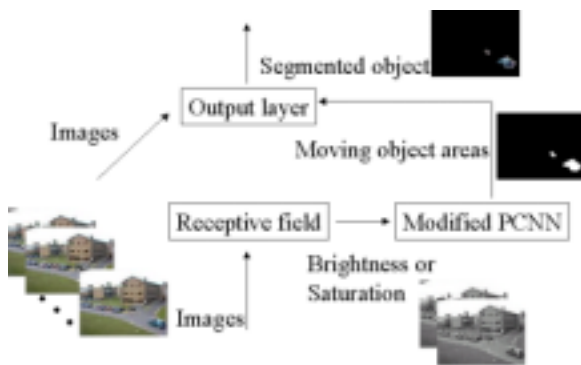


図 4 抽出手順

提案手法による複数移動物体の抽出手順を図4に示す。まずカラー動画からフレーム画像の輝度・彩度を256階調で計算する(人物の肌などを対象とした場合に、明度の情報を用いたほうが有効である場合がある。特殊な撮影装置が必要となるが、対象物によっては、赤外情報等を用いることも可能である)。輝度・彩

度から適切なものを選択し、改良型 PCNN の外部刺激として入力する。最後に PCNN の各ニューロンの出力(1or0)と対応するフレーム画像の画素の積をとる。この時、発火しているニューロン群にあたる部分が動物体領域として抽出される。

評価実験

本章では、提案手法の有効性を示すため、室内で両手を振っている人物を撮影した動画、および、屋外の駐車場で歩行者と通行車両を撮影した動画¹に適用した結果について述べる。

実験は、Pentium III processor 650 MHz を載した DOS-V 互換機上で MATLAB5.3 を用いて行なった。パラメータは全ての実験において、Gaussian kernel radius(M, W, R) = 3, (M は抑制結合) $\beta = 0.4$, $\theta = 0.5$, $\log 2 / \alpha_L = 0.2$, $\log 2 / \alpha_F = 0.1$, $\log 2 / \alpha_p = 0.01$, $V_L = 1.4$, $V_F = 0.1$, $V_p = 9.0$ とした。なお、室内で撮影した動画による実験では、外部刺激として彩度を選択し、屋外で撮影した動画による実験では、輝度を選択している。

室内で撮影した動画による実験

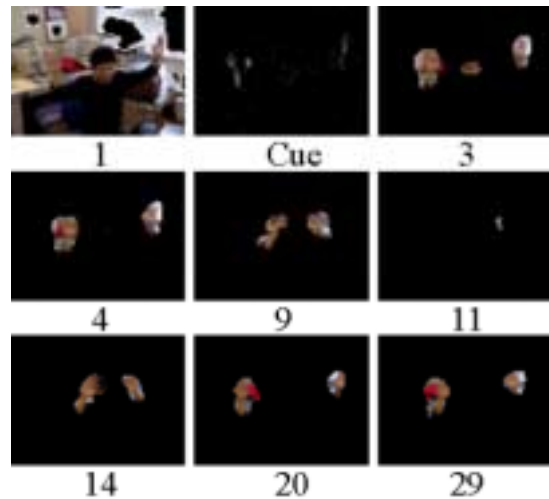


図 5 室内で撮影した画像の処理結果

室内の動画(80x60画素,全30フレーム)から、左右に振られている人物の手を抽出した結果の一部を図5示す。図5-1が第一フレームの抽出結果である。PCNN に対し最初に外部刺激を与えるとすべてのニューロンが反応するため、入力画像のそのものが出力されている。画面上の人物は挙げている両手を左右にお互い反対方向になるように振っている。図5-cueは、第1・第2フレームの彩度の変化値を示している。

PCNN はこの値に基づいた外部刺激を入力としているが、実際には動きのない部分でも、明るさの変化などがノイズとして現れており、また、移動物体は部分的に抽出されている程度である。図5-4は第4フレ

¹ The PETS2001 datasets (The University of Reading, UK)を用いた。

ームの抽出結果である。両手と頭の一部のみが抽出されている。

図 5-4, 図 5-9 は第 4 フレームから第 9 フレームの抽出結果を示している。この間両手は画面中心部に向かって移動しており、両手とも抽出できている。その後、第 10,11 フレームは両手の移動方向が変わる部分であり、抽出がうまく行えなかった(図 5-11)。第 14 から 20 フレームにおいて、両手は再び画面の両端方向へ移動しており、両手とも抽出されている(図 5-20)。そして、第 21 から 24 フレームにおいて、再び手の移動方向が変わり、抽出はできなかったが、第 25 フレーム以降第 29 フレームまで、手が移動している間は抽出できている(図 5-29)。

屋外で撮影した動画像による実験

屋外の動画像(データセット名: dataset1¥_test¥_camera1, フレーム番号:2490-2640[5 フレーム間隔],115x87 画素)から動方向に進む歩行者と車両を抽出した結果の一部を図 6 に示す。

左上図 6-1 が第一フレームの抽出結果である。PCNN に対し最初に外部刺激を与えるとすべてのニューロンが反応するため、入力画像そのものが出力されている。紙面に向かって左奥の歩行者は画面の左方向へ歩いており、T 字路の奥から車両 1, 2 が連なって歩行者の方向へ曲がって進行している。図 6-cue は、第 1・第 2 フレームの輝度の変化を示している。輝度の変化には実際には動きのない部分でも、明るさの変化などが多くノイズとして現れており、また、移動物体は部分的に抽出されているだけである。図 6-2 は第 2 フレームの抽出結果である。外部刺激のノイズ周辺の領域も抽出している。図 6-4 は第 4 フレームの抽出結果である、図 6-4 にみられたノイズの周りの領域は抽出されておらず、歩行者、車両 1,2 を含む領域のみが抽出されている。その後、図 6-29 に示す第 29 フレームまで、移動物体を抽出できている。なお、歩行者は第 17 と 18 フレーム間で画面の外に出ているのでその後は抽出されていない。

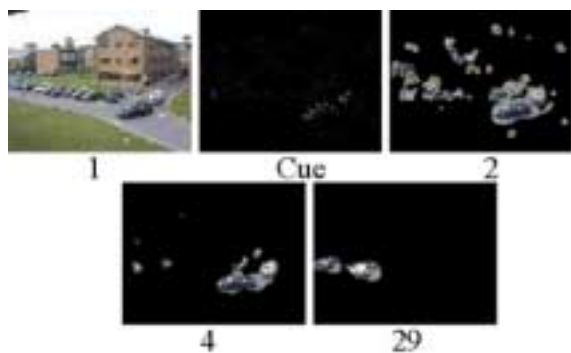


図 6 屋外で撮影した画像の処理結果

考察

以上、提案手法により、複数の移動物体を含む領域を正確に抽出できること実験により示したが、ここで、

その有効性や問題点について考察する。一般に実環境では、本実験からわかるように、移動物体の特徴が現れると仮定している差分情報に照明条件の変化などに起因する多くのノイズを含んでいる。提案手法は、このようなノイズに対して、Linking 部の入力による抑制や前のニューロン状態などが、PCNN のダイナミクスにうまく作用することで、ニューロンの反応自体が順応し抽出を行わなくなる。PCNN が順応するノイズは小さなものやランダムに出現するものである。この様子は各実験の初期の処理過程によく現れている。ノイズへの順応は PCNN ダイナミクスそのものによるため、今回の実験で示したように撮影場所が異なるデータを処理した場合でも、パラメータを調整する必要がなく、実環境で撮影したデータを処理する上で大変有効であると考えられる。ただし、提案手法による移動物体の抽出では対象物以外の領域を多少含んでいる。主な原因としては、影と対象物を区別できないことと、ネットワークが即座に順応できないことが考えられる。そのため、影と対象物進行方向の逆側に尾びれのような残像が抽出される傾向がある。

なお、今回室内で撮影した動画像の処理には彩度を用いているが、これは、背景と手の色が類似していたため輝度情報にはほとんどさが現れなかったためである。また、室内で撮影した動画像において、第 11 から 13 フレーム、第 21 から 24 フレームの間は手の抽出が行えなかった。これは、手の移動方向が変わる前後にあたるフレームであるため、移動速度の減少により十分な外部刺激が得られなかったと考える。

おわりに

本論文では、複数移動物体の抽出を目的として、PCNN のセグメント抽出着目し、動画像を入力とする PCNN を提案した。提案手法をパラメータの調整することなく屋内外の撮影条件の異なる動画像に適応し、複数移動物体の抽出に有効であることを示した。

参考文献

- [1]R.Eckhorn, H.J.Reitboeck, M.Arndt, and P.Dicke, 1990, Feature linking via synchronization among distributed assemblies: simulations of results from cat cortex, Neural Comput. 2, 293-307.
- [2]John L. Johnson, 1994, Pulse-coupled neural nets: translation, rotation, scale, distortion, and intensity signal invariance for images, APPLIED OPTICS, Cvol. 33, No. 26, 6239-6253
- [3]Thomas Lindblad and Jason M.Kinser, Image Processing using Puls-Coupled Neural Networks, Springer.
- [4]H. S. Ranganath and G. Kuntimad, 1999, Object Detection Using Pulse Coupled Neural Networks, IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 10, NO. 3.
- [5]Fred Delcomyn, Foundations of Neurobiology, W.H. Freeman and Company, New York. (1998)