

画像を入力とするニューラルネットの学習における 方位選択性入力への付加

大分大学 沢津橋由人 柴田克成

Addition of orientation-selective inputs in the learning using a neural network with image inputs

Yoshito Sawatsubashi, Katsunari Shibata, Oita University

Abstract :

In the authors' laboratory, aiming to flexible and purposive recognition, image signals are put into a neural network directly as inputs and the network is trained based on reinforcement learning. Since an experimental result suggested a limitation in the learning from raw image signals, the addition of orientation-selective inputs is examined in both supervised learning and reinforcement learning in this paper. Improvement of the learning performance and generalization ability by the addition could be observed for the images with a low resolution, but the difference cannot be clearly observed for the images with a higher resolution due to the performance improvement by the increase of resolution even without the addition.

1 まえがき

ロボットが、膨大な情報を得ることが可能な視覚センサを用いる場合、通常は、設計者がタスクに必要な情報を抽出し、行動計画等に用いる。しかし、人間の設計力の限界を考えると、ロボットが実世界で柔軟に状況を判断し、行動するためには、並列処理と学習が可能なニューラルネットを用いて自律的に学習することが有効であると考えられる。

そこで本研究では、人間による干渉を避け、視覚センサ信号を直接ニューラルネットに入力し、センサからモータまでを機能モジュールに分割することなく報酬や罰に基づいて強化学習するというアプローチで研究を行って来た¹⁾。文献²⁾では、カメラに映った矢印の方向に可動カメラを回転させると報酬が得られるタスクをリカレントネットで学習することで、矢印パターンの種類によらず、タスクに必要な向きを抽出して、記憶し、行動に反映できるようになることを確認した。

しかし、先行研究では、矢印パターンの種類が複数の場合、カメラの位置をずらすとうまく学習できなかった。このことから、生の画像信号から学習することに限界があると考え、汎用的かつ柔軟性を阻害しない入力を付加することで学習能力を向上させることを検討した。そこで、方位選択性入力、つまり、特定の方位のエッジに選択的に反応する入力を付加することを考えた。たとえば、矢印の向きを認識する場合、初期位置を水平にずらすと、特定の領域の画素値によってその向きを判断することは容易ではないが、画像の下方に横軸から反時計回りに45度の向きのエッジがあり、上方に135度方向のエッジがあれば右向きの矢印と容易に認識できると考えられる。さらに、エッジは画素間の差を取るため、照明条件の変化にも強いと考えられる。また、このようなことは、矢印以外の形状認識や文字認識でも有効であると考えられる上、実際に、生物の一次視覚野には、方位選択性セルがあることは、良く知られていることである。

そこで本研究では、画像をニューラルネットに入力して学習する際に、方位選択性入力を付加する。そして、先行研究を参考に、矢印の向きを認識に必要な教師あり学習と強化学習のそれぞれにおいてその効果を検証した。

2 矢印方向判別の教師あり学習

2.1 学習方法とタスク設定

まず始めに、方位選択性入力の付加が対象物の認識において有効に働くのか、矢印方向判別の教師あり学習をすることで検証を行った。

矢印方向判別タスクでは、Fig.1のように、カメラで実際にキャプチャしておいた矢印画像を入力として階層型のニューラルネットに与え、矢印の方向が右か左かを出力するタスクである。

学習に用いた矢印として Fig.2 の4種類用意し、さらに、キャプチャする位置も左右に5°まで0.1°刻みでずらし左右の向きの違いも含めて全部で808パターンから毎回ランダムに選んで入力した。

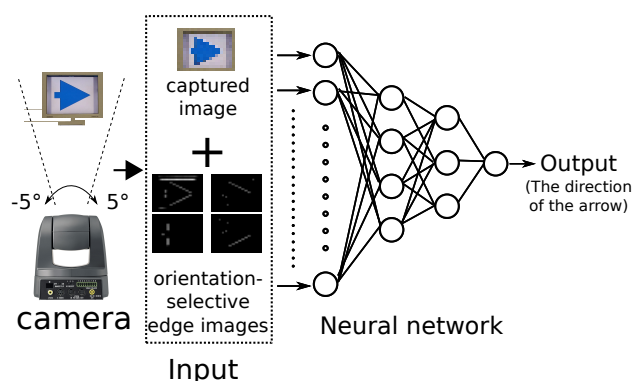


Fig. 1: Identification learning of arrow direction by a neural network with image inputs with the addition of orientation-selective inputs.

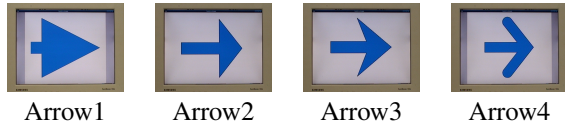


Fig. 2: Arrow patterns used for learning.

画像入力、生画像の各画素のRGBそれぞれについて0~255の値を-0.5~0.5に変換して与えた。一方、方位選択性入力を付加する場合は、まずグレースケールに変換した画像をFig.3のような角度の異なるエッジに反応する4つの3×3のフィルタを適用し、絶対値を取ることで、反転したエッジも同時に検出させ、それを0~255に線形変換した。さらに、特定の傾きのエッジにより選択的に反応するように変更後の画素値の85を下限とし、矢印の青と隣接する背景の白の色によってできるエッジの値である170を画素値の上限として85~170を0~1に線形変換し、入力に与えた。

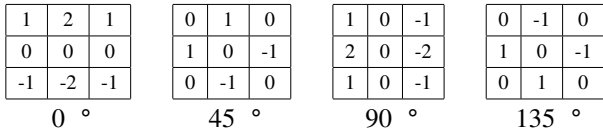


Fig. 3: 4 filters to generate orientation-selective inputs.

ニューラルネットは入力数が多いことから5層とした。入力数は、生画像のみの場合は画素数を40×30=1200とし、RGBのため3倍した3600個の信号を入力とした。方位選択性入力を付加する場合は、入力数を揃えるために生画像とエッジ画像共に画素数を24×18=432に落としてRGBと4方向のエッジで3024とした。中間層から出力層の各層のニューロン数は、入力層側から順番に500-100-20-1と設定した。ニューラルネットの初期重み値は、最下位の中間層から順に±0.1, ±0.2, ±0.5の範囲でランダムに与え、出力層は0とした。バイアス用の定常値入力は0.1とし、学習係数は0.5とした。中間層と出力層のニューロンは、-0.5~0.5の値域のシグモイド関数を出力関数とし、教師信号は、矢印が右方向のとき0.4、左方向のとき-0.4を与えBP(Error Back Propagation)法により学習させた。出力と教師信号の差の絶対値が0.1以下の条件を全画像パターン連続で満たした場合に学習を終了した。

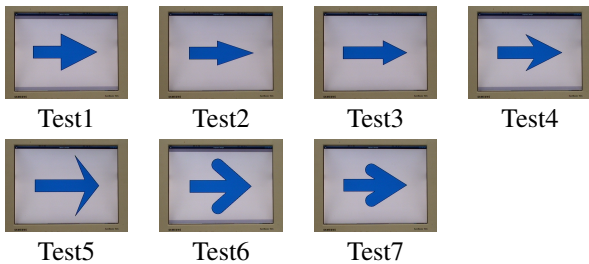


Fig. 4: Arrow patterns used for test.

学習終了後は、方位選択性入力の付加の有無それぞれのニューラルネットを用い、テストパターンとしてFig.4の矢印をキャプチャ位置を1°間隔で左右5°までの画像を入力として汎化能力の検証を行った。

2.2 学習結果

Table1に、方位選択性入力の付加の有無による学習終了までの試行回数を、乱数系列を変えた50回分平均したものと、学習後の汎化能力の検証としてテストパターンの平均二乗誤差の比較を示す。Table1をみると、方位選択性入力を付加した方が、画像の解像度が大きく低下しているにもかかわらず平均試行回数が1000試行程度少ないことがわかる。また、テストパターンの結果は、完全に認識できていないわけではないが、方位選択性入力を付加した場合のほうが誤差が減っていることがわかり、汎化能力の向上も見られた。

Table 1: Comparison of average number of steps until successful termination, and average error for the test patterns.

	方位選択性なし	方位選択性あり
試行回数 (標準偏差)	6414(763)	5306(682)
出力誤差 (標準偏差)	0.038(0.010)	0.017(0.007)

ここで、方位選択性入力の付加が有効に機能しているのかを調べるために、最下位中間層ニューロンの重み値を

$$pixel_{color,i,j} = \frac{w_{after,color,i,j} - w_{before,color,i,j}}{\max_{color,i,j} |w_{after,color,i,j} - w_{before,color,i,j}|} \times 127 + 128 \quad (1)$$

の式にしたがって画像化した。pixelは重み値画像の画素値、 w_{before} , w_{after} は学習前と後の重み値である。colorはR,G,B,グレースケールのどれかの色を、(i,j)は当該画素の画像中での位置を表す。また、初期重み値の影響を取り除き、学習による重み値の変化を見るため、学習後の重み値から学習前の重み値を引いている。

Fig.5に最下位の中間層の1つのニューロンについて画像化した重み値の例を示す。

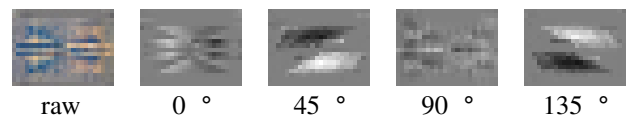


Fig. 5: Visualization of weight change during learning in one hidden neuron.

Fig.5の生画像に対する重み値のカラー画像は、左右の矢印が合成したようになっており、矢印の部分的な領域に反応するようになってきている。Fig.5の45°,135°のエッジに対する画像では、画像の上下に矢の斜めのエッジが重なってできたような黒と白の領域がみえる。まえがきでも述べたように、この中間層ニューロンが入力画像の

上部と下部に 45° または 135° のエッジを検出することで矢印の方向を判別していると考えられる。このような表現は他の中間層ニューロンにも多く見られることから、方位選択性入力に矢印方向の判断に貢献していると考えられる。

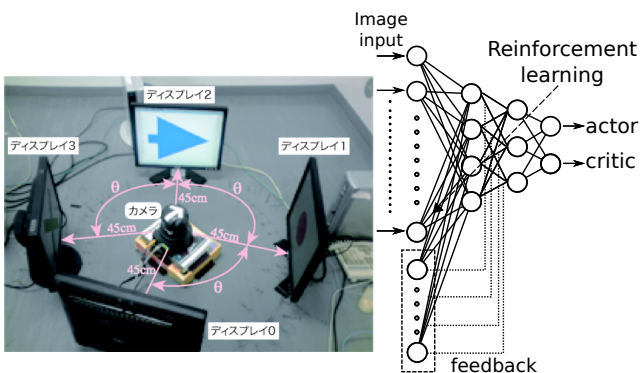
3 矢印方向に基づくカメラ移動の強化学習

3.1 学習方法とタスク設定

次に先行研究²⁾で行った実機を用いた強化学習において、方位選択性入力の付加の有効性を検証した。

カメラ移動タスクは、まず Fig.6 のように可動カメラと 4 台のディスプレイを用意し、ディスプレイ 1 または 2 に矢印、矢印方向の隣のディスプレイに赤い \times 、反対方向の隣のディスプレイに緑の \times を表示させる。そして、カメラが矢印をみて移動し \times を捉えたら報酬、 \times なら罰を与えて学習することで、カメラの画像からの矢印の認識と矢印方向へのカメラの移動を獲得させる。矢印は、前タスクと同様に Fig.2 の中から毎回ランダムに選択し表示させた。カメラの初期位置は、表示矢印を中心に左右 5° まで 1° 間隔でランダムに設定した。また Fig.6 のようにディスプレイを配置することで、ディスプレイ 1 が左矢印のときとディスプレイ 2 が右矢印のときに、途中同じ背景が画像入力とされるため、背景からは正しいカメラの回転方向を判別することができず、矢印の向きを記憶する必要があるタスクとした。

また、学習はレベル 1~7 に分けレベル 1 では Fig.6 の角度を 52° に設定した。1 試行で 40 ステップ経過したり、罰を 1 試行で合計 10 回与えられた場合は中断し、50 試行連続で (レベル+15) ステップ以内に報酬を得た場合に次のレベルに移行した。レベルが 1 上がるにつれて 8° づつ増やし、レベル 7 で 100° になるまで徐々に難易度をあげた。



Recurrent Neural Network

Fig. 6: Reinforcement learning task in which the system gets a reward when the system moves its camera to the arrow-pointing direction and finds the red circle²⁾.

このタスクでは、Elman 型のリカレントニューラルネットを用いた強化学習のアルゴリズムで学習を行い、記憶に基づく行動を自律的に獲得させる。強化学習として Actor-Critic を使い、強化学習に基づいて自動生成された教師信号でニューラルネットを学習した。

学習としては、矢印判別タスクと同様に画像の RGB 値と方位選択性入力をリカレントネットに与え、このときの時刻を t とする。出力は状態評価部の Critic と動作生成部の Actor があり、カメラは Actor の出力と試行錯誤量である乱数 rnd を加えた値に基づいて移動する。試行錯誤量 rnd は $-0.5 \sim 0.5$ の範囲とし、学習中に報酬が与えられると範囲を小さくし、報酬が与えられないと大きくする設定とした。また Actor と試行錯誤量 rnd の合計が $-0.5 \sim 0.5$ の範囲を越える場合は ± 0.5 で切った。1 ステップのカメラの移動量は出力の値域の $-0.5 \sim 0.5$ を $-9.6 \sim 9.6$ に線形変換した値となる。移動後の時刻 $t+1$ における出力により、以下の式のように Critic の教師信号 $C_{d,t}$ と Actor の教師信号 $A_{d,t}$ を強化学習に基づいて生成する。

$$TDerror_t = r_{t+1} + \gamma C(x_{t+1}) - C(x_t) \quad (2)$$

$$C_{d,t} = C(x_t) + TDerror_t = r_{t+1} + \gamma C(x_{t+1}) \quad (3)$$

$$A_{d,t} = A(x_t) + rnd \times TDerror_t \quad (4)$$

ここで、 γ は割引率、 r は報酬、 x が状態で $C(x_t)$ は状態評価値、 $A(x_t)$ は動作信号である。報酬が与えられた場合の Critic は 0.9 とするが、出力の値域が $-0.5 \sim 0.5$ のため Critic の教師信号 $C_{d,t-1}$ には 0.9 から 0.5 を引いた 0.4 を与えて学習を行い、状態評価値として Critic を計算する場合は出力に 0.5 を加える。また罰を与える場合は、罰がない時の教師信号に -0.1 を加える。こうして求められた教師信号により、時刻 t における入力に対し BPTT (Back Propagation Through Time) 法を用いて学習を行う。

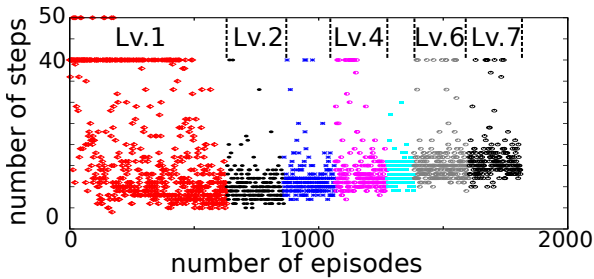
リカレントニューラルネットワークは 5 層に設定した。入力数は、生画像のみの場合は画素数を $44 \times 33 = 1452$ で RGB のために 4356 とし、方位選択性入力を付加する場合は、入力数を揃えるために生画像の画素数は 32×24 に Fig.3 のフィルタを使用したエッジ画像は 24×18 に設定し、合計 4032 とした。中間層から出力層のニューロン数は、入力層側から順番に 750-150-30-2 と設定した。

リカレントネットの初期重み値は、最下位の中間層から順に ± 0.1 , ± 0.2 , ± 0.5 の範囲でランダムに与え、出力層は 0 とした。フィードバック部分の重み値に関しては、セルフフィードバック部は 4、それ以外は 0 とした。バイアス用の定常値入力は 0.1 とし、学習係数はフィードバック部を 0.125 で他は 0.5 とし、BPTT 法による学習でさかのぼる時刻は最大 25 とした。

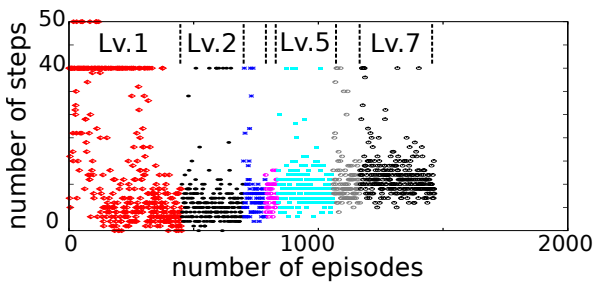
3.2 学習結果

Fig.7 に生画像のみの入力と方位選択性入力を付加したときの、試行回数に対する各試行でのステップ数を表す

学習曲線を示す。各試行では、40ステップでゴールできない時にはそこで打ちきっているが、罰を10回与えられて試行を終了したときは、50stepのところプロットした。また汎化能力の検証ために、それぞれの学習後のリカレントネットを用いて Fig.4 のテストパターンによる矢印方向に基づくカメラ移動タスクを行った。カメラの初期位置・表示ディスプレイ・テストパターンの組み合わせで計 308 パターンを検証し、Fig.8 にはそのときにカメラが 40 ステップ以内に移動して矢印方向の を捉えられなかった全パターンを示す。



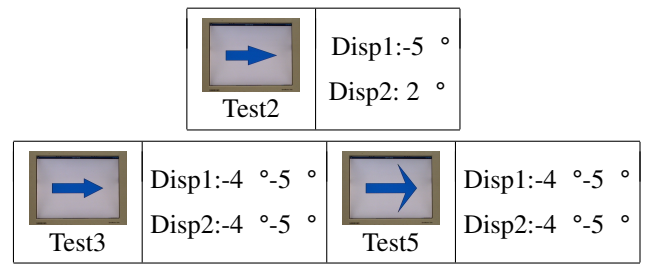
(a) without orientation-selective inputs



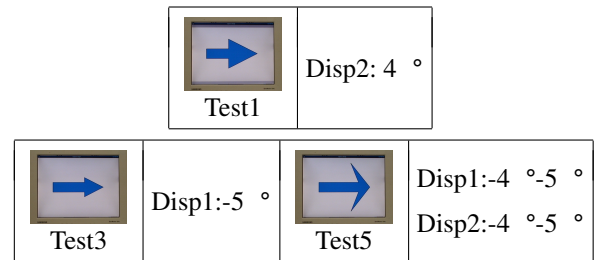
(b) with orientation-selective inputs

Fig. 7: Comparison of learning curve depending on the addition of orientation selective inputs.

Fig.7 の結果をみると、先行研究²⁾ではできなかった生画像のみの入力での学習ができたことがわかる。これは、先行研究²⁾と違い入力画像の画素数を上げたためと考えられる。しかしながら、方位選択性入力を付加した場合の方が特にレベル1の場合に学習は速かった。Fig.8 を見ると、テストパターンで認識できなかったのは生画像のみの入力の場合が 10 パターン、方位選択性入力を付加した場合が 6 パターンであった。Test5 のパターンは矢印方向にずれると学習中に矢が出なかった部分に矢が出るために汎化が元々利かなく考えられる。そのためこのパターンを除くと、失敗パターン数は 2 と 6 となり少し差が見られる。また、失敗したパターンは Fig.8 のようにいずれも右矢印の場合であった。これは、先に左矢印が学習し、左矢印に比べ右矢印の学習が足りなかったためだと考えられる。以上の結果は、1 回のみ結果であるため、信頼性を上げるためにはさらに学習回数が必要である。



(a) without orientation-selective inputs



(b) with orientation-selective inputs

Fig. 8: All the failed patterns in 308 test patterns. For each pattern, arrow image, presented display, and shift angle are shown.

4 まとめ

本研究では、画像をニューラルネットに入力して学習する際に、矢印の向き認識が必要な教師あり学習と強化学習において方位選択性入力を付加することによる効果の検証をした。

結果、方位選択性入力の付加は、教師あり学習では学習能力の向上がみられた。また Fig.5 の $45^\circ, 135^\circ$ のようにエッジ画像に繋がるニューロンには矢印方向に反応する表現が確認できた。次に強化学習では教師あり学習と違い入力画素数を上げて学習を行ったため、生画像のみでも初期位置のずれに対応できあまり大きな差は見られなかった。しかし画素数が小さい場合生画像のみの入力と同等、それ以上に近い結果から方位選択性入力の付加は認識において有効に働いていると考えられる。

このことから方位選択性入力の付加は、解像度を上げて効果は小さくなったが、動きの自由度が大きくなると、エッジの情報が効果をあげる可能性が大きいのではないかと考えている。

謝辞

本研究は、日本学術振興会科学技術研究費補助金基盤研究 (B)#19300070 および (C)#23500245 の補助を受けた。

参考文献

- 1) 柴田克成:強化学習とニューラルネットによる知能創発, 計測と制御, Vol. 48, No. 1, pp. 106-111, 2009
- 2) 柴田克成, 沢津橋由人, 宇都宮浩樹:強化学習によるパターンの意味付けと記憶に基づく行動の獲得, SICE九州支部学術講演会, pp. 243-246, 2010