

速度分散最小化変換による マルチモーダル情報からの状態表現の獲得

Acquiring state representation from multimodal information by slow feature analysis

- 西川 輝彦 (阪大) 港 隆史 (ATR)
荻野 正樹 (阪大) 正 浅田 稔 (阪大, JST ERATO)

Akihiko NISHIKAWA, Osaka University, 2-1, Yamadaoka, Suita, Osaka
Takashi MINATO, ATR, 2-2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto
Masaki OGINO, Osaka University, 2-1, Yamadaoka, Suita, Osaka
Minoru ASADA, Osaka University, JST ERATO Asada Project

This paper proposes a hierarchical model which is composed of a slow feature analysis (SFA) network to extract multi-modal representation of a humanoid robot. The experiment with humanoid robot shows that the network can integrate multi-modal information and detect semantic features by the extraction of the slowly varying features from the high-dimensional input sensory signal, and it shows that the multi-modal representation is useful as state representation for reinforcement learning compared with using state representation without the integration of the multi-modal information.

Key Words: Slow Feature Analysis, Multimodal representation, Reward prediction

1. はじめに

近年、人間のように多くのセンサを配置した自律型ヒューマノイドロボットの開発が盛んに行われている。ロボットは外界から得られるセンサ情報から必要な情報を抽出し、それを利用してある状態に対する行動を決定する。従来はあらかじめ設計された計算式に基づいて状態空間を構成していたが、ロボットが複雑な環境の中で適応的に行動するためには状況に応じて柔軟に状態空間を構成する必要がある。このため人間が行っているようにセンサ情報から自律的に必要な情報を抽出することが求められる。

人間は周囲の空間情報や自身と環境との相互作用を把握し、適切に運動を制御することで様々なタスクを行うことができる。これらの能力は自身の運動を行う際に知覚される様々な感覚情報を、大脳皮質連合野において時間的、空間的に統合することで獲得されると考えられる。このような情報の処理は背側皮質視覚路によって行われ、特に頭頂葉での処理が大きく関わっていると考えられている。この領域では多感覚性を示すニューロンの存在が確認されていることから、視覚と体性感覚が統合されると考えられている [1]。以上のことからマルチモーダルな情報の統合によって局所的な感覚情報から自己の状態や周囲の空間情報を表現する大域的な情報に変換されると考えられるが、このようなマルチモーダルな表現がどのようにして獲得されるのかは明らかにされていない。

認知発達ロボティクス [2] の分野では、ロボットにセンサ情報から自律的にマルチモーダルな表現を獲得させることが試みられてきた [3] [4]。これらの研究では異なるセンサモダリティの同期性に基づいた対応学習により視覚表現や視空間表現を確立しているが、静的な関係しか扱っていないため学習の契機は設計者が与える必要があった。しかし、入力される感覚情報には身体や環境のダイナミクスが反映されるため、時間遷移を考慮することにより入力信号そのものが持つ情報として、学習

の契機を与えることなく自律的に必要な情報が抽出できると考えられる。

入力信号の統計的性質から内在する情報を抽出する学習アルゴリズムに、Slow Feature Analysis (SFA) [5] がある。SFA は信号の速度分散を最小化する変換を行うことで、高速に変化する多次元の入力信号から低速に変化する特徴量を抽出することができる。例えば、ある線分より物体の形、姿勢より運動といった情報は低速に変化することから、SFA を用いることにより高次の情報が表現されると考えられる。SFA は、下位層のニューロンは特定の空間だけに反応するため高速に興奮状態を変えるが、上位層では興奮状態をあまり変えずに時間的、空間的に広い情報を表現しているという神経系の階層構造に基づいている。

本研究では、自己の身体や外界に関するセンサ情報を効果的に統合することによって自己の状態が適切に学習できるという考えのもとに、SFA を用いてロボットがセンサ情報から自律的にマルチモーダルな表現を獲得可能なモデルを提案する。これを強化学習と組み合わせることによって、獲得された表現が適切に報酬を予測できる状態表現となっていることを示す。

2. SFA によるマルチモーダル情報統合モデル

本研究では感覚情報として視覚情報、関節角度情報、触覚情報、姿勢情報を扱い、これらを統合してマルチモーダルな特徴抽出を行うことを目的として SFA の階層ネットワークを構成する。その処理過程を Fig.1 に示す。下位層を感覚入力層、上位層を統合層と名付ける。まず、感覚入力層に入力される感覚情報を時系列データとして記録し、モダリティごとに SFA を行う。この処理によって、各モダリティにおいて特徴抽出がなされる。次にそれらの出力を並べ、上位層に入力して SFA を行う。この処理はマルチモーダルな情報の統合に相当し、上位層における SFA の出力として各モダリティの情報が統合された

結果, さらに大域的な情報が抽出されると考えられる. 以上の処理によりネットワークが構造化され, 次に入力される感覚情報は学習されたネットワークによって変換される.

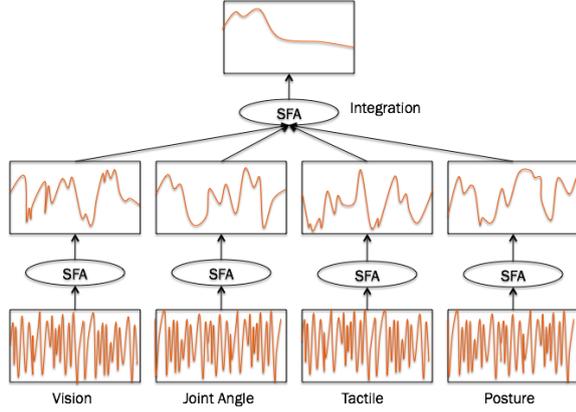


Fig.1 SFAによるマルチモーダル情報の統合過程

2.1 Slow Feature Analysis [5]

Slow Feature Analysis (SFA) は, 多次元の時系列入力信号から低速に変化する特徴量を抽出する教師なし学習アルゴリズムである. 入力信号 $\mathbf{x}(t) = [x_1(t), \dots, x_N(t)]^T$ に対して出力信号 $y_j(t)$ を低速にする変換 g_j を求める問題を解く. 結合荷重を \mathbf{w}_j^T とすると, 出力信号は

$$y_j(t) = g_j(\mathbf{x}(t)) = \mathbf{w}_j^T \mathbf{x} \quad (1)$$

と表される. 目的関数は, 次式で定義される.

$$\Delta(y_j) := \langle y_j^2 \rangle_t \quad \text{is minimal} \quad (2)$$

$\langle \cdot \rangle$ は平均, \dot{y} は y の微分を表す. これを, 以下の制約条件のもとで解く.

$$\langle y_j \rangle_t = 0 \quad (\text{zero mean}) \quad (3)$$

$$\langle y_j^2 \rangle_t = 1 \quad (\text{unit variance}) \quad (4)$$

$$\forall i < j, \langle y_i y_j \rangle_t = 0 \quad (\text{decorrelation}) \quad (5)$$

式 (2) は信号の微分の二乗に等しく, 信号の時間変化の尺度となる. Δ は高速に変化する信号に対しては大きくなり, 定常信号に対しては $\Delta = 0$ である. またこの式によって 1 番目の出力信号から最も低速な信号の順に出力される. 式 (2) で表される最適化問題を解くことは, 信号の共分散行列 \mathbf{A} と信号の微分の共分散行列 \mathbf{B} を計算し, 線形代数で知られる一般化固有値問題を解くことに帰着される.

$$\begin{aligned} \mathbf{A}\mathbf{W} &= \mathbf{B}\mathbf{W}\mathbf{\Lambda} \\ \text{with } \mathbf{A} &:= \langle \dot{\mathbf{x}}\dot{\mathbf{x}}^T \rangle_t \\ \text{and } \mathbf{B} &:= \langle \mathbf{x}\mathbf{x}^T \rangle_t \end{aligned} \quad (6)$$

$\mathbf{W} = (w_1, \dots, w_N)$ は固有ベクトルを並べた行列, $\mathbf{\Lambda}$ は固有値 $\lambda_1, \dots, \lambda_N$ を対角成分とする行列である.

3. 特徴抽出に基づく状態価値学習モデル

前節のモデルにより獲得された状態表現の有用性を検証することを目的として, 特徴抽出に基づく状態価値学習モデルを提案する. 入力される感覚情報の処理を行う SFA の階層ネットワークと, その出力を状態として状態価値と結びつけるニューラルネットワークから構成される. モデルの全体図を Fig.2 に示す. まず, 感覚経験を通して前章で述べた SFA ネットワークを構造化しておく. そのネットワークを用いて入力されるマルチモーダルな情報から自己の状態を表現する特徴抽出を行い, 強化学習を行うことにより状態価値関数を学習する. その後は学習した価値関数に基づき, 感覚情報から現在の状態価値を推定することが可能である.

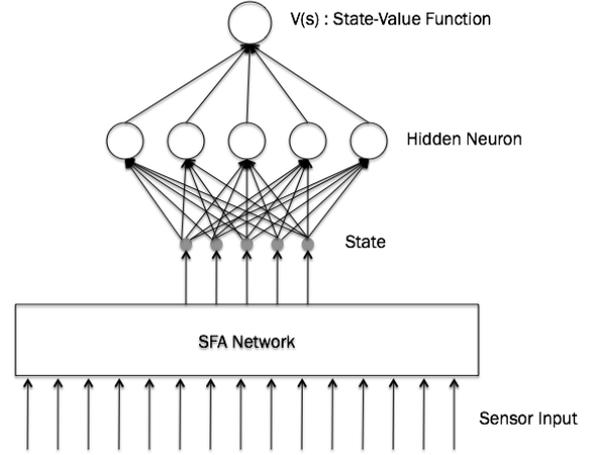


Fig.2 状態価値学習モデル

3.1 状態価値

ある状態 $s(t)$ を起点として方策 π に従って報酬 $r(t)$ を受け取る時の状態価値 $V(s(t))$ は,

$$\begin{aligned} V(s(t)) &= E[r(t) + \gamma r(t+1) + \gamma^2 r(t+2) + \dots] \\ &= E[r(t)] + \gamma V(s(t+1)) \end{aligned} \quad (7)$$

と定義される. γ は割引率と呼ばれ, $0 \leq \gamma \leq 1$ である. 状態価値 $V(s(t))$ は学習率 $\alpha (0 \leq \alpha \leq 1)$ を用いて以下のように更新される.

$$V(s(t)) \leftarrow V(s(t)) + \alpha \Delta V(s(t)) \quad (8)$$

$$\Delta V(s(t)) = r(t) + \gamma V(s(t+1)) - V(s(t)) \quad (9)$$

このとき求められる状態価値の推定誤差 $\Delta V(s(t))$ を TD 誤差と呼ぶ. 方策に従って行動をとることで状態が変化し, 報酬が与えられることで状態価値が更新される.

本研究では連続かつ多次元の状態空間を扱うため状態価値 $V(\mathbf{s})$ を関数で近似する. SFA ネットワークの出力 $y_i (i = 1, 2, \dots)$ をそれぞれ状態変数 $s_i (i = 1, 2, \dots)$ とする. 状態ベクトル $\mathbf{s} = (s_1, s_2, \dots)$ を入力, 状態価値関数 $V(\mathbf{s})$ を学習データとし, 中間層を 1 層有する 3 層のニューラルネットワークで近似器を構成する. 状態価値関数は次式で表される.

$$V(\mathbf{s}) = f^{(2)} \left(\sum_{j=0}^N w_{kj}^{(2)} f^{(1)} \left(\sum_{i=0}^M w_{ji}^{(1)} s_i \right) \right) \quad (10)$$

$f^{(1)}, f^{(2)}$ はシグモイド関数である。実際には出力が1つのノードしか持たないが便宜上添字 k を記述した。各層間の重み w_{kj} , w_{ji} は Back Propagation(BP) 法を用いて修正する。

3.2 学習方法

ロボットは目標物に向かって運動し、目標物に到着した時刻 t で報酬 $r(t)$ が与えられる。同時に、現時刻までの全ての状態 $s(0), s(1), \dots, s(t)$ をニューラルネットワークに入力し、状態価値の学習を行う。この一連の過程を1エピソードとし、状態価値関数の更新は1エピソードにつき1回のみ行うものとする。学習を繰り返し行うことで、自己の状態と状態価値を適切に結びつけることが可能になると考えられる。

3.3 状態価値の推定

学習後のSFAネットワーク、状態価値関数を用いてある時点における自己の状態から状態価値を推定する。入力される感覚情報は学習時と同様の順序を辿ることにより階層的に処理される。まず、センサ情報を各モダリティにおけるSFAに入力し、学習済みの結合荷重により変換する。出力として抽出される情報は各モダリティの状態を表現している。それらを統合して次の層に入力することにより自己の状態表現に相当する出力が得られる。最後に自己の状態表現から価値関数を用いて状態価値が算出される。推定される状態価値が高い程、報酬を期待できる可能性が高い状態にあるといえる。

4. ヒューノイドロボットによる実験

4.1 実験設定

ロボットは運動を行い、その間1ステップごとにそれぞれのセンサ情報を観測する。記録した時系列データから、提案モデルにより特徴抽出と強化学習を行う。実験プラットフォームにはM3-Neony [6]を用いて、はいはい運動による目標物到達タスクを扱う。M3-Neonyの正面像をFig.3、実験の様子をFig.4に示す。センサ情報として、視覚情報にはオプティカルフローを用いる。320×240ピクセルの単眼の画像を7×7の領域に分割し、各領域においてブロックマッチング法により2次元のオプティカルフローを算出する。SFAにはそれらを並べた98次元のベクトルを入力する。関節角度情報には全身22個のモータのそれぞれの関節角度の指令値を用いる。触覚情報には全身90個の触覚センサの値を用いる。姿勢情報として2軸ジャイロセンサと3軸加速度センサを用いる。



Fig.3 M3-Neony

4.2 特徴量の抽出

ロボットは最初うつ伏せに倒れた状態から手足を使って起き上がり、はいはいをしながら前進するという運動を行い、開始

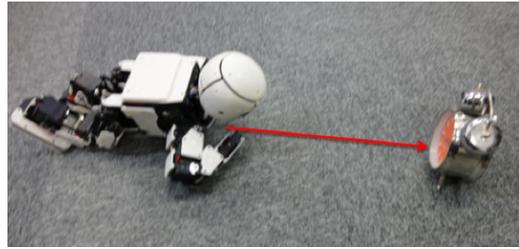
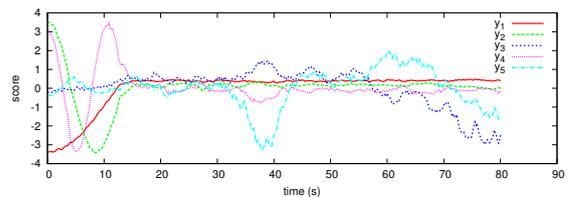


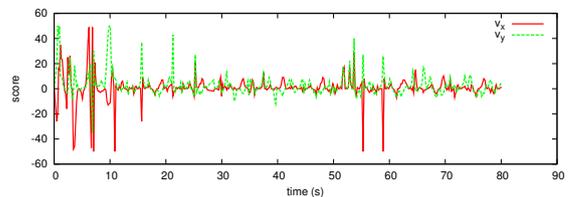
Fig.4 実験の様子

後401ステップで運動、センサ情報取得を停止した。この時点まで80.087sを要した。この一連の運動を行った間のセンサ情報から、提案モデルにより特徴抽出を行った。統合層の出力結果において最も低速に変化する(固有値が小さい)5成分をFig.5(a)に示す。生のセンサデータの例として1領域のオプティカルフローの x, y 成分 (Fig.5(b)) のグラフと比較すると、SFAによって短期的な感覚情報の変化にあまり影響されない特徴量が抽出できたといえる。

出力 y_1 は起き上がり運動時のみ負の値を取り、それ以降(15s以降)のはいはい運動時は常に0.4付近の値を取ったことから、この出力から運動状態の識別を行うことが可能である。同様に y_2 と y_4 の値も運動状態の変化に応じて変化し、はいはい運動時は一定の値を取ることが確認できる。これに対して、出力 y_3 や y_5 は運動状態が変化しても値はあまり変化せず、はいはい運動を行っている間に値の変化が見られた。特に y_3 は目標物に近づくにつれて負の値が大きくなった。これらの結果から、統合層の出力は運動状態や空間情報を表現している可能性があると考えられる。



(a) 統合層の出力



(b) オプティカルフローの x, y 成分

Fig.5 マルチモーダル情報の統合結果

4.3 状態価値の学習

提案モデルを使ってロボットが状態価値を学習できるかを検証した。ロボットは前小節と同様に目標物に向かってはいはいをしながら前進し、開始後400ステップで報酬 $r(t) = 1$ を与えられ、運動、センサ情報取得を停止した。その他の場合は

報酬は常に 0 とした. SFA ネットワークの出力ノードのうち 20 個をこのニューラルネットワークの入力層のノードとして用い, 学習データを式 (8) で与えられる状態値 $V(s)$ として学習を行った.

50 エピソード終了後の状態値を Fig.6 に示す. 赤色の線が式 (8) により与えられる理論値であり, 緑色が実際に学習した状態値を表す. 2 つのグラフがほぼ一致していることから, 状態値は適切に学習できたといえる.

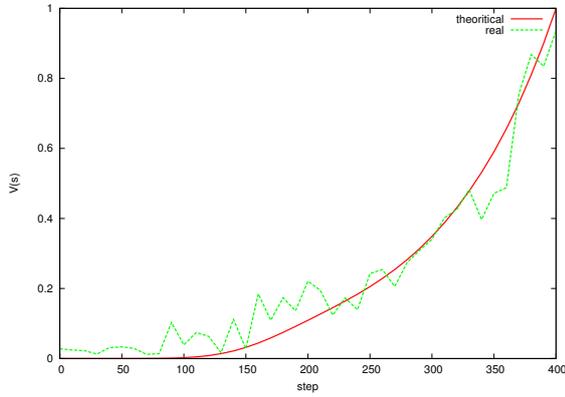


Fig.6 状態値の学習結果

4.4 状態値の推定

学習した状態値関数を用いて, 新規エピソードにおいて状態値を推定できるかについて検証する. 学習時から少しずれた経路を目標物に向かって前進し, 開始後 400 ステップで運動, センサ情報取得を停止した. この時点までの 10 ステップごとの推定状態値 $V^*(s)$ の推移を Fig.7 に示す.

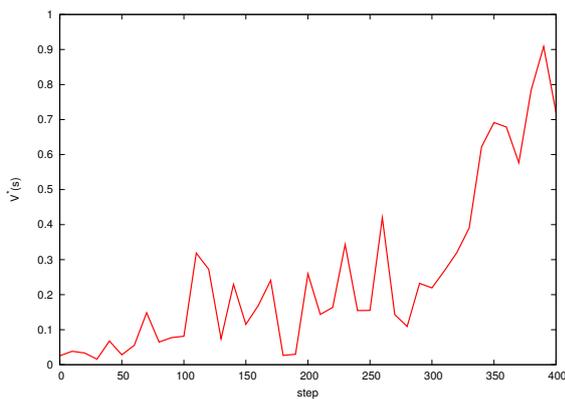


Fig.7 新規エピソードにおける推定状態値 (統合あり)

グラフは報酬の与えられる開始後 400 ステップに向かって概ね右肩上がりになっていることから, 学習した経路から少しずれた経路においても学習した価値関数を使って状態値を推定できていることが確認できる.

次に, マルチモーダル情報の統合の効果を調べるために, 統合層が存在しない, すなわちマルチモーダル情報の統合を行わずに状態値の学習を行った結果との比較を行う. 視覚, 関節角度, 触覚, 姿勢の情報に対してモダリティごとに SFA を行ったものをそのままニューラルネットワークに入力した. 学習率等の条件は同じ値に設定した. 新規エピソードにおける推定結

果を Fig.8 に示す. 1 ステップごとの値の変動が大きく, 正しく推定が行われなかった. このことから, 身体や環境の変化に対してロバストな状態表現の獲得のためにはマルチモーダルな情報の統合と低速に変化する情報の抽出が重要になると考えられる.

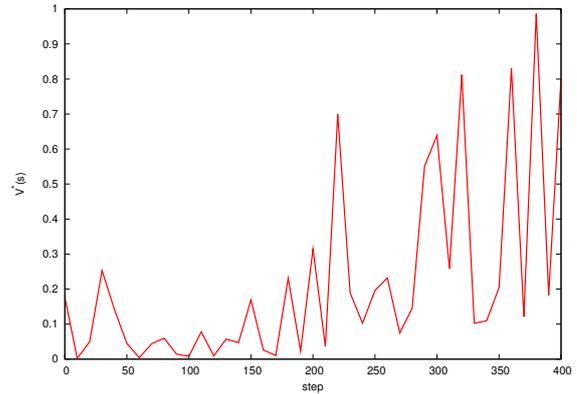


Fig.8 新規エピソードにおける推定状態値 (統合なし)

5. おわりに

本研究ではマルチモーダルな情報を段階的に統合し, 低速に変化する特徴量を抽出することにより, 運動状態や空間情報に対応するような表現が感覚運動経験から学習できる可能性を示した. また獲得されたマルチモーダルな表現は, 連続な状態空間を扱うヒューマノイドロボットの実環境中での強化学習に有用であり, 低速に変化する特徴量を状態変数に用いた学習によって適切に報酬を予測できることが確認できた. 今後の課題としては, 獲得した状態表現を用いて次の感覚入力や運動コマンドを予測できるモデルの設計を考えている.

文献

- [1] A. Iriki, M. Tanaka, S. Obayashi, and Y. Iwamura. Self-images in the video monitor coded by monkey intraparietal neurons. *Neuroscience Research*, Vol. 40, pp. 163–173, 2001.
- [2] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida. Cognitive developmental robotics: a survey. *IEEE Transactions on Autonomous Mental Development*, Vol. 1, No. 1, pp. 12–34, 2009.
- [3] M. Hikita, S. Fuke, M. Ogino, and M. Asada. Cross-modal body representation based on visual attention by saliency. *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 2041–2046, 2008.
- [4] S. Fuke, M. Ogino, and M. Asada. VIP neuron model: Head-centered cross-modal representation of the peri-personal space around the face. *Development and Learning, 2008. ICDL 2008. 7th IEEE International Conference on*, pp. 145–150, 2008.
- [5] L. Wiskott and T. Sejnowski. Slow feature analysis: Unsupervised learning of invariances. *Neural Computation*, Vol. 14, No. 4, pp. 715–770, 2002.
- [6] T. Minato, F. DallaLibera, S. Yokokawa, Y. Nakamura, H. Ishiguro, and E. Menegatti. A baby robot platform for developmental robotics. *Proc. of the Workshop on Synergistic Intelligence: Approach to Human Intelligence through Understanding and Design of Cognitive Development (IEEE/RSJ International Conference on Intelligent Robots and Systems)*, 2009.