

Slow Feature Analysis を利用したマルチモーダル情報の統合による身体表現の獲得

Acquiring body representation from multimodal information by slow feature analysis

○¹ 荻野 正樹, ² 西川 輝彦, ² 浅田 稔
○¹ Masaki Ogino, ² Akihiko Nishikawa, ² Minoru Asada
¹ 関西大学 ² 大阪大学
¹Kansai University ²Osaka University

Abstract It is still uncovered how body and spatial representation is calculated with various kinds of sensor information involved with motion. In human brain, the higher layer of the cortex is thought to represent wider spatial area and longer temporal period. This paper proposes a hierarchical model that integrates multi modal information that is acquired when an agent interacts with the world. In each layer, the slowly changed features are extracted from input signals. The simulation experiment shows that multimodal information related to self movement is transformed into lower dimensional data that changes slowly.

1 はじめに

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ているが、その獲得過程に関しては明らかにされていない。認知発達ロボティクス [1] 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 [?] [2]。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触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。西川らは、Slow Feature Analysis (SFA) [4] を利用し、入力される信号の時間遷移を考慮した状態空間の構築方法の提案を行った。彼らはヒューマノイドロボットの視覚、体性感覚、触覚などのマルチモーダル情報を階層的に Slow Feature Analysis を適用して統合することにより、強化学習に有用な状態表現を構築が可能であることを示している [3]。しかし、彼らは環境に対する自己の位置についての情報表現にとって SFA が有用であることを示したが、自己の身体表現にとって有用であるかどうかについては示していない。本研究では、そのための第一歩として、自己の身体の運動に関わるセンサー情報が SFA を使ったマルチモーダル情報の統合により、低次の低速の情報に変換されることを示す。

2 Slow Feature Analysis Network

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 [4]。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。

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

$\langle \cdot \rangle$ 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。

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

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

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

本研究では、Fig. 1 に示すように上記の SFA を階層的に組み合わせ、視覚情報と体性感覚情報（関節角度）の統合により、自己の身体情報を表現することを目指す。

3 シミュレーション実験

SFA Network の有効性を示すためにシミュレーション実験を行った。実験では3次元力学シミュレーター

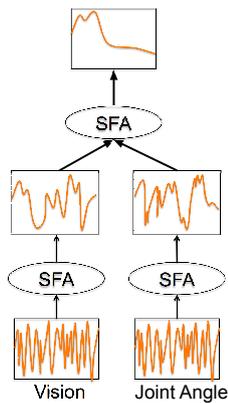


図 1: SFA によるマルチモーダル情報の統合

上で構築したヒト型のロボットを用いた。ロボットは胴体 3 自由度、腕、脚にそれぞれ 10 自由度で合計 23 自由度を持つが、実験では左腕 5 自由度の関節角度のみを SFA への入力として用いた。画像情報として右目の位置から 320×240 ピクセルの画像を得てロボットから得られる画像とし、その画像を 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。ロボットは左手を口周辺と前下方向の間を 2 秒間で往復させる運動を 5 回行う。この運動の間に得られるセンサ情報を、333 [msec] 間隔で取得し、SFA への入力信号とした。

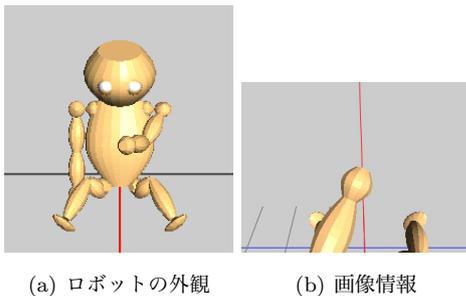
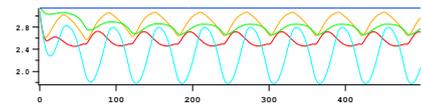
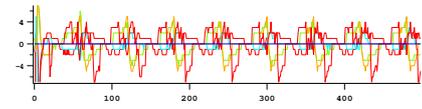


図 2: シミュレーションに使用したロボット

Fig. 3 に SFA に入力する前の関節角度情報、視覚情報を示す。また Fig. 4 に SFA により関節角度情報、視覚情報の信号を統合した後の、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。図より運動が安定する 250 ステップまでは大きな振幅の振動が見られるものの、それ以降は腕の上下運動が 2 つの曲線によって近似されることがわかる。また、画像には自己の身体の情報局所的に現れるにも関わらず、SFA によって全体として滑らかな情報へと変換されていることがわかる。



(a) 関節角度



(b) 画像のフロー情報 (一部)

図 3: SFA への入力信号

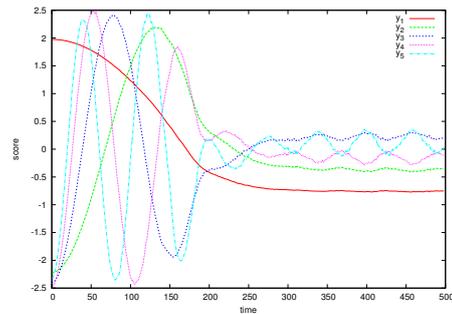


図 4: SFA によるマルチモーダル情報の統合

4 まとめ

本研究では、ヒト型ロボットの自己の身体運動に関わるマルチモーダル情報が、SFA の階層的なネットワークにより低次元の曲線へと変換されることを示した。今後は、より多様な動作を含んだデータをもとに統合表現の一般化と、頭頂葉で明らかになっている様々な座標系との比較を行い、その運動学習への適用の有用性について検証を行う予定である。

参考文献

- [1] M. Asada et al.: Cognitive developmental robotics: a survey. IEEE Transactions on Autonomous Mental Development, Vol. 1, No. 1, pp. 12-34, 2009.
- [2] S. Fuke et al.: VIP neuron model: Head-centered cross-modal representation of the pri-personal space around the face. ICDL 2008, pp. 145-150, 2008.
- [3] 西川ら: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, ROBOMECH2011, 2P2-M02, 2011.
- [4] L. Wiskott et al.: Slow feature analysis: Unsupervised learning of invariances. Neural Computation, Vol. 14, No. 4, pp. 715-770, 2002.

連絡先

荻野 正樹

E-mail: ogino@res.kutc.kansai-u.ac.jp