

内的観測に基づく関係同定を反映した相互適応系設計

— 人間-ロボット系における協調歩行の実現 —

東京工業大学 ○宮川透 三宅美博

Intra-observation based Identification and Control to the dynamical system characterized by unseparable interaction

Tokyo Institute of Technology ○ Toru MIYAGAWA Yoshihiro MIYAKE

Abstract - This paper discusses the method of system identification from the viewpoints of internal subsystems and its application to designing a mutual adaptive system includes human and the engineering system, as an example, bipedal locomotive robot.

Key Words : Intra-observation, unseparable interaction, mutual entrainment, Dual centers.

1 序

実世界に存在する多くのシステムは要素間の分離不可能な相互作用を特徴とする動的な系であり、Fig.1の4つのクラスに問題を分類することができる。Knownは完全情報、Separableは本来分離できないサブシステム間の相互作用を分離することで疑似的に問題を解決することを意味する。観測の視点を系の外部に持ち系全体の完全な情報に基づいて制御する外的観測に対し、内的観測とは系の内部に位置する観測者であるサブシステムがそれぞれ独自に、局所的で不完全な観測に基づいて行動することを指す。

我々は相互作用を双方向の同時的なものとして扱うためにリズムの相互引き込み(固有周期の異なる非線形振動子間の相互作用によって見かけ上の周期が一致すること)という現象に着目し、ロボットと人間の協調歩行などのタスクを設定した研究によって、相互引き込みに基づく協調歩行の達成や機能分散の実現などを示して来た。しかし従来の研究ではロボットの制御則は設計者が試行錯誤を経て予め与えており、恣意性が残されていた。本研究は、内的な観測に基づいて自己と環境とを統合した関係の同定を行い、その結果得られた内部モデルによって関係を操作することで目的を達成する、相互適応的なシステム設計の提案と検証を試みる。

2 モデル概略

今回提案するのは Fig.2 に示されるシステムである。 $s\Omega(t) = \{s\omega_r(t), s\omega_l(t)\}$, $s \in \{R'_1, R_2, V_1, V_2\}$ であり、 ω はシステム s の時定数、添字の r, l は右足、左足を、 (t) は時刻 t を意味している。

ロボット身体 R'_1 は神経振動子と筋骨格系から構成される2足歩行モデル [1] である。 R'_1 と R_2 は実時間で常時相互作用し、変化している。これに対し V_1, V_2 はそれぞれ R_1 内部に生成した R_1 と R_2 の像であり、 R_1 の接地の瞬間に仮想的な時間スケールで計算され動作する結合非線形振動子、具体的には (1) 式のように van der Pol 方程式に変形を加えたものとした。 $D_{rl} \geq 0$ は左右の足に見立てた振動子間の結合強度、 $\varepsilon \geq 0$ は減衰項である。

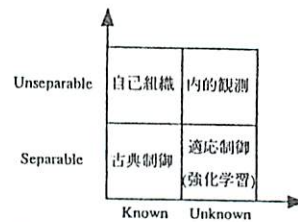


Fig.1

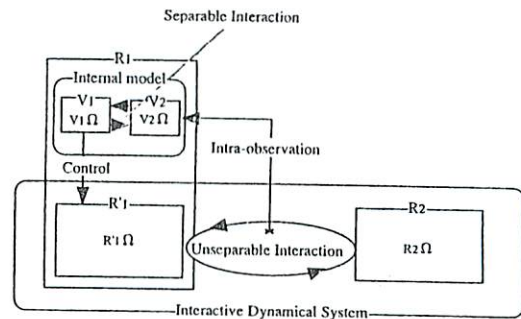


Fig.2

$$\begin{cases} \dot{x}_{ir} = y_{ir}, \\ \dot{y}_{ir} = \varepsilon(1 - x_{ir}^2)y_{ir} - v_1\omega_r^2 x_{ir} \\ \quad + D_{rl}(x_{il} - x_{ir}) + pulse_{ir}, \\ \dot{x}_{il} = y_{il}, \\ \dot{y}_{il} = \varepsilon(1 - x_{il}^2)y_{il} - v_l\omega_l^2 x_{il} \\ \quad + D_{rl}(x_{ir} - x_{il}) + pulse_{il}, \\ i \in \{1, 2\}. \end{cases} \quad (1)$$

ここでは内的観測の立場を採っているため、 R_1, R_2 間の関係性を反映した指標である位相差 $R_{12}pd(t)$ が唯一の観測量 $X(t)$ となる。この条件下で次のアルゴリズムを実行する。内部モデルでは $V_1\Omega, V_2\Omega$ の一方を固定し他方を変化させ、仮想的な時空で相互作用を分離することで問題を解いている。2. の条件文は位相差一定、つまり引き込みの状態にあるか否かを判定している。

1. 内的観測により、 R_1 は $X(t) = R_{12}pd(t)$ を得る。
2. if $|R_{12}pd(t) - R_{12}pd(t-1)| \leq \text{閾値}$
 - (a) V_2, V_1 間の位相差を $v_{21}pd(vt)$ とするとき、固定した $v_1\Omega(t-1)$ の下で

$$\min (R_{12}pd(t) - v_{12}pd(vt))$$

となる $v_2 \Omega(t)$ を求める。 vt は仮想的時間。

- (b) V_1 、 V_2 間の位相差を $v_{12}pd(vt)$ とするとき、固定した $v_2 \Omega(t)$ の下で目的

$$v_{12}pd(vt) = 0$$

を最もよく満たす $v_1 \Omega(t)$ を求める。

- (c) $v_1 \Omega(t)$ に対応する $R_1^i \Omega(t)$ へと変化させる。

- (d) flag = 1

3. 時間ステップ t を $t+1$ に進める。

4. R_1^i の行動に応じて R_2 は行動を変更し、 $X_{(t+1)}$ が生じる。

5. 内的観測により、 R_1 は $X_{(t+1)}$ を得る。

6. if flag == 1

- (a) (1) 式におけるモデル化誤差を逐次的に修正するため、 $v_{12}pd(t)$ と $R_{12}pd(t+1)$ を用いて誤差関数

$$E = (R_{12}pd(t+1) - v_{12}pd(t))^2$$

を定義し、内部モデルとして利用した結合非線形振動子系の正常な動特性を保證する区間内で E を減少させるようパラメータを変化させる。

- (b) flag = 0

7. 2. に戻る。

3 実験結果

現在、身体のみからなる $Robot_2$ を R_2 として用い、内部モデルを組み込んだ $Robot_1$ とロボット 2 体で計算機シミュレーションに基づいた実験を行っている。 $Robot_1$ に与えた目的は協調歩行、端的には $R_{12}pd(t) = 0$ の達成である。 $Robot_2$ の Ω を実験開始後 10 秒で変化させ、 $Robot_1$ 、 2 の相互作用は 60 秒から開始し、180 秒で再び $R_2 \Omega$ を変化させ計 300 秒間の実験を行った。Fig.3a は $Robot_1$ 、 2 の右足の歩行周期、Fig.3b は両者間の位相差、Fig.3c はモデル化誤差のグラフである。

また環境同定という従来手法に比して今回提案した関係同定の有効性を検証するため、強化学習を組み込んだロボットを用いた比較実験も行った。Fig.4a、Fig.4b はそれぞれ右足の歩行周期、位相差のグラフである。リアルタイム性が要求されることから強化学習のアルゴリズムは経験強化型的一种である Profit Sharing を用い、合理性定理 [2] を満たす強化関数を採用し、環境変動に対処するため加重の減衰を採り入れた。

Fig.3a と Fig.3b から、相互作用開始後間もなく相互引き込みの状態となり、この結果位相差=0 を満たすよう内部モデル (具体的には V_1 、 V_2 の時定数 Ω) が更新され、160 秒頃に目的が達成されている。これは相互作用開始から 100 秒ほど経過しており、適応速度が速いとは言えない。しかし Fig.3c に見られるように適応過程で内部モデルの誤差が減少した結果、180 秒時点での $Robot_2$ の変化にはほぼ瞬時に適応している。

一方、Fig.4a と Fig.4b を見ると、目的は完全には達成されておらず、また $Robot_2$ の突然の変化への対応は鋭敏ではない。強化学習における学習パフォーマンスの向上には状態空間や報酬の設計などが関連しているが、学習の速度と精度の双方を向上させるのは困難であり、また加重という形で過去の履歴が現在の行動に強力に影響するため、減衰の導入によっても動的に変化する環境

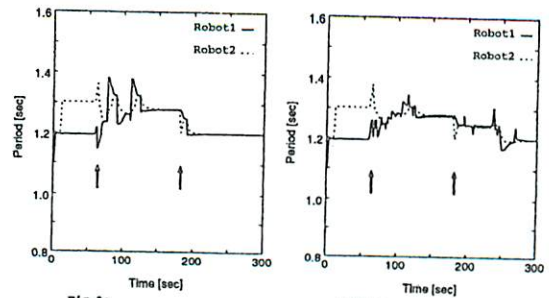


Fig. 3a

Fig. 4a

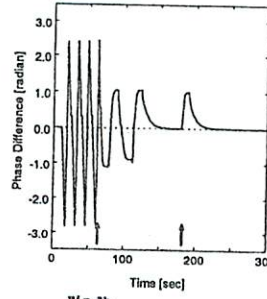


Fig. 3b

Fig. 4b

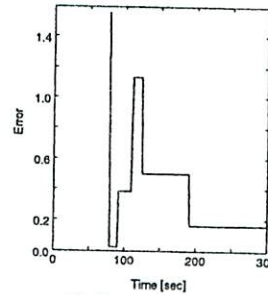


Fig. 3c

に対処するには困難が伴う。 R_2 として人間を想定した場合は試行毎、また試行中の揺らぎが不可避であることなどを併せ考えると、今回我々が提案したモデルは分離不可能な相互作用を特徴とする動的環境に対して有効であると思われる。

4 結語

本報告では、歩行ロボット 2 体でのシミュレーションの結果、実際の相互作用を行う身体系、関係性を反映し身体系を拘束する内部モデル、という 2 中心から構成されるシステム設計の有効性を検証し、確認した。今後は当初の目的である人間を含む系へと実験対象を拡張していきたいと考えている。

参考文献

- [1] G. Taga, Y. Yamaguchi, and H. Shimizu: Self-organized control of bipedal locomotion by neural oscillators in unpredictable environment, *Biol. Cybern.*, 65, pp.147-159 (1991).
- [2] 宮崎, 山村, 小林: 強化学習における報酬割当ての理論的考察, *人工知能学会誌*, Vol.9, No.4, pp.580-587 (1994).