

視覚情報に基づいたヒューマノイドのリズム歩行パラメータの学習

Learning Rhythmic Walking Parameters Based on Visual Information

加藤 豊 (阪大) 青野 正弘 (阪大) 正 荻野 正樹 (阪大)
正 高橋 泰岳 (阪大, 阪大 FRC) 正 細田 耕 (阪大, 阪大 FRC) 正 浅田 稔 (阪大, 阪大 FRC)

Yutaka KATO, Osaka University, 2-1, Yamadaoka, Suita, Osaka
Masahiro AONO, Osaka University
Masaki OGINO, Osaka University
Yasutake TAKAHASHI, Osaka University, HANDAI Frontier Research Center
Koh HOSODA, Osaka University, HANDAI Frontier Research Center
Minoru ASADA, Osaka University, HANDAI Frontier Research Center

This paper presents a method for generating vision-based humanoid behaviors by reinforcement learning with rhythmic walking parameters. The walking is stabilized by a rhythmic motion controller. The learning process consists of two stages: first one is building an action space so that infeasible combinations of walking parameters are inhibited. The second one is reinforcement learning with the constructed action space and the state space consisting of visual features and posture parameters to find feasible action. The method is applied to a situation of the RoboCupSoccer Humanoid league.

Key Words: humanoid, reinforcement learning, rhythmic walking, RoboCupSoccer

1 はじめに

人間が普段生活するような複雑な環境において、2脚歩行ロボットの適応的な歩行を実現するためには視覚によって環境を観測し、その状況に応じて歩行運動を調節するような適応的歩行制御器を考える必要がある。従来手法^{1, 2)}では、視覚情報に基づいた歩行の調節を設計者がプログラムしていた。しかし、時々刻々と変化する環境中で設計者があらかじめ起こり得る全ての状況を予測し、歩行の調節を記述することは一般的に困難であると考えられる。

本研究では、視覚情報に基づいた歩行の調節をロボット自身が外界との相互作用を通して自律的に獲得する学習的アプローチを用いる。振動子モデルの原理に基づき設計されたリズム歩行制御器によって2脚歩行は実現されているものとし、視覚情報によってその歩行パラメータを調節する方策を自律的に獲得する。学習過程は2段階で構成され、第1段階でロボットの転倒や脚の干渉を防ぐためにロボットの姿勢に応じた行動空間を構成する。第2段階で、構成された行動空間と視覚と姿勢で構成された状態空間を用いて強化学習³⁾により環境に応じた歩行パラメータを調節する方策を獲得する。

提案手法をサッカータスクに適用し、その有効性を検証する。

2 2脚歩行ロボットのモデル

Fig. 1 に示す2脚歩行ロボットを考える。ロボットは1リンクの胴体と、左右各4リンクの腕、左右各6リンクの脚から構成される。それぞれのリンクは図に示されるように1回転自由度のジョイントで結合されている。各脚先には4個の床の垂直抗力を検出する反力センサが付いており、胴体上部には魚眼レンズを装着した CCD カメラが搭載されている。

3 視覚情報に基づいた歩行運動の調節

Fig. 2 に視覚情報による歩行運動の調節機構の概要図を示す。全体の構造は2層で構成され、視覚情報により歩行運動を調節する機構を上位層、歩行運動を生成する機構を下位層とする。上位層は1つの学習器から成り、学習器は視覚とロボットの姿勢から構成された状態を認識し、ゴール状態により報酬を発生させ、動的計画法³⁾を

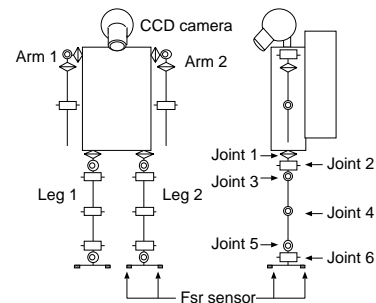


Fig.1 Model of biped locomotion robot

用いて歩行運動を調節する行動を学習する。下位層は振動子系を用いたリズム歩行制御器から構成され、上位層から出力された行動に応じて歩行パラメータを変更し、歩行運動を生成する。センサシステムは、脚裏に取り付けられた反力センサと各関節に取り付けられた角度センサ、視覚情報を得るための CCD カメラから構成される。

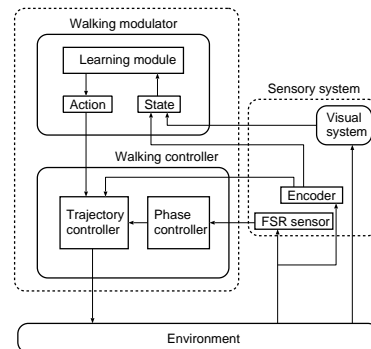


Fig.2 Biped walking system with visual perception

3.1 リズム歩行制御器

土屋ら⁴⁾は、振動子モデルの原理に基づいた歩行ロボットの歩行制御器を設計し、各脚に対して設置された振動子間の相互引き込みにより、環境に適応した歩行を生成している。本研究では、土屋らの振動子モデルに基づいてリズム歩行制御器を設計する。Fig. 3に示すように、歩行制御器は各脚の往復運動を制御する軌道制御器と、各脚の位相関係を制御する位相制御器から構成される。軌道制御器は位相制御器からの指令値に従ってモーターコマンドを生成し、モータを駆動する。位相制御器は振動子で構成され、脚先に取り付けられた反力センサ信号によるフィードバック信号を受け取り、各振動子の位相を変化させる。

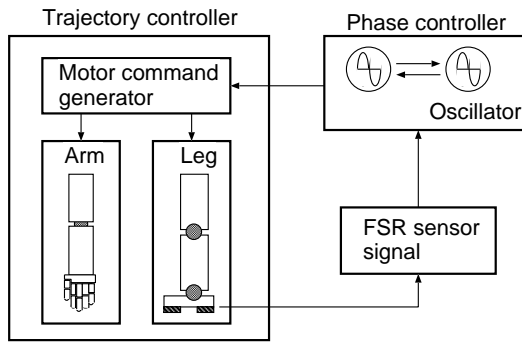


Fig.3 Walking controller

歩行は以下のように生成される。

1. 各脚の関節に対する目標軌道は、対応する振動子の位相の関数として与えられ、遊脚軌道、支持脚軌道の2つの軌道を切替えることにより生成される。
2. 位相は位相制御器で計算され、反力センサからのフィードバック信号と振動子間の相互作用により制御される。
3. 歩行パラメータ(回転角度、歩幅)を変化させることにより、異なる歩行運動を生成する。

3.2 ロボットの姿勢に応じた行動空間の構成

歩行制御器は歩行パラメータを変更することにより、異なる歩行運動を生成することができる。しかし、歩行パラメータを適切に決定しなければ、ロボットがバランスを崩して転倒する、脚同士が干渉する等の問題が生じる。これらを防ぐためにロボットの姿勢に応じた行動空間を構成する。以下にその手順を示す。

1. 両脚支持期にロボットの姿勢を観測する。
2. 学習器はランダムに行動を決定し、歩行制御器に入力する。歩行制御器は入力された行動に従い歩行運動を実行する。
3. 転倒や脚の干渉が起こらなければ手順1に戻る。
4. 転倒や脚の干渉が起こると、学習器は手順1で観測した姿勢の時に、手順2で選択した行動を取らないように学習し、手順1に戻る。

3.3 歩行パラメータの学習

学習器は、観測した状態における最適な行動(歩行パラメータの調節)を環境との相互作用を通して獲得していく。学習器は状態 s で行動 a をとった時に次状態 s' に遷移する確率 $P_{ss'}^a$ を環境との相互作用を通して更新し、動的計画法を用いて次式により状態 s で行動 a を取った時の行動価値 $Q(s, a)$ を計算する。

$$Q(s, a) = \sum_{s'} P_{ss'}^a [\mathcal{R}_s + \gamma \max_{a'} Q(s', a')] \quad (1)$$

ここで、 \mathcal{R}_s は状態 s における報酬の期待値である。行動価値関数 $Q(s, a)$ より、最適な政策 $f(s)$ は次式で求められる。

$$f(s) \leftarrow a \text{ such that } \max_{a' \in A} Q(s, a') \quad (2)$$

4 実験結果・考察

提案手法の有効性を検証するために実機を用いて実験する。具体的なタスクは、ゴールにボールをシュートすることである。実験で使用する2脚歩行ロボットは富士通オートメーション株式会社が研究プラットフォームとして開発した小型ヒューマノイドHOAP-1⁵⁾を使用する。そして、環境を観測するためにHOAP-1の頭部に魚眼レンズを装着したCCDカメラを搭載する。

タスクを実行する前に、ロボットの姿勢に応じた行動空間を構成する。獲得された行動空間を用いて、タスクを行った結果の一例をFig. 4に示す。図より、ロボットが歩行を調節してボールをキックできる位置に移動していることがわかる。以上より、視覚情報に基づいた歩行運動の調節を環境との相互作用を通して自律的に獲得することができた。



Fig.4 Result of experiment

5 結論

本研究では、2脚歩行を振動子モデルの原理に基づき設計されたリズム歩行制御器によって実現し、視覚情報によってその歩行パラメータを調節する方を自律的に獲得する手法を提案した。提案手法をサッカータスクに適用し、視覚情報に基づいて歩行を適応的に調節する方を獲得できることを示した。今後の課題としては、学習器の状態空間をロボット自身が学習を通して自律適に構築する事が考えられる。

参考文献

- [1] G. Taga, Y. Yamaguchi, and H. Shimizu. Self-organized control of bipedal locomotion by neural oscillators in unpredictable environment. *Biological Cybernetics*, Vol. 65, pp. 147-159, 1991.
- [2] H. Kimura, S. Akiyama, and K. Sakurama. Realization of Dynamic Walking and Running of the Quadruped Using Neural Oscillator. *AUTONOMOUS ROBOTS*, Vol. 7, No. 3, pp. 247-258, 1999.
- [3] Richard S. Sutton, Andrew G. Barto, 三上貞芳, 皆川雅章 共訳. 強化学習. 森北出版株式会社, 2000.
- [4] 土屋和雄, 青井信也, 辻田勝吉. 振動子系を用いた2脚歩行ロボットの歩行制御. *SICE SI2002 講演論文集* (2), pp. 151-152, 2002.
- [5] 村瀬有一, 安川 裕介 境克司, 植木美和. 研究用小型ヒューマノイドの設計. 第19回日本ロボット学会学術講演会, pp. 789-790, 2001.