

AFOによる鬼ごっこ

梅本雅之¹, 西岡良太¹, 山形佳史¹, 川村秀憲¹, 大内東¹ 高谷敏彦²,
1, 北海道大学大学院情報科学研究科
2, リコーソフトウェア

概要

AFO(Autonomous Flying Object)とは、3次元空間を移動可能な小型バルーンロボットである。近年、バルーンロボットをエンタテインメントの分野に応用する研究が盛んになっている。エンタテインメント性により注目すると複数ロボットによる戦略的なゲームが考えられる。本研究では、戦略的なゲームとして鬼ごっこをモデル化したゲームを定義し、シミュレータ上で共進化GAで学習させたニューラルネットワークを用いた制御により、行動戦略を自律的に獲得する。環境やモータ特性の違いを生かした戦略を獲得し、フェイント行動や壁に追い込む行動戦略を獲得した。

Tag game of Autonomous Flying Objects

Masayuki Umemoto¹, Ryota Nishioka¹, Yoshihumi Yamagata¹, Hidenori Kawamura¹
, Azuma Ohuchi¹, Toshihiko Takaya¹

1, Graduate School of Information Science and Technology, Hokkaido University
2, Ricoh Software Inc.

Abstract

Autonomous Flying Object (AFO) is an autonomously controlled indoor robot that is developed to realize autonomous flights for an entertainment. The balloon robot can fly freely in a three dimensional space. In this paper, the tag game by two robots is concerned as an application of the entertainment flights. The complexity or highly technical strategy of behaviors for the game can be an important factor for the entertainment. The combination of genetic algorithm and neural networks is used to co-evolve the behavioral strategies for the robots. The complexity of the evolved behaviors and techniques is investigated in response to the space and motor characteristics.

1. はじめに

AFO(Autonomous Flying Object)とは浮力を利用して空中に浮き、3次元空間を移動することができる小型バルーンロボットである。階段や段差などの床面に依存しない移動が可能であり、バルーンの形状もある程度自由に設計できる。また、他の飛行ロボットと比較して安全性も高い為、安全に楽しめるエンタテインメントを提供できるロボットとして注目を集めている。バルーンロボットを用いた研究として、拍手に反応して動作を変化させるインタラクティブロボットの開発[1]や、複数ロボットが協調したシンクロ飛

行などエンタテインメント分野での研究が盛んに行われている。更にエンタテインメント性をより追求すると複数ロボットによる高度な戦略を持つプレイヤー同士でのゲームを実現する事で新たなエンタテインメントを提供する重要な要素の一つだと考えられる。ゲームという性質上、相手の戦略に対してより強固な戦略を自律的に獲得する必要性が生じる。これはエンタテインメントに限らず、マルチエージェント環境において有効な自律的な行動獲得の実現が可能であると考えられる。そこで、本研究ではバルーンロボットを用いた戦略的な行動を伴うゲームの実現を目指す。ゲ

ームを実現するには、戦略的な行動を伴った制御や戦略を獲得する仕組みが必要である為、本研究ではニューラルネットワークによる制御を導入する。また、ニューラルネットワークのパラメータの学習方法として、共進化遺伝的アルゴリズム(以下、共進化GA)により、相手の戦略に対し強固な戦略を自律的に獲得する。ゲームとして他のゲームにも応用可能だと考えられる鬼ごっこを対象とし、バルーンロボットをモデル化したシミュレータを用いて最適な環境を検討する。また、獲得した行動戦略によるバルーンロボットの振る舞いについて分析を行う。

2. 鬼ごっこゲームの定義

鬼ごっこゲームのモデルについて説明する。追跡者と逃走者2個体のプレイヤーにより、半径 $R[m]$ ・高さ $H[m]$ の仮想的に制限された3次元円柱空間内で制限時間 $T[sec]$ 試行する。また初期状態として、 $d_{init}[m]$ 離れた状態からスタートし、制限時間内に捕獲に成功する(追跡者勝利)か $T[sec]$ 経過した時点(逃走者勝利)でゲームを終了とする。

プレイヤーの目的として追跡者は制限空間内で逃走者を短時間で捕獲することを目指し、逃走者は制限空間内で追跡者から長時間逃げる事を目的とする。このゲームの目的関数を以下のように定式化する。

追跡者	$1 - \frac{1}{T} \sum_t \frac{d_t}{d_{init}} + \frac{ct}{T} - P_p$
逃走者	$\frac{1}{T} \sum_t \frac{d_t}{d_{init}} - \frac{ct}{T} - P_e$

ここで、 ct は捕獲時間、 d_t はタイムステップ当りのプレイヤー間の距離、 P は空間サイズを超過した罰則を表す。また、パラメータの下付き文字 p は追跡者(pursuer)、 e は逃走者(evader)を表すものとする。

3. バルーンロボット

本研究で対象とするバルーンロボット[2]のハードウェア構成は、ヘリウムを充填させたバルーン部と制御を行う駆動部からなる。更に、駆動部は位置情報を取得するセンサ部、センサ情報から制御出力を計算する制御部、制御出力を用いてプロペラを回転させる推進部から構成される。ハードウェア構成を図1に示す。

3.1 バルーン部

小型バルーンロボットは、ペイロードの制約が大きい為、容量が大きく浮力を獲得しやすい円柱型のバルーンを利用する。また、円柱形を用いる理由として、室内の狭い空

間を3次元に移動させることが可能、水平方向の移動に対して空気抵抗の偏りが少ないことが挙げられる。駆動部の重量に対して必要な浮力を算出し、直径 $94[cm]$ 、高さ $80[cm]$ の寸法とする。

3.2 センサ部

飛行船に搭載したカメラで、床面上に配置したランドマークを認識することで現在位置を取得する。

カメラはRGB16bitのカラー画像を $160 \times 144[px]$ の解像度で CPU に取り込むことが可能である。

3.3 制御部

制御部は T-Engine ボードと RBTMC ボードから成る。T-Engine ボードはセンサ部で取得した現在位置を基に制御出力を計算し、RBTMC ボードでモータに制御出力を送る。また、制御出力の計算は $0.3[sec]$ 間隔で計算される。

3.4 推進部

推進部は3次元方向に移動可能な様に $ch0 \sim ch5$ の合計6つのモータとプロペラが十字に組まれたロット上に固定されている。 $ch0 \sim ch3$ のモータは独立に水平方向の推力及びヨー角に対する回転運動。また、 $ch4$ 及び 5 は2個のモータを同期して垂直方向の推力を発生させる。また、これらのモータの ON-OFF を切り替えるタイミングを $0.01[sec]$ 間隔に設定することで $0.3[sec]$ 内では正回転・逆回転ともに30段階の推力を発生させることが可能である。

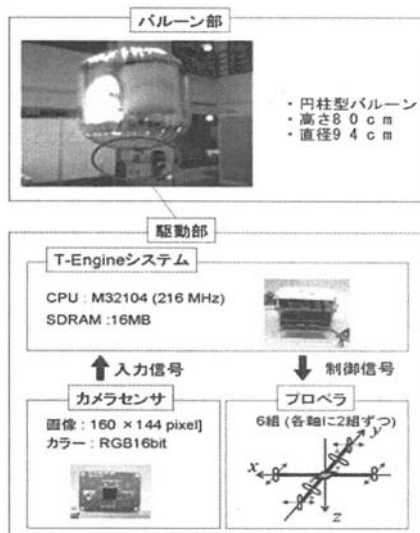


図1 Hardware Configuration

3.5 シミュレータ

バルーンロボットをモデル化し、運動方程式に基づき飛行船の状態量を算出するシミュレータを開発した。以下に用いた運動方程式を示す。

$$M\dot{v} = Av + F + B + \tau$$

ここで、 M はバルーンロボットの質量行列の項、 \dot{v} は加速度ベクトル、 A は空気抵抗、 F は遠心力とコリオリ力による項、 B は浮力と重力による項で τ はバルーンロボットの推力を表す。運動方程式は3次元上での絶対座標系 (X, Y, Z 座標)の力のつり合いと軸周りの回転座標 (X 軸, Y 軸, Z 軸周り)の力のモーメントのつり合いを考慮する。また、バルーンロボットは自己の座標系を持っており、回転の変換式により絶対座標系に変換する。算出した飛行船の状態量を基に OpenGL により、視覚化する。

4. 制御方法

バルーンロボットの制御方法として PID による制御方法 [3] が多くのバルーンロボットで研究されており、あらかじめ与えられたタスクを実現する制御ではその有効性が確認されている。しかし、従来の PID 制御では、動的な環境の中で、他のバルーンロボットの戦略に対応したり、最適な行動を自律的に獲得したりする事は非常に困難である。その為、本研究では、動的な環境に柔軟に対応できるニューラルネットワークによる制御方法を採用する。また、ニューラルネットワークの学習法として相手の戦略に効果的な戦略を獲得する必要性から共進化GAを用いてネットワークを最適化する。

4.1 ニューラルネットワークによる制御

本論文では、入力層15、中間層12、出力層6のニューラルネットワークを適用し、制限された空間内で相手を追跡(または逃走)する制御を行う為、ニューラルネットワークへの入力として、機体間の相対座標 (x, y, z) 仮想的な壁との相対座標 (w_x, w_y, w_z) 移動偏差 (dx, dy, dz) と前回の出力を入力とする。また、出力層は水平方向の4つのプロペラ ($ch0 \sim ch3$) への推力と1つの垂直方向の推力 ($ch4, ch5$ は同期)を返し、1つは内部記憶として出力には影響せず、そのまま入力層に戻す。リカレントや内部記憶構造を持つ事で行動戦略を内蔵した制御を実現する。ニューラルの構造を図2に示す。

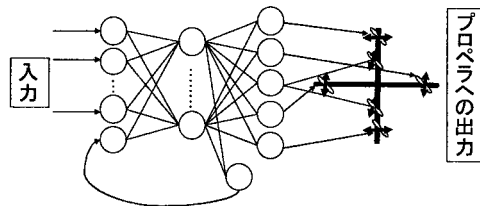


図2 ニューラルネットワーク

4.2 共進化GA[4]

相手の戦略に効果的な戦略を獲得する為共進化GAを用いてニューラルネットワークのパラメータの学習を行う。遺伝子としてニューラルネットワークの結合係数と閾値を持つ個体を考え、逃走者と追跡者各々にN個の個体集合を用意する。遺伝子操作は simple-GA と同様に行い、選択はトーナメント選択、交叉は1点交叉、突然変異は $[-1, 1]$ の一様分布に従う乱数と置き換える。

また、交叉確率 0.2、突然変異確率 0.6 で選択し、個体数はプレイヤー各々に対し、100個体、世代数は1000とする。

また評価値の算出は、図2の様に対戦させた結果得られた目的関数より相対的に算出する。T 世代における逃走者の個体の評価値 fit_t の算出方法は

1. T-1 世代の追跡者の最優良個体とT世代の逃走者N個体に対戦させてゲームで定義した目的関数より評価値を算出
2. T-1 世代の追跡者の2番目に優秀な個体とT世代の逃走者N個体に対戦させて評価値を算出
3. T-1 世代の追跡者の3番目に優秀な個体とT世代の逃走者N個体に対戦させて評価値を算出
4. 各個体に対して 1~3の評価値の平均をとりその個体の評価値とする

これは、相手の最優良個体1個体とだけ共進化を行うと、同じ箇所を堂々巡りをしている可能性や進化が不安定になり上手く進化を遂げない為、相手の個体3個体と共進化を行った。追跡者の場合も同様に対戦を行い、評価関数に従い評価値を算出し、世代数遺伝子操作を繰り返す、個体を進化させる。図3に共進化の様子を示す。

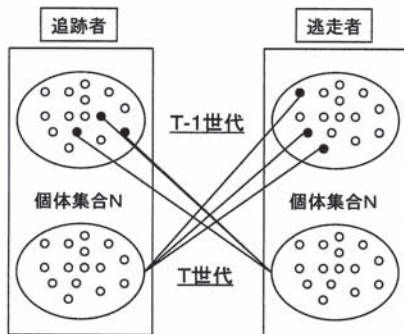


図3 共進化

T世代のN個体全てとT-1世代の相手の優秀個体3個体の対戦を行い、3回の評価値の平均をとり評価値を算出する

5. 実験

鬼ごっこゲームの環境を設定する際、追跡者と逃走者のバランスが均衡しているとき程、戦略が有効に作用し、新たな高度な戦略が獲得され易いと考えられる。また、環境要因は戦略を進化させる上で重要な要素のうちの一つである、まず最初に、同じ運動特性の2機体による鬼ごっこを行い、環境を変化させた際の環境への適応限界と振る舞いの変化を分析する。また、自然界に見られるように各個体の運動特性もまた捕食者・被捕食者で独自の進化が起こり、戦略の進化と共に変化してきた。よって異なる運動特性の2機体を用いて鬼ごっこを試行し、運動特性の違いによる振る舞いの変化を分析し、最適なゲーム環境を構築する。

ゲームの環境として空間の高さ制限を H を $5[m]$ と水平方向と比較して小さな値を設定しているが、これはパルーンロボットは床面のランドマークから位置を算出する為、垂直方向の制御には限界がある。また、鬼ごっこゲームというエンタテインメント飛行を想定しているため、人が視覚的に捕らえられる範囲内で飛行を行う必要があると考えられる。よってここでは、 $5[m]$ という環境を想定する。

5.1 空間サイズを変更

同じ運動特性で環境の異なる中でゲームを行う。

ゲームの環境の各パラメータは以下の様に設定する。

制限時間: $T = 300[sec]$

空間サイズ: 半径 R , 高さ $H=5$ の円柱空間内

初期距離: $d_{mit} = R$

R を $5\sim 50[m]$ まで $5[m]$ 間隔で変更して実験

学習済みのニューラルネットワークを用いて環境サイズを

変更して100試行ゲームを行い、捕獲回数と平均捕獲時間を測定する。実験の初期配置は図4の通りである。

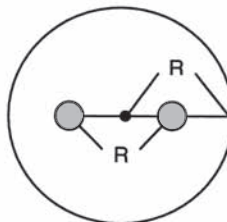


図4 2機体のプレイヤーの初期配置

環境サイズ R を $5\sim 50$ まで変更し、学習されたNNで100ゲーム試行した、結果を図4に示す。図4は横軸に空間サイズ R 、縦軸に捕獲回数(図中▲)と平均捕獲時間(図中■)をプロットする。

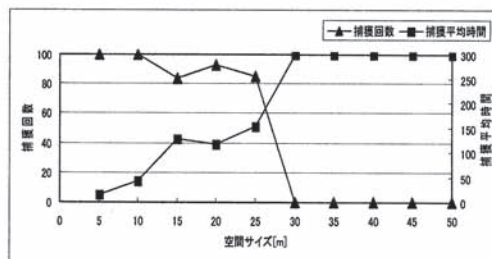


図5 捕獲回数と捕獲平均時間

結果より、 R が 10 以下のときは、追跡者は短時間で必ず捕獲する事が出来ている。 $R=15\sim 25$ の空間サイズ内において追跡者と逃走者のカバランスが均衡しつつあり、それ以上になると、逃走者は必ず逃げ切る事が可能となっている。カバランスが均衡している方がより効果的な行動戦略を獲得できると考えられるので、ここで、 $R=15$ のときの2機体による鬼ごっこを試行し、得られた行動戦略に基づいた振る舞いを分析する。

今回は軌跡を分析する為に衝突を行わない設定にし、制限時間 $T[sec]$ 試行した軌跡を示す。XY 平面上の軌跡を図6、速度の時間推移を図7、2点間距離と衝突範囲の時間推移を図8に示す。図8より一方的に距離が縮められ、 $120[sec]$ で捕獲されている事が分かり、図6より逃走者は初期状態からプロット間隔が非常に狭く初期位置からの移動距離が非常に少なく、追跡者に壁に追い込まれて捕獲されている、これは蛇行しながら追跡者が移動していることや、図9の速度推移において追跡者の速度が大きく振動

しながら推移していることから、追跡者は正回転と逆回転を交互に切り替える事で、逃走者の移動距離と移動可能な方向を制限し、壁に追い込む戦略を獲得している事が確認できる。

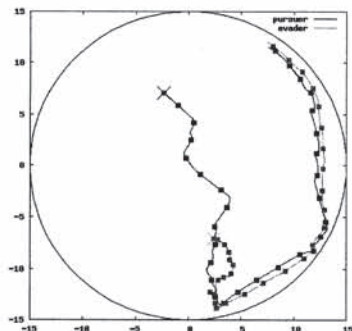


図6 XY平面上での軌跡

(×:初期位置, ■10[sec]間隔)

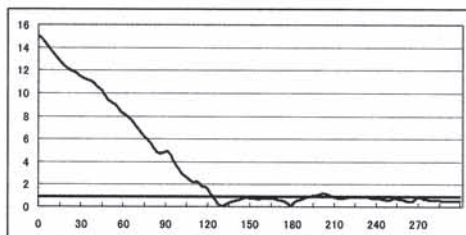


図7 XY平面での2点間距離と衝突範囲の時間推移

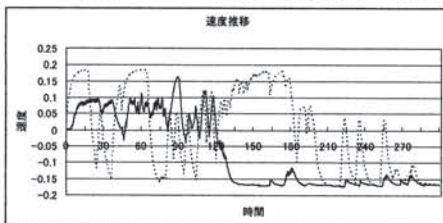


図8 速度の時間推移

初期状態の違いで捕獲可能かどうかを物理的に判断すると、今、試行時間が 300[s]でバレーンロボットの最高速度が約 0.2[m/s]なので、限界移動可能距離は 60[m]である。従って、初期距離が $d_{init}=R$ で図4より壁までの最短距離が $\frac{R}{2}$ なので、 $\frac{3}{2}R > 60$ だと直線距離で逃げるだけで逃げ切れてしまう。従って、 $R > 40$ の際には合理的なプレイヤーを捕獲することは不可能である。また、逃走者は壁との相対距離を考慮しながら、逃走する必要があるが、合理的な追跡者は壁との距離を考慮せずに逃走者だけを追跡する

ことが可能となるので、距離を縮めて最終的には壁に追い込んで捕獲することが可能だと考えられる。

よって、探索空間が狭い追跡者の方が進化が早く進み、逃走者の十分な進化の前に、より強固な戦略を獲得してしまい、逃走者の進化が停滞し、全体としての共進化も効果的に作用せず、行動戦略も獲得できない。

よって、実験②として逃走者に有利な運動特性を与えてゲームを行い、逃走者と追跡者の効果的な共進化により、高度な行動戦略の獲得を目指す。

5.3 運動特性を変更

ゲームの環境の各パラメータは①と同様に設定し、

実験①の結果より、 $R=25[m]$ と固定し運動特性を変更する。運動特性の変更は逃走者の加速度を追跡者の4倍にし、最高速度は同じ速度に設定する、つまり、加速度の違いによる振る舞いの変化を分析する。

運動性能を変更した結果獲得した行動戦略による特徴的な飛行船の振る舞いについて示す。XY平面上での軌跡を図9に示し、速度の時間推移を図10に示す。

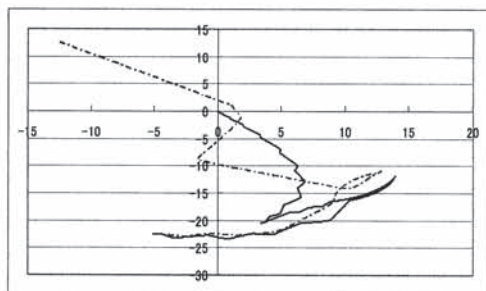


図9 XY平面での軌跡(実線:逃走者, 破線:追跡者)

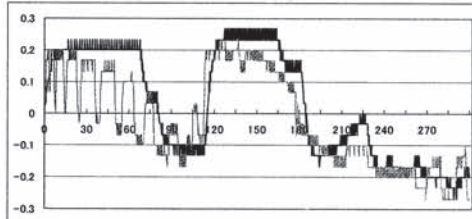


図10 速度の時間推移 (黒:追跡者、灰色:逃走者)

逃走者は加速度の違いを生かして、小刻みに振動しながら追跡者から逃走する行動戦略を獲得した。また、飛行船の特性として慣性の影響が強く働くという性質があり、逃走者の加速度を活かした大きくフェイントをするという戦略に対して、追跡者は慣性の影響により引っ張られ、捕獲に失敗している。

6. おわりに

本研究では、共進化GAで学習させたニューラルネットで制御された2機体のバルーンロボットによる鬼ごっこゲームを提案し、シミュレータ上でその振る舞いを観察した。ニューラルネットによる制御で飛行船の特性・空間サイズや運動特性が異なるなど多様な環境の変化に対応し、最適な行動戦略を自律的に獲得することを確認した。また、共進化の過程で相手を壁に追い込む戦略やフェイント行動を創発し、高度な行動戦略を自律的に獲得した。また環境により、生成された高度な行動戦略に対し、有効な行動戦略を共進化により自律的に獲得することを確認した。

謝辞

本研究を進めるにあたって多大なご協力、ご助言をいただいた、リコーソフトウェア株式会社に深く感謝いたします。

参考文献

- [1]川村秀憲, 角田久雄, 山本雅人, 高谷敏彦, 大内東, “ホバリング制御に基づくエンタテインメントバルーンロボットの開発”, 知能・情報・フuzzy学会論文誌, Vol. 17, No. 2, pp. 29-37, 2005.
- [2]高谷敏彦, 川村秀憲, 皆川良弘, 山本雅人, 大内東, “屋内小型自律飛行船の自動巡回システム”, 22nd Fuzzy System Symposium, 8A1-1, pp. 685-688, 2006
- [3]角田久雄, 川村秀憲, 山本雅人, 高谷敏彦, 大内東, “カメラ搭載型バルーンロボットシステムの開発と PD 制御による位置制御の実現”, 情報処理学会論文誌, 45 巻, 6 号, pp. 1715-1725, 2004.
- [4]D. Cliff and G. F. Miller, “Co-evolution of pursuit and evasion II : Simulation methods and results.”, In Proc. of the 4th International Conference on Simulation of Adaptive Behavior: From Animals to Animals 4, pp. 506-515 (1996).