論　文

ポテンシャルゲーム理論的姿勢協調 —同期・平衡の達成*

畑中　健志†・後藤　達彦†・藤田　政之†

Potential Game Theoretic Attitude Coordination —Attitude Synchronization and Balanced Circular Formation*

Takeshi Hatanaka†, Tatsuhiko Goto† and Masayuki Fujita†

In this paper we consider potential game theoretic attitude coordination. We especially focus on two ordered configurations: “synchronization” and “balanced circular formation”. We first show that both problems constitute potential games by employing some global and individual objective functions, and a learning algorithm called Restrictive Spatial Adaptive Play (RSAP) leads robots to the ordered configurations with high probability even in the presence of mobility constraints. We moreover show that the problem also constitutes a group-based potential game and convergence of distribution of actions to stationary one can be accelerated. Finally, the effectiveness of the schemes is demonstrated by simulation.

1. はじめに

安価で消費電力の低いセンサや無線通信機器などの要素技術の発展によってセンサネットワーク，ロボティックネットワーク [1]，交通ネットワーク，電力ネットワークに代表されるネットワークのスマート化が望まれている．これらのネットワークにおいては，個々の構成要素が限られた情報交換のもとで協調的に動作し，所望の目的を達成することが求められる．このような目的を達成する制御は協調制御とよばれ，合意問題，同期問題 [2–4]，被覆問題 [5]，フォーメーション制御問題 [6]，資源配分問題など，様々な協調制御問題が提案されてきた．このうち，本論文では姿勢協調問題 [2–4] を扱う．

姿勢協調問題は 2 次元および 3 次元空間上の剛体の姿勢を，分散制御則によってある秩序だった配置に収束させる問題である．姿勢協調に関する研究は，自然界の鳥や魚の群れの解析 [2] という自然科学的な好奇心に端を発するが，海洋調査のためのモバイルセンサネットワークへの適用も検討されている [3]．さらに，ビジュアルセンサネットワーク [7] におけるカメラの姿勢制御も本協調制御問題の有力な適用対象である [8]．本論文では，姿勢協調問題のうち，とくに同期と平衡という代表的な姿勢配置を考える．これらを達成する制御則はすでに存在するが [2,3]，とくに姿勢平衡については大域的に達成を保証する方法は報告されていない．さらに，障害物などの行動制約のもとで協調を達成する方法は存在しない．

他方，協調制御問題を非協力ゲームとして捉え，その均衡解への収束を制御目的の達成と同一視する枠組が文献 [9] によって提案されている．ゲーム理論を用いる利点は，故障や環境変化に対するロバスト性，通信要求の削減，大域的な収束の保証など，これまでの協調制御にはない機能を実現できる点にある．とくに，文献 [9] では，代表的な協調制御問題の多くがポテンシャルゲーム [10] とよばれる特殊なゲームによって表現されることを示すとともに，目的を達成する均衡解への到達手法として“ Restrictive Spatial Adaptive Play (RSAP) ”なる学習アルゴリズムを提案した．その後，協調制御に対するゲーム理論的アプローチが複数提案されている [11–13]．

本論文では，ポテンシャルゲーム理論に基づいて，姿勢同期と姿勢平衡なる姿勢協調を考察する．この目的のために，まずこれらの姿勢協調問題をゲームとして表現する．つぎに，適当な仮定のもとでこれらの姿勢協調ゲームがポテンシャルゲームを構成することを証明する．これにより，学習アルゴリズム RSAP を用いれば，十分時間が経過した後，行動制約のもとでの姿勢協調が初期状態によらず高確率で達成される．ただし，RSAP では 1 ステップに 1 台のエージェントしか動くことができず，

* 原稿受付　2010 年 11 月 5 日
† 東京工業大学 大学院 理工学研究科　Graduate School of Science and Engineering, Tokyo Institute of Technology; 2-12-1 Ookayama Meguro-ku, Tokyo 152-8550, JAPAN
Key Words: cooperative control, potential game.

上の行動選択が実現されるまでに時間を要するという問題がある．そこで, 群ポテンシャルゲームの概念を姿勢協調問題に適用し，RSAP を加速する新たな学習アルゴリズムを提案する．最後にシミュレーションにより，本研究の有効性を検証する．

本論文の構成は以下の通りである．まず，**2.** では本論文で考える姿勢協調問題を定式化したのち (**2.1, 2.2**)，それをゲームとして捉える枠組みを提示する (**2.3**)．ここで，基本的なポテンシャルゲームの構造は

- ポテンシャルゲームを構成する利得関数の設計
- 利得関数を利用した行動決定アルゴリズムの設計

から構成される [14]．本論文では，**3.** において姿勢協調問題における利得関数の設計を行い，**4.** において既存の行動決定アルゴリズムを高速化する新たな学習アルゴリズムを提案する．**5.** ではシミュレーションによって，本論文の枠組みと提案アルゴリズムの有効性を検証し，最後に **6.** に結論を述べる．

2. 問題設定

2.1 ロボティックネットワーク

本論文では，2 次元空間に存在する N 台のエージェント $\mathcal{V}:=\{1,...,N\}$ から構成されるロボティックネットワーク [1] を考える．エージェント $i\in\mathcal{V}$ の姿勢を $\theta_i\in[0,2\pi)$，エージェント i と j の相対姿勢を $\theta_{ij}=\theta_i-\theta_j$ と表記する．本論文を通して，各エージェントの姿勢の動きは次式の運動学モデルに従うものとする．

$$\theta_i(k+1)=a_i(k),\ a_i(k):=\theta_i(k)+u_i(k)$$

ここで，$u_i(k)$ はエージェント i に対する速度入力であるが，本論文では簡単のため $a_i(k)=\theta_i(k)+u_i(k)\in\mathcal{A}_i\subset[0,2\pi)$ を設計対象とし，$a_i(k)\in\mathcal{A}_i\subset[0,2\pi)$ を時刻 k における行動とよぶこととする．集合 $\mathcal{A}_i$ は行動集合とよばれ，すべてのエージェントの行動集合を集めて $\mathcal{A}:=\mathcal{A}_1\times\cdots\times\mathcal{A}_n$ と定義し，これを行動群集合とよぶ．また，すべてのエージェントの行動をまとめて $a=(a_1,\cdots,a_N)\in\mathcal{A}$ と表す．さらに，i 以外の行動群を

$$a_{-i}:=(a_1,...,a_{i-1},a_{i+1},...,a_n)$$

と表記する．このとき，行動群 a は $a=(a_i,a_{-i})$ と表現することができる．最後に，相対姿勢と同様に，行動の偏差を $a_{ij}=a_i-a_j$ と表記する．

本論文では，エージェントのハードウェア制約や障害物の存在ゆえに生じる可動範囲の制限を表現するために，エージェント i が行動 a_i を選択している際の実行可能な行動集合を拘束行動集合 $R_i(a_i)\subset\mathcal{A}_i$ と表記する．すなわち，各エージェント i は各時刻 k において行動集合 $R_i(a_i(k-1))=R_i(\theta_i(k))$ の中から行動 $a_i(k)$ を選択し，これが次の時刻 $k+1$ の i の姿勢 $\theta_i(k+1)$ となる．本論文を通して以下の仮定をおく．

（仮定 1） 拘束行動集合は，以下の条件を満たす．

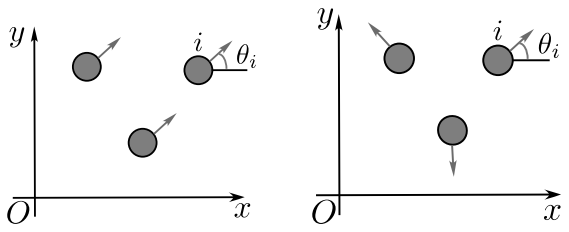

Fig. 1 Attitude coordination (Left: Synchronization, Right: Balanced circular formation)

可逆性: 任意の $i\in\mathcal{V}$ および行動 $a_i^1,a_i^2\in\mathcal{A}_i$ に対して $a_i^2\in R_i(a_i^1)\longleftrightarrow a_i^1\in R_i(a_i^2)$ が成立する．

可解性: 任意の $i\in\mathcal{V}$, 行動 $a_i^0,a_i^m\in\mathcal{A}_i$ に対して, $\exists(a_i^0\cdots a_i^m)$ s.t. $a_i^l\in R_i(a_i^{l-1})$, $l\in\{1,\cdots,m\}$ が成立する． □

可逆性はある時刻に姿勢を変化させた場合，それ以降の時刻において元の姿勢に戻ることができることを意味する．可解性は，$\mathcal{A}_i$ 内のいかなる姿勢も，一度の遷移では到達できなくとも数回の移動を繰り返せば到達できることを意味する．ハードウェア制限や障害物による行動制約を想定する限り，これらの仮定は一般に満たされる．

つぎに，エージェント間の通信構造をグラフ $G=(\mathcal{V},\mathcal{E})$ を用いて表現する．ここで，エージェント i の通信可能な近傍を $\mathcal{N}_i:=\{j\in\mathcal{V}|\ (i,j)\in\mathcal{E}\}$ と表記し，これに自身を加えた集合を $\tilde{\mathcal{N}}_i:=\mathcal{N}_i\cup\{i\}$ と定義する．本論文ではグラフ G に以下の仮定をおく．

（仮定 2） グラフ G は時不変，無向，連結である． □

本論文の文脈において，グラフが無向であるとはエージェント間に双方向通信を仮定することに相当し，グラフが連結であるとはすべてのエージェントが任意のエージェントまで枝を通って到達することができることを意味する．グラフに関する詳細は文献 [1] を参照されたい．

2.2 姿勢協調問題

本論文の目的は，以下の二つの姿勢配置を達成する制御アルゴリズムを提案することである．

【定義 1】（姿勢協調 [2]） エージェント群 $\mathcal{V}$ は

$$\theta_i=\theta_j\quad\forall i,j\in\mathcal{V} \tag{1}$$

を満たすとき，姿勢同期を達成するという．また，

$$\sum_{i=1}^{N}\cos\theta_i=0,\ \sum_{i=1}^{N}\sin\theta_i=0 \tag{2}$$

を満たすとき，姿勢平衡を達成するという． □

すなわち，姿勢同期とはすべてのエージェントの姿勢が一致することを指し，姿勢平衡とはエージェントの姿勢がつり合うことを意味する (Fig. 1)．

文献 [2] では，魚の群れ行動の模倣という観点から，ほかにも様々な姿勢協調が定義されているが，同期と平衡はその中でも代表的なものである．また，モバイルセンサネットワーク [3] やビジュアルセンサネットワーク [7] といった応用を想定すれば，これらの協調はとくに有

用性が高い．なお，姿勢の表現の相違から，**定義 1** は文献 [2] の定義とは異なるが，これらは本質的に等価である．

定義 1 の姿勢協調を達成する制御則はすでに複数報告されている（たとえば，文献 [2,3] 参照）．しかしながら，とくに姿勢平衡については，大域的に達成を保証する方法は報告されていない．同期のみに限れば大域的な収束を保証することは可能であるが，協調のパターンごとに方法論を変更するよりは，統一的な方法論が確立されることが望ましい．さらに，いずれの問題に関しても障害物などの行動制約のもとで協調を達成する方法は存在しない．

2.3 姿勢協調ゲーム

本論文では，姿勢協調問題をゲームとして捉える．この目的のために，これ以降，本論文では行動集合 $\mathcal{A}_i$ は有限集合とし，個々のエージェントは有限個の行動の候補の中から実際に実行する行動を選択する．これはゲーム理論的協調制御では一般的な緩和であり [9]，とくに工学的な適用対象を限定するものではない．なお，協調制御をゲームとして捉える枠組は文献 [9,14] に詳述されており，一般論に興味がある読者はそちらを参照されたい．

戦略型ゲームはエージェント $\mathcal{V}$，行動群集合 $\mathcal{A}$ および利得関数 $\{U_i\}_{i\in\mathcal{V}}$, $U_i:\mathcal{A}\to\mathcal{R}$ の組として定義される．個々のエージェントは自身の利得関数 U_i を最大化するように行動 $a_i(k)$ を選択する．なお，ゲーム理論の目的は，与えられた利得関数に対して Nash 均衡解 [9,14] とよばれる均衡解を求めることである．

前節において姿勢協調の $\mathcal{V}$，および $\mathcal{A}$ は定義済みであるので，設計すべきは利得関数 $\{U_i\}_{i\in\mathcal{V}}$ およびそれを利用した $a_i(k)$ の決定アルゴリズムである [14]．利得関数 U_i の設計に際して注意すべき点は以下のことである: (i) 姿勢協調問題では通信がグラフ G によって制限されるため，$U_i(a)$ は常に近傍からの通信情報を用いて計算できる必要がある．(ii) (1) 式および (2) 式を満足する行動群が均衡解となる $U_i(a)$ を設計する必要がある．

まとめると，本論文の目的は行動制約のもとで大域的に姿勢協調を達成する利得関数 $\{U_i\}_{i\in\mathcal{V}}$ および $a_i(k)$ の決定アルゴリズムを設計することである．

3. 姿勢協調ポテンシャルゲーム

3.1 ポテンシャルゲーム

本論文では，とくに次で定義されるゲームに着目する．

【定義 2】（ポテンシャルゲーム [9,10]） エージェント集合 $\mathcal{V}$，行動群集合 $\mathcal{A}$ および利得関数 $\{U_i\}_{i\in\mathcal{V}}$, $U_i:\mathcal{A}\to\mathcal{R}$ からなるゲーム Γ を考える．このとき，ある関数 ϕ が存在して任意の $i\in\mathcal{V}$，$a_i',a_i''\in\mathcal{A}_i$ に対して，

$$U_i(a_i'',a_{-i})-U_i(a_i',a_{-i})=\phi(a_i'',a_{-i})-\phi(a_i',a_{-i}) \quad (3)$$

を満たすとき，Γ はポテンシャルゲームを構成する． □

ポテンシャルゲームは Nash 均衡解が容易に計算できるという点で扱いやすいクラスの問題である [10]．ポテンシャル関数 ϕ はグループとしての利得と解釈されることが多い．この意味で，ポテンシャルゲームにおいては，エージェント i のみが行動を変化させたとき，利得関数の増加がグループとしての利得の増加を意味する．

文献 [9] では，ポテンシャル関数としてそれを最大化する行動群が協調制御の目的と一致するような関数を構成することで，多くの協調制御をポテンシャルゲームとして解釈することができることを示している．

3.2 ポテンシャル関数

本節ではまず，姿勢協調ゲームにおけるポテンシャル関数を定義する．前節の通り，ポテンシャル関数はその最大値を与える行動群 a が，(1) 式あるいは (2) 式を達成するものでなければならない．

この目的のために，ラプラシアン姿勢ポテンシャル [2]

$$W(\theta)=\frac{1}{2N}\sum_{i=1}^{N}\Big(d_i-\sum_{j\in\mathcal{N}_i}\cos\theta_{ij}\Big)$$

を導入する．ここで，d_i は近傍 $\mathcal{N}_i$ の要素数である．以降，表記の簡単化のために $D:=\sum_{i\in\mathcal{V}}d_i$ とおく．ポテンシャル $W(\theta)$ は以下の性質を満足する [2]．

- $W(\theta)$ が最小のときエージェント群は同期する．
- G が環状グラフ ($\mathcal{N}_1=\{N,2\}$, $\mathcal{N}_N=\{1,N-1\}$, $\mathcal{N}_i=\{i-1,i+1\}$, $i\neq\{1,N\}$) であれば，$W(\theta)$ が最大のとき，エージェント群は姿勢平衡を達成する．

以上の性質から，本論文では同期，平衡のポテンシャル関数 $\phi_{\rm s}$, $\phi_{\rm b}$ をそれぞれ

$$\phi_{\rm s}(a)=-W(a) \quad (4)$$
$$\phi_{\rm b}(a)=W(a) \quad (5)$$

と定義する．なお，以降の議論には必要としないためその仮定は明記しないが，姿勢平衡を考える場合には常に環状グラフを想定するものとする．

3.3 利得関数の設計

本節では，姿勢協調問題に対する利得関数を設計し，結果として与えられるゲームがポテンシャルゲームを構成することを証明する．

【定理 1】 エージェント群 $\mathcal{V}$，行動集合 $\{\mathcal{A}_i\}_{i=1}^{N}$ に加えて利得関数を

$$U_i(a)=\frac{1}{N}\sum_{j\in\mathcal{N}_i}\cos a_{ij} \quad (6)$$

とする姿勢同期ゲーム $\Gamma_{\rm s}$ を考える．このとき仮定 2 のもとで，ゲーム $\Gamma_{\rm s}$ はポテンシャル関数 (4) に対するポテンシャルゲームを構成する．また，利得関数を

$$U_i(a)=-\frac{1}{N}\sum_{j\in\mathcal{N}_i}\cos a_{ij} \quad (7)$$

とする姿勢平衡ゲーム $\Gamma_{\rm b}$ を考える．このとき仮定 2 のもとで，ゲーム $\Gamma_{\rm b}$ はポテンシャル関数 (5) に対するポテンシャルゲームを構成する． □

（証明）仮定 2 のもとで，ポテンシャル関数は

$$\phi_{\rm s}(\theta)=-\frac{1}{2N}\Big(D-\sum_{j\in\mathcal{N}_i}2\cos a_{ij}-\sum_{j\neq i}\sum_{l\in\mathcal{N}_j/\{i\}}\cos a_{jl}\Big)$$

と書き換えることができる．行動を $a_i^{'}$ から $a_i^{''}$ に変化させたときのポテンシャル関数の変化は，

$$\begin{aligned}&\phi_{\rm s}(a_i^{''},a_{-i})-\phi_{\rm s}(a_i^{'},a_{-i})\\&=\frac{1}{2N}\sum_{j\in\mathcal{N}_i}2\cos(a_i^{''}-a_j)-\frac{1}{2N}\sum_{j\in\mathcal{N}_i}2\cos(a_i^{'}-a_j)\\&=U_i(a_i^{''},a_{-i})-U_i(a_i^{'},a_{-i})\end{aligned}$$

となる．ポテンシャルゲームの定義 (3) 式より，定理が示される．ゲーム $\Gamma_{\rm b}$ についても同様に証明できる． □

明らかに，(6) 式および (7) 式は $\{a_j\}_{j\in\tilde{\mathcal{N}}_i}$ のみから計算できる．また，上述の通り，ポテンシャル関数 $\phi_{\rm s}$ および $\phi_{\rm b}$ の最大点はそれぞれ姿勢同期，姿勢平衡を達成する．さらに，ポテンシャル関数を最大化する行動群は Nash 均衡解となる [9] ので，姿勢同期ゲーム $\Gamma_{\rm s}$，姿勢平衡ゲーム $\Gamma_{\rm b}$ は **2.3** の要求 (i)，(ii) を同時に満足する．ただし，均衡解は必ずしもポテンシャル関数を最大にしない．姿勢平衡の場合には，たとえば $N=3$ における $a_1=a_2=0$, $a_3=\pi$ など，ポテンシャル関数を最大化しない均衡解の存在を容易に確認できる．また姿勢同期に関しても，行動制約が存在する場合には，[9] の合意問題と同様にそのような均衡解が存在する可能性がある．

以上で，姿勢協調ゲームがともにポテンシャルゲームを構成することが示された．文献 [9] ではポテンシャル関数を最大化する特定の均衡解に到達させるための学習アルゴリズム RSAP を提案している．これを用いれば，十分時間が経過したのち，各エージェントは初期状態によらず，行動制約の存在下においてポテンシャル関数を最大化する行動群を高確率で選択することが保証される．

しかしながら，RSAP は一回の反復で 1 台のエージェントしか行動を変化できないため，とくに N が大きい場合には，後述の各行動の確率分布の所望の分布への収束に時間を要する．そこで，同時に複数のエージェントの行動を変更することで学習アルゴリズムの高速化を図る．ただし，ポテンシャルゲームに対しても，複数エージェントによる利得を増加させる行動変化は必ずしもポテンシャル関数を増加させない．そこで，次節では群ポテンシャルゲーム [9] の概念を導入し，姿勢協調ゲームがこれを構成することを示したのち，この性質を利用した学習アルゴリズムを提案する．

（注意 1）関数 (4)，(5) は正負を逆転させるだけで二つの姿勢協調問題を統一的に扱えるという利点をもつ．ただし，ポテンシャル関数の採り方は一意ではない．たとえば，姿勢同期問題を合意問題と同一視すると，$\phi_{\rm s}=-(1/2N)\sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{N}_i}\|a_i-a_j\|^2$，$U_i(a)=\sum_{j\in\mathcal{N}_i}\|a_i-a_j\|^2$ はポテンシャルゲームを構成する．姿勢平衡に関しても，$2\pi/N$ なるバイアスを設けて同様のノルムを用いればポテンシャルゲームを構成できる．ただし，この場合，利得関数中に N が表れ，個々のエージェントが全エージェント数を常に把握する必要が生じる．

4. 学習アルゴリズムの高速化

本節では学習アルゴリズムの収束を加速するため，以下に示す群ポテンシャルゲーム [9] の概念を用いる．

4.1 群ポテンシャルゲーム

群ポテンシャルゲームでは，エージェント群 $\mathcal{V}$ および行動集合 $\mathcal{A}_i$ に加えて $\mathcal{V}$ の部分集合族 $\mathcal{V}_i$, $i=1,\cdots,m$ の存在を仮定し，各集合 $\mathcal{V}_i$ に利得関数を対応させる．すなわち，利得関数を $U_{\mathcal{V}_i}\colon\mathcal{A}\to\mathcal{R}$ と表記する．

【定義 3】（群ポテンシャルゲーム [9]） エージェント群 $\mathcal{V}$, 行動集合 $\mathcal{A}_i$, 集合族 $\mathcal{V}_i$, $i=1,\cdots,m$ および $\{U_{\mathcal{V}_i}\}_{i=1}^m$ から構成されるゲーム Γ を考える．任意の $\mathcal{V}_i$, $a'_{\mathcal{V}_i}$, $a''_{\mathcal{V}_i}\in\prod_{i\in\mathcal{V}_i}\mathcal{A}_i$, $a_{-\mathcal{V}_i}=(a_j)_{j\notin\mathcal{V}_i}\in\prod_{i\notin\mathcal{V}_i}\mathcal{A}_i$ に対して

$$\begin{aligned}&U_{\mathcal{V}_i}(a''_{\mathcal{V}_i},a_{-\mathcal{V}_i})-U_{\mathcal{V}_i}(a'_{\mathcal{V}_i},a_{-\mathcal{V}_i})\\&\qquad=\phi(a''_{\mathcal{V}_i},a_{-\mathcal{V}_i})-\phi(a'_{\mathcal{V}_i},a_{-\mathcal{V}_i})\end{aligned}\qquad(8)$$

を満たす関数 $\phi\colon\mathcal{A}\to\mathbb{R}$ が存在するとき，ゲーム Γ は，群ポテンシャルゲームを構成するという． □

(3) 式はポテンシャル関数の増分と個々の利得の増分が等しいことを意味するのに対して，(8) 式は群の利得変化と等しくなる．次節以降，この性質を利用して群ごとに行動を変化させるアルゴリズムを提案する．

4.2 群利得関数の設計

本論文では，集合 $\mathcal{V}_i$ を $\tilde{\mathcal{N}}_i$ とする．このとき，つぎの定理が成り立つ．

【定理 2】 エージェント群 $\mathcal{V}$，行動集合 $\{\mathcal{A}_i\}_{i=1}^N$，集合族 $\{\tilde{\mathcal{N}}_i\}_{i=1}^N$ に加えて，群利得関数を

$$\begin{aligned}U_{\tilde{\mathcal{N}}_i}(a)=&\frac{1}{N}\sum_{j\in\mathcal{N}_i}\cos a_{ij}+\frac{1}{N}\sum_{j\in\mathcal{N}_i}\sum_{r\in\mathcal{N}_j/\mathcal{N}_i}\cos a_{jr}\\&+\frac{1}{2N}\sum_{j\in\mathcal{N}_i}\sum_{l\in\mathcal{N}_i\cap\mathcal{N}_j}\cos a_{jl}\end{aligned}$$

とする姿勢同期群ゲーム $\Gamma_{\rm gs}$ を考える．このとき仮定 2 のもとで，ゲーム $\Gamma_{\rm gs}$ はポテンシャル関数 (4) に対する群ポテンシャルゲームを構成する．また，群利得関数を

$$\begin{aligned}U_{\tilde{\mathcal{N}}_i}(a)=&-\frac{1}{N}\sum_{j\in\mathcal{N}_i}\cos a_{ij}-\frac{1}{N}\sum_{j\in\mathcal{N}_i}\sum_{r\in\mathcal{N}_j/\mathcal{N}_i}\cos a_{jr}\\&-\frac{1}{2N}\sum_{j\in\mathcal{N}_i}\sum_{l\in\mathcal{N}_i\cap\mathcal{N}_j}\cos a_{jl}\end{aligned}$$

とする姿勢平衡群ゲーム $\Gamma_{\rm gb}$ はポテンシャル関数 (5) に

対する群ポテンシャルゲームを構成する． □

（証明）以降 $U_{\tilde{\mathcal{N}}_i}$ を簡単に U_i と表記する．仮定 2 より，ポテンシャル関数は，

$$\phi_{\mathrm{s}}(\theta)=-\frac{1}{2N}\Big(D-\sum_{j\in\mathcal{N}_i}2\cos a_{ij}-\sum_{j\in\mathcal{N}_i}\sum_{r\in\mathcal{N}_j/\mathcal{N}_i}2\cos a_{jr}-\sum_{j\in\mathcal{N}_i}\sum_{l\in\mathcal{N}_i\cap\mathcal{N}_j}\cos a_{jl}-\sum_{r\in N/\mathcal{N}_i}\sum_{m\in\mathcal{N}_k/\mathcal{N}_i}\cos a_{rm}\Big)$$

と表現できる．$\mathcal{V}_i$ 内のエージェント群が，行動集合を $a'_{\tilde{\mathcal{N}}_i}$ から $a''_{\tilde{\mathcal{N}}_i}$ に変え，ほかのエージェントが行動 $a_{-\tilde{\mathcal{N}}_i}$ を行うとき，ポテンシャル関数の変化は，

$$\begin{aligned}&\phi(a''_{\tilde{\mathcal{N}}_i},a_{-\tilde{\mathcal{N}}_i})-\phi(a'_{\tilde{\mathcal{N}}_i},a_{-\tilde{\mathcal{N}}_i})\\&=\frac{1}{2N}\sum_{j\in\mathcal{N}_i}\Big(2\cos a''_{ij}+\sum_{r\in\mathcal{N}_j/\mathcal{N}_i}2\cos(a''_j-a_r)\\&+\sum_{l\in\mathcal{N}_i\cap\mathcal{N}_j}\cos a''_{jl}-2\cos a'_{ij}\\&-\sum_{r\in\mathcal{N}_j/\mathcal{N}_i}\cos(a'_j-a_r)-\sum_{l\in\mathcal{N}_i\cap\mathcal{N}_j}\cos a'_{jl}\Big)\\&=U_i(a''_{\tilde{\mathcal{N}}_i},a_{-\tilde{\mathcal{N}}_i})-U_i(a'_{\tilde{\mathcal{N}}_i},a_{-\tilde{\mathcal{N}}_i})\end{aligned}$$

となる．よって，定義 (8) 式から定理が証明される．姿勢平衡についても同様に証明できる． □

以上より，姿勢協調問題は群ポテンシャルゲームを構成する．すなわち，群 $\tilde{\mathcal{N}}_i$ が U_i を増加させる行動をとれば，ポテンシャル関数も増加する．次節ではこの性質を利用して，群全体が同時に行動を変化し，かつ (4) 式，(5) 式を高確率で達成する学習アルゴリズムを提案する．

4.3 学習アルゴリズムの提案

本節では，姿勢協調群ポテンシャルゲームに対する行動決定アルゴリズム Algorithm 1 を提案する．

Algorithm 1 の概要を以下に示す．まず，一様分布に従って無作為に一つのエージェント $i\in\mathcal{V}$ を選択する．エージェント i は U_i の計算に必要な情報を近傍から収集したのち，試行行動 $\hat{a}_{\tilde{\mathcal{N}}_i}$ を以下の確率にしたがって選択する．ここで，$\tilde{\mathcal{R}}_i(a_{\tilde{\mathcal{N}}_i}):=\prod_{j\in\tilde{\mathcal{N}}_i}R_j(a_j)\setminus\{a_j\}$ である．

$$\Pr[\hat{a}_{\tilde{\mathcal{N}}_i}=a_{\tilde{\mathcal{N}}_i}]=1/\tilde{z}_i\quad\text{if } a_{\tilde{\mathcal{N}}_i}\in\tilde{\mathcal{R}}_i(a_{\tilde{\mathcal{N}}_i}(k-1)) \tag{9}$$

$$\Pr[\hat{a}_{\tilde{\mathcal{N}}_i}=a_i(k-1)]=1-|\tilde{\mathcal{R}}_i(a_{\tilde{\mathcal{N}}_i}(k-1))|/\tilde{z}_i \tag{10}$$

自然数 $\tilde{z}_i$ は，障害物がない場合の選択可能な行動の数 z_i を用いて $\tilde{z}_i=\prod_{j\in\tilde{\mathcal{N}}_i}z_i$ と定義される．すなわち，ハードウェア制約に起因する制限がなく，自由に姿勢を変更できる状況下では $z_i=|\mathcal{A}_i|$ である．

最後に，試行行動 $\hat{a}_{\tilde{\mathcal{N}}_i}$ を実際に実行するか否かを以下の確率にしたがって決定する．

$$\Pr[a_{\tilde{\mathcal{N}}_i}(k)=\hat{a}_{\tilde{\mathcal{N}}_i}]=\frac{\exp\{\beta U_i(\hat{a}_{\tilde{\mathcal{N}}_i},a_{-\tilde{\mathcal{N}}_i}(k-1))\}}{\exp\{\beta U_i(\hat{a}_{\tilde{\mathcal{N}}_i},a_{-\tilde{\mathcal{N}}_i}(k-1))\}+\exp\{\beta U_i(a(k-1))\}} \tag{11}$$

$$\Pr[a_{\tilde{\mathcal{N}}_i}(k)=a_{\tilde{\mathcal{N}}_i}(k-1)]=\frac{\exp\{\beta U_i(a(k-1))\}}{\exp\{\beta U_i(\hat{a}_{\tilde{\mathcal{N}}_i},a_{-\tilde{\mathcal{N}}_i}(k-1))\}+\exp\{\beta U_i(a(k-1))\}} \tag{12}$$

Algorithm 1 :Restrictive Spacial Adaptive Play with Neighbor-based Decisions

初期状態: $k:=0$

1: ランダムに一つのエージェント $i\in\mathcal{V}$ を選択する．エージェント i は，選ばれた事実を近傍のエージェント $\mathcal{N}_i$ に伝える．

2: 近傍のエージェント $j\in\mathcal{N}_i$ は，その近傍 $\mathcal{N}_j$ の情報 $\{a_l(k-1)\}_{l\in\tilde{\mathcal{N}}_j}$ を集め，エージェント i に伝える．

3: エージェント i は，一つの試行行動集合 $\hat{a}_{\tilde{\mathcal{N}}_i}$ を確率 (9)，(10) に従って選択する．

4: エージェント i は，近傍からの情報をもとに群目的関数を計算する．

5: エージェント i は，確率 (11)，(12) に従って行動群 $a_{\tilde{\mathcal{N}}_i}(k)$ を選択する．

6: エージェント i は近傍に $a_{\tilde{\mathcal{N}}_i}(k)$ を伝える．

7: $j\in\tilde{\mathcal{N}}_i$ は行動 $a_{\tilde{\mathcal{N}}_i}(k)$ を実行する．

8: $k\leftarrow k+1$ として Step 1 へ．

(11) 式は，試行行動 $\hat{a}_{\tilde{\mathcal{N}}_i}$ のグループ利得が現在の行動と比べて大きければ試行行動を優先的に実施することを意味する．パラメータ $\beta>0$ はその優先度を指定し，これが大きければ利得が大きい行動の優先度が増大する．

(11) 式において利得関数が現在よりも劣る行動をとる確率が 0 ではないことに注意されたい．これにより，行動群が ϕ の局所的最大点に陥った場合にそこから脱け出すことができるが，同様に大域的最大点，すなわち姿勢協調の達成状態からも脱する．β を小さく設定すれば，局所最大点に留まる時間が短縮される一方で姿勢協調状態から脱する確率も高まり，逆に大きく設定すれば姿勢協調状態に留まる確率は高まるものの，到達時間は増大するというトレードオフの関係がある．これは，探索と知識利用のジレンマ [15] とよばれる現象であり，RSAP を含め，多くの学習アルゴリズムが内包する問題である．

4.4 収束性解析

本節では，Algorithm 1 の収束性解析を行う．RSAP と同様に，Algorithm 1 において行動群自体は収束しないものの各行動群がとられる確率分布は時間の経過とともに定常分布に収束する．

まず，集合 $\mathcal{A}$ 内の行動群に順番をつけて $a^{(1)},\cdots,a^{(|\mathcal{A}|)}$ と表現する．つぎに，時刻 k において行動群 $a^{(i)}$ が実行される確率を $\mu_i(k)$ と表現し，$\mu_1(k),\cdots,\mu_{|\mathcal{A}|}(k)$ を横に並べた行ベクトルを $\mu(k)$ と表記する．このとき，$\mu(k)$ の遷移は次式のマルコフ連鎖として表現される．

$$\mu(k+1)=\mu(k)P \tag{13}$$

ここで，行列 P は遷移行列を表し，その (i,j)-成分 P_{ij}

は $\Pr[a(k+1)=a^{(j)}|\ a(k)=a^{(i)}]$ である．

Algorithm 1 を用いたとき，$\mu=\mu P$ を満足する $\mu(k)$ の定常分布 μ に関して以下の定理が成立する．

【定理 3】 仮定 1，2 を満足する姿勢協調群ポテンシャルゲームを考える．このとき，Algorithm 1 を用いるならば，$\mu(k)$ の定常分布 μ は次式で与えられる．

$$\mu_i=\frac{\exp\{\beta\phi(a^{(i)})\}}{\sum_{\bar{a}\in A}\exp\{\beta\phi(\bar{a})\}} \tag{14}$$

□

（証明）仮定 1 より，Algorithm 1 により導出されるマルコフ連鎖 (13) は既約かつ非周期的であるので唯一の定常分布が存在する [16]．よって，(14) 式で与えられる μ が詳細つり合い条件

$$\mu_j P_{jl}=\mu_l P_{lj}\quad \forall j,l \tag{15}$$

を満たすことを証明すれば十分である．

行動群 $a^{(j)}$ と $a^{(l)}$ がある i に対して $\tilde{\mathcal{N}}_i$ のみ異なる，つまり $a^{(j)}_{-\tilde{\mathcal{N}}_i}=a^{(l)}_{-\tilde{\mathcal{N}}_i}$ かつ $a^{(j)}_{\tilde{\mathcal{N}}_i}\neq a^{(l)}_{\tilde{\mathcal{N}}_i}$ の場合を考える．そのような i が存在しなければ $P_{jl}=P_{lj}=0$ である．Step 1 においてエージェント i が選ばれる確率が $1/n$ であり，(9) 式において試行行動として $a^{(l)}_{\tilde{\mathcal{N}}_i}$ が選択される確率は $1/\tilde{z}_i$ である．よって，遷移確率 P_{jl} は

$$P_{jl}=\Pr[a(k)=a^{(l)}|a(k-1)=a^{(j)}]=\sum_{i\in\mathcal{V}}\mathrm{sgn}\{A\} \tag{16}$$

$$A:=\frac{1}{n\tilde{z}_i}\frac{\exp\{\beta U_i(\hat{a}_{\tilde{\mathcal{N}}_i},a^{(j)}_{-\tilde{\mathcal{N}}_i})\}}{\exp\{\beta U_i(\hat{a}_{\tilde{\mathcal{N}}_i},a^{(j)}_{-\tilde{\mathcal{N}}_i})\}+\exp\{\beta U_i(a^{(j)})\}}$$

となる．$\mathrm{sgn}(x)$ は，もし $(\hat{a}_{\tilde{\mathcal{N}}_i},a^{(j)}_{-\tilde{\mathcal{N}}_i})=a^{(l)}$ ならば $\mathrm{sgn}(x)=x$ となり，そうでなければ 0 となる関数である．ここで，

$$\lambda=\frac{1}{n}\frac{1}{\sum_{\bar{a}\in\mathcal{A}}\exp\{\beta\phi(\bar{a})\}}\times\frac{1}{\exp\{\beta U_i(\hat{a}_{\tilde{\mathcal{N}}_i},a^{(j)}_{-\tilde{\mathcal{N}}_i})\}+\exp\{\beta U_i(a^{(j)})\}}$$

を定義すると，(14) 式および (16) 式から，

$$\mu_j P_{jl}=\lambda\sum_{i\in\mathcal{V}}\mathrm{sgn}\Big\{\frac{1}{\tilde{z}_i}\exp(\beta(U_i(a^{(l)})+\phi(a^{(j)})))\Big\}$$

を得る．(8) 式より，$U_i(a^{(l)})+\phi(a^{(j)})=U_i(a^{(j)})+\phi(a^{(l)})$ となるので，次式を得る．

$$\begin{aligned}\mu_j P_{jl}&=\lambda\sum_{i\in N}\mathrm{sgn}\Big\{\frac{1}{\tilde{z}_i}\exp(\beta(U_i(a^{(j)})+\phi(a^{(l)})))\Big\}\\&=\mu_l P_{lj}\end{aligned}$$

よって，条件 (15) が成立し，定理は証明される． □

添字 $i=1,\cdots,|\mathcal{A}|$ のうち，対応する $a^{(i)}$ が姿勢協調を達成する集合を $\mathcal{I}_{AC}$ とおくと，時刻 k における姿勢協調の達成確率は $\sum_{i\in\mathcal{I}_{AC}}\mu_i(k)$ と定義される．**定理 3** は，Algorithm 1 を用いることで，十分時間が経過したのち，その確率が (14) 式で与えられる μ によって $\sum_{i\in\mathcal{I}_{AC}}\mu_i$ で近似できることを意味する．(14) 式において，分母が定数であることに注意すると，ポテンシャル関数 $\phi(a^{(i)})$ が大きい行動に対応する μ_i が大きな値をもつことがわかる．ここで，$\beta\to\infty$ とすれば，定常分布 μ のすべての重みはポテンシャル関数を最大化する行動群におかれるが，β は **4.3** 末尾のジレンマを考慮して，一定の値よりも小さく設定しなければ現実には機能しない．

姿勢協調はポテンシャル関数の最大値を与えること，(14) 式が初期状態に依存しないこと，定理は行動制約の存在を考慮している点に注意すると，**定理 3** は，任意の初期状態に対して，十分時間が経過したのち，行動制約のもとでの姿勢協調を高確率で達成することを意味する．すなわち，高確率という緩和はあるものの，本論文の目的はおおよそ達成される．

定理 3 は文献 [9] の結論と同じである．つまり，Algorithm 1 の利点は各反復において複数のエージェントの行動を変更するにもかかわらず，RSAP と同じ定常分布を得ることである．これにより，確率分布の収束時間の短縮が期待できる．ただし，収束時間に関する厳密な評価は今後の課題とし，本論文ではシミュレーションによって提案アルゴリズムの有効性を検証するに留める．

5. シミュレーション

本節では，数値シミュレーションを通して本論文の枠組みと提案アルゴリズムの有効性を検証する．

5.1 姿勢平衡

本節では，姿勢平衡問題に対するシミュレーション結果を示す．行動集合 $\mathcal{A}_i=\{2\pi l/5\}_{l=0}^{4},\ i=1,2,3,4,5$ をもつ 5 台のロボットを考える．また，各エージェントは 1 ステップに $\pm 2\pi/5$ しか動くことができないという行動拘束を加える ($z_i=3$)．通信構造を表すグラフは環状グラフとする．また，すべてのエージェントの初期姿勢は，大域的な収束を保証しない従来法では姿勢平衡を達成できないように，0 と設定した．

まず，RSAP を用いた際の姿勢の時間応答を示す．パラメータ β は，ポテンシャル関数を最大にする配置から移動する確率が 0.025 以下になるように，(15) 式から逆算して $\beta=14.5$ とする．設定の詳細は [8] を参照されたい．Fig. 2 に各エージェントの姿勢，Fig. 3 にポテンシャル関数の応答を示す．Fig. 2 の網掛けは姿勢平衡が達成されている時間区間を示す．図より，60 ステップ近辺において最初の平衡状態に達し，80 ステップ辺りでその状態を脱する．その後，探索行動を繰り返しながら，160, 220，280 ステップ近辺で再び姿勢平衡に達する．

これらの時刻における姿勢平衡のパターンは必ずしもほかの時刻のものとは等しくない．すなわち，探索行動のために一時的にはポテンシャル関数に関して不利な行動が選択されるため，大域的な姿勢平衡への到達と異な

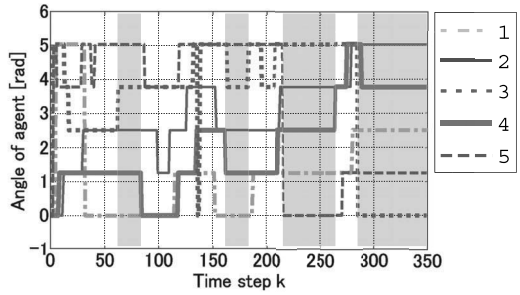

Fig. 2 Time responses of orientations with RSAP (Balanced circular formation)

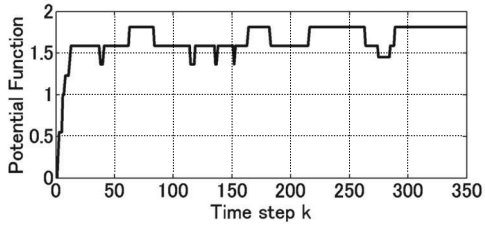

Fig. 3 Time response of potential function with RSAP (Balanced circular formation)

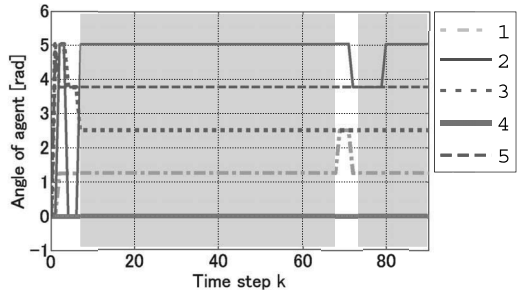

Fig. 4 Time responses of orientations with Algorithm 1 (Balanced circular formation)

る平衡パターン間の移動が実現される．平衡パターン間の移動は役割分担とみなせば耐故障性の観点から興味深い挙動である．ただし，その代償として，これらの時間区間以外の時刻に見られるとおり，姿勢平衡が達成されない時間区間の存在を許容する必要がある．

Figs. 4, 5 に Algorithm 1 を用いたときのシミュレーション結果を示す．ここでは，ポテンシャル関数を最大にする配置から移動する確率が 0.025 以下になるように $\beta = 21.2$ と設定した．Figs. 3, 5 における x 軸のスケールの違いに注意すると，Algorithm 1 では 7 ステップにおいてすでに最初の平衡状態に達しており，平衡状態への到達速度が大幅に改善されていることがわかる．また，この図のみから判断するのは正確性を欠くが，開始直後に RSAP に見られる揺動が大幅に抑制されている事実から，分布の収束自体も加速していることが予想される．

5.2 姿勢同期

最後に，ビジュアルセンサネットワークにおける姿勢制御を想定して，障害物が存在する状況における姿勢同期の達成を検証する．ここでは，4 台のカメラが

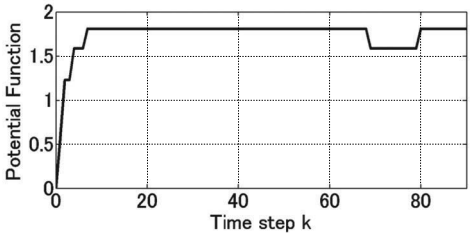

Fig. 5 Time response of potential function with Algorithm 1(Balanced circular formation)

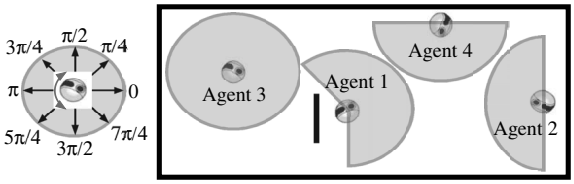

Fig. 6 Example of environment

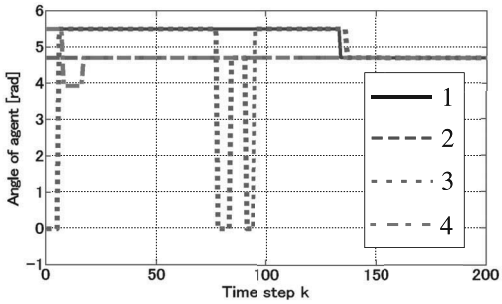

Fig. 7 Time response of orientations with RSAP (Synchronization)

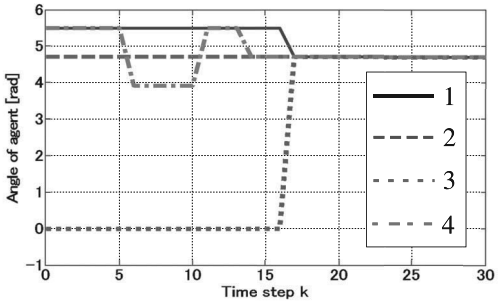

Fig. 8 Time responses of orientations with Algorithm 1 (Syncronization)

存在する状況を考え，Fig. 6 左図のように行動集合を $\mathcal{A}_i = \{l\pi/4\}_{l=0}^{7}$, $i = 1,2,3,4$ と設定する．ハードウェア制約により，各ロボットは 1 ステップに $\pm\pi/4$, $\pm\pi/2$ までしか動けないものとする ($z_i = 5$)．ここで，壁や障害物の影響により，各カメラの姿勢は Fig. 6 右図のように拘束されるものとする．たとえば，

$$R(a_2) = \{0,\ \pi/4,\ \pi/2,\ 3\pi/4,\ 3\pi/2,\ 7\pi/4\}\ \forall a_2 \in \mathcal{A}_2$$

である．当然，従来法ではこのような問題は扱えない．なお，本シミュレーションでは通信構造を $\mathcal{E} = \{(1,2),(1,3),(2,4)\}$ と仮定し，エージェント 1,2,3,4 の初期状態をそ

れぞれ $7\pi/4$, $3\pi/2$, 0, $7\pi/4$ [rad] と設定する．

Fig. 7 に RSAP ($\beta = 45.01$) を用いた場合の，Fig. 8 に Algorithm 1 ($\beta = 74.4$) を適用した場合の姿勢の時間応答を示す．図より，両手法とも最終的に同期が達成されていることが確認できる．よって，提案した枠組みが障害物が存在する状況においても十分機能することが確認できる．最後に，両アルゴリズムの同期達成時刻を比較すると Algorithm 1 は 18 ステップ目で同期を達成するのに対して，RSAP は達成までに約 140 ステップ要している．このことから，提案アルゴリズムが到達を加速していることが確認できる．さらに，**5.1** 同様，開始直後の揺動の抑制から分布の収束の加速が予想される．

6. おわりに

本論文では，姿勢協調問題にポテンシャルゲームを適用した．まず，適切なポテンシャル関数と目的関数を用いることにより，姿勢協調問題をポテンシャルゲームとして表現できることを示した．さらに，姿勢協調問題が群ポテンシャルゲームを構成することを示し，複数のエージェントを同時に動かすことで，定常分布に収束する速さを加速するアルゴリズムを提案した．最後に，本論文の有効性および妥当性をシミュレーションにより検証した．本論文においては割愛したが，検証実験については文献 [8] を参照されたい．

参 考 文 献

[1] F. Bullo, J. Cortes and S. Martinez : *Distributed Control of Robotic Networks*, Princeton University Press (2009)

[2] D. A. Paley, N. E. Leonard, R. Sepulchre, D. Grunbaum and J. K. Parrish: Oscillator models and collective motion; *IEEE Control Systems Magazine*, Vol. 27, No. 4, pp. 89–105 (2007)

[3] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni and R. E. Davis: Collective motion, sensor networks, and ocean sampling; *Proc. of the IEEE*, Vol. 95, No. 1, pp. 48–74 (2007)

[4] Y. Igarashi, T. Hatanaka, M. Fujita and M. W. Spong: Passivity-based attitude synchronization in SE(3); *IEEE Transactions on Control Systems Technology*, Vol. 17, No. 5, pp. 1119–1134 (2009)

[5] 吉村, 東, 杉江：マルチエージェントシステムのブロードキャスト制御；第 39 回 SICE 制御理論シンポジウム, pp. 249–252 (2010)

[6] 根，福島，松野：予測時刻間の衝突回避を考慮した複数移動体のモデル予測編隊制御；計測自動制御学会論文集, Vol. 46, No. 7, pp. 383–390 (2010)

[7] H. Aghajan and A. Cavallaro (Eds.): *Multi-Camera Networks: Principles and Applications*, Academic Press (2009)

[8] T. Goto, T. Hatanaka and M. Fujita: Potential game theoretic attitude coordination on the circle: Synchronization and balanced circular formation; *Proc. of the 2010 IEEE Multi-conference on Systems and Control*, pp. 2314–2319 (2010)

[9] J. R. Marden, G. Arslan and J. S. Shamma: Cooperative control and potential games; *IEEE Transacions on Systems, Man and Cybernetics, Part B: Cybernetics*, Vol. 39, No. 6, pp. 1393–1407 (2009)

[10] D. Monderer and L. Shapley: Potential games; *Games and Economic Behavior*, Vol. 14, No. 1, pp. 124–143 (1996)

[11] M. Zhu and S. Martinez: Distributed coverage games for mobile visual sensors (i): Reaching the set of nash equilibria; *Proc. of the 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, pp. 169–174 (2009)

[12] 寺岡，潮，金澤，林：ボロノイ分割による被覆制御へのポテンシャルゲームの応用; 電子情報通信学会技術報告, Vol. 110, No. 89, pp. 43–46 (2010)

[13] M. Saito, T. Hatanaka and M. Fujita: Decision dynamics in cooperative search based on evolutionary game theory; *Communications in Information and Systems*, Vol. 11, No. 1, pp. 57–70 (2011)

[14] R. Gopalakrishnan, J. R. Marden and A. Wierman: An architectural view of game theoretic control; *Proc. of ACM Hotmetrics 2010*, Third Workshop on Hot Topics in Measurement and Modeling of Computer Systems (2010)

[15] R. Bogacz, E. Brown, J. Moehlis, P. Holmes and J. D. Cohen: The physics of optimal decision making: A formal analysis of performance in two-alternative forced choice tasks; *Psychological Review*, Vol. 113, No. 4, pp. 700–765 (2006)

[16] C. M. Grinstead and J. L. Snell: *Introduction to Probability* (2nd Edition), American Mathematical Society (1997)

著 者 略 歴

畑中 健志 (正会員)

本論文誌 p. 164 参照

後藤 達彦

2009 年東京工業大学工学部制御システム工学科卒業，2011 年同大学大学院機械制御システム専攻博士前期課程修了．同年 4 月 (株) 東芝に入社，現在に至る．ゲーム理論を用いた協調制御の研究に従事．

藤田 政之 (正会員)

本論文誌 p. 164 参照