

ロボカップサッカーを題材とした  
人間－ロボット協調システムに関する研究

九州工業大学大学院  
生命体工学研究科 生命体工学専攻  
博士論文

指導教員: 石井和男

富永萌子  
2021年3月

© 2020  
富永萌子

審查委員会

主査: 石井和男

副査: 我妻先生

副査: 林英治先生

## **謝辞**

本研究を遂行するにあたり、熱心なご指導と有益なご助言を賜りますとともに、研究者としての心得を御教授してくださいました九州工業大学大学院生命体工学研究科人間知能システム工学専攻 石井和男教授に心から感謝します。

# 目 次

謝辞	iv
<b>第 1 章 序論</b>	<b>3</b>
1.1 緒言 . . . . .	3
1.1.1 ロボットの社会実装 . . . . .	3
1.1.2 協調行動 . . . . .	4
1.1.3 ロボット学習 . . . . .	4
1.1.4 人間ロボット協調 . . . . .	5
1.1.5 ロボカップ . . . . .	6
1.2 本研究の目的 . . . . .	9
1.3 本論文の構成 . . . . .	10
1.4 変数, 記号の約束 . . . . .	11
<b>第 2 章 人間のチーム行動解析</b>	<b>13</b>
2.1 緒言 . . . . .	13
2.2 階層型ニューラルネットワーク (MLP) . . . . .	14
2.3 自己組織化マップ (SOM) . . . . .	15
2.3.1 競合学習 . . . . .	15
2.3.2 アルゴリズム . . . . .	18
2.4 解析実験 . . . . .	22
2.4.1 観察対象試合 . . . . .	22
2.4.2 状況分析における観察要素の設定 . . . . .	24
2.4.3 チーム行動の有無による変化 . . . . .	26
2.4.4 チーム内での仲間との位置関係 . . . . .	31
2.4.5 人間同士によるフットサル試合とロボット同士によるサッカー試合の比較解析 . . . . .	35

2.4.6 結論	44
<b>第3章 自己組織的チーム行動獲得システム</b>	<b>46</b>
3.1 緒言	46
3.2 テンソル自己組織化マップ (TSOM)	47
3.3 TSOM を用いた学習方法検証	49
3.4 TSOM を用いた環境認識と行動決定	50
3.5 人間と同じ行動判断アルゴリズムの創発	51
<b>第4章 ロボットへの実装実験</b>	<b>53</b>
4.1 緒言	53
<b>第5章 考察及び結論</b>	<b>55</b>
5.1 緒言	55
5.2 今後の展望	55

## 表 目 次

1	Observation . . . . .	22
2	SOM parameters for experiment (Observation target game A) . . . . .	26
3	Role sharing by the relative distance between the player and the ball among the players in the team (Observation target game A) . . . . .	27
4	Role sharing by the relative distance between the player and the ball among the players in the team (Observation target game A) . . . . .	31
5	SOM parameters for experiment(Observation target game B,C) . . . . .	35

## 図 目 次

1	Human-Robot Symbiosis System . . . . .	8
2	RoboCup Middle Size League 2016 . . . . .	8
3	Observation of SOM algorithm . . . . .	17
4	3D View of $\sigma_\lambda = 3$ . . . . .	20
5	2D View of $\sigma_\lambda = 3$ . . . . .	20
6	3D View of $\sigma_\lambda = 2$ . . . . .	20
7	2D View of $\sigma_\lambda = 2$ . . . . .	20
8	3D View of $\sigma_\lambda = 1$ . . . . .	20
9	2D View of $\sigma_\lambda = 1$ . . . . .	20
10	Time variation of standard deviation . . . . .	21
11	Observation target game A . . . . .	23
12	Observation target game B . . . . .	23
13	Observation target game C . . . . .	23
14	The Observation Environment of the Game . . . . .	25
15	Defining the position by the relative distance within the team relative to the ball . . . . .	28
16	The feature map expressed by the gradation display (Observation target game A) . . . . .	29
17	Role sharing by the relative distance between the player and the ball among the players in the team (Observation target game A) . . . . .	29
18	The feature map expressed by the gradation display (Observation target game B).eps . . . . .	30
19	Role sharing by the relative distance between the player and the ball among the players in the team (Observation target game B) . . . . .	30

20	The feature map expressed by the gradation display (Observation target game B) . . . . .	33
21	The area of the rectangle constituted with the player of the defense team as the apex (Observation target game B) . . . . .	33
22	The area of the rectangle constituted with the player of the offense team as the apex (Observation target game B) . . . . .	34
23	Distance of the ball against the goal attacked by the offense team(Observation target game B) . . . . .	34
24	TRY01 red to red . . . . .	37
25	TRY02 red to red . . . . .	37
26	TRY03 blue to blue . . . . .	38
27	TRY04 red to blue . . . . .	38
28	TRY05 red to red . . . . .	39
29	TRY06 blue to blue . . . . .	39
30	TRY07 red to red . . . . .	40
31	TRY08 blue to blue . . . . .	40
32	TRY09 blue to blue . . . . .	41
33	TRY08 red to blue . . . . .	41
34	The feature map expressed by the gradation display (Observation target game B) . . . . .	42
35	A color-coded figure with the team try finally chosen a shot(Black : Euclidean distance $\geq 0.6$ ) . . . . .	42
36	The feature map expressed by the gradation display (Observation target game C) . . . . .	43
37	A color-coded figure with the team try (Black : Euclidean distance $\geq 0.6$ ) . . . . .	43
38	Data image of tensor SOM . . . . .	47

## Abstract

With the progress of technology, the realization of a symbiotic society with human beings and robots sharing the same environment has become an important subject. An example of this kind of systems is soccer game. Soccer is a multi-agent game that requires strategies by taking into account each member's position and actions. In this paper, we discuss the results of the development of a learning system that uses SOM to select behaviors depending on the situation. This system can reproduce the action selection algorithm of all players in a certain team, and the robot can instantly select the next cooperative action from the information obtained during the game. Because of this system, common sense rules was shared to learn an action selection algorithm for a set of agents, not only a team consisting of robots alone, but also a group of heterogeneous agents consisting of humans and robots.

# 第1章

## 序論

# 第1章 序論

## 1.1 緒言

### 1.1.1 ロボットの社会実装

近年、安全安心で持続可能な社会の実現、少子高齢化対応や第一次産業を始めとした産業基盤の構築等の社会的な課題に対する解決策として、ロボットの社会実装が期待されている。ここで、ロボットの社会実装の社会的期待について考えるため、工場内のロボット有用を例にとる。産業用ロボットの普及に伴い、工場内のロボット実装には労働者を守るための安全対策が複数関わってくる。代表的なものとして、労働安全規則や国際標準化機構 (International Organization for Standardization: ISO)、日本工業規格 (Japanese Industrial Standards: JIS) などが挙げられる。労働安全衛生規則第150条の4では、

事業者は、産業用ロボットを運転する場合（教示等のために産業用ロボットを運転する場合及び産業用ロボットの運転中に次条に規定する作業を行わなければならない場合において産業用ロボットを運転するときを除く。）において、当該産業用ロボットに接触することにより労働者に危険が生ずる恐れのあるときは、さく又は囲いを設ける等当該危険を防止するために必要な措置を講じなければならない。）

とある。この規則により、人間とロボットが同じ工場内で作業する場合において、その間には、柵や囲いを設ける必要があり、協働することは困難であった。しかし、規制改革実施計画及び近年の技術改心が加味され、平成25年12月24日付基発1224第2号通達によりいかが明記された。

産業用ロボットを使用する事業者が、労働安全衛生第28条の2による危険性等の調査（以下「リスクアセスメント」という。）に基づく措置を実施し、産業用ロボットに接触することにより労働者の危険の生ずるおそれが無くなったと評価できるときは、本条の「労働者に危険が生ずるおそれがあるとき」に該当しないものとすること。（以下省略）

したがって、現在では安全が確保できていると確認できる環境下であれば、人間とロボットが協働作業可能となった。

このように、法整備という観点から観察しても、社会的期待としては人間とロボットが同じフィールドで安心して作業や生活さえも行える将来が望まれている。しかし、度ボットにか

んする研究開発では、「人とロボットの関係性」についてどこまで考えられているであろうか。ロボットの社会実装には、社会的期待と研究開発との方向性の乖離を避け、研究成果をわかりやすく社会に還元・提示する必要がある。そのため、人間とロボットをどのように共存させるか、人間ロボット共生社会のあり方を議論すべきである。

### 1.1.2 協調行動

協調行動は、異なる自律エージェントが共通のタスクを実行しながらコミュニケーションを図るときに重要な側面である。多くの場合、単一のエージェントでのタスクを実行は必ずしも効率的とは言えず、ここ数年で、困難な問題を解決するためにマルチエージェントシステム(MAS)を研究している研究者も多い。マルチエージェントシステムは、複数のエージェント(自律エージェント)が協調行動によって共通の目標を達成しようとするシステムである。センサーから取得したデータに基づいてリアルタイムで決定を下すことによって、エージェント同士に加え、環境と相互作用する。

MASのテストベッドとして有名なRoboCupは、世界的ロボット大会であり、強化学習やニューラルネットワークなどの学習方法を使用してマルチエージェントの協調を一つの課題としている。RoboCup[?]はランドマークプロジェクトであり、2050年までに、人間のサッカーワールドカップチャンピオンチームに勝つロボットチームの制作プロジェクトである例としてこれまでに、Sandholm and Crites[?]によれば、十分な測定データと行動が利用可能であれば、強化学習は反復囚人のジレンマに対して最適な手法であると示された。さらに、Arai[?]は、環境がグリッドとしてモデル化されている場合のマルチエージェントシステムにおける追跡問題に関するQラーニングとProfit Sharing[?]での適性比較を行い、マルチエージェント環境での学習には、Profit Sharingが適していると結論づけている。しかし、これらは実ロボットへの適応までは検討されていない。

### 1.1.3 ロボット学習

学習アルゴリズムの中で、教師なし学習はMASシステムにとって有望な方法である。教師なし学習の利点は、ロボットが環境に関する以前の情報やロボット自身の前提知識を持っている必要がないことである。しかし、学習には多くのテストや経験が必要であるため、設計者は、状態空間とアクション空間を定義するために、試合中に考慮すべき重要なパラメータと変数を明確に定義することがコストと時間を削減するために重要である。状態空間の設計は、ロボットが行動空間で何ができるかによって異なる。同時に、状態空間はロボットの能力とそれ自身の行動空間に従って定義される。2つの空間は互いに相互接続している。効率的な方法で状態

空間を設計するために, Asada[?] は, 状態空間を最初に 2 つの主な状態に分割し, 次に状態数を増やしながら再帰的に多数の層に分割する方法を提案した. しかし, いくらかの偏りの問題は依然としてロボットの行動に影響を及ぼし, 何らかの誤った行動および習慣を引き起こす可能性があると示された.

#### 1.1.4 人間ロボット協調

前節で述べたように, 複数の実ロボット間での協調行動の獲得においては, 様々な議論がなされてきた. しかし, 人間ロボット共生社会を見据えた際に, これまでの学習に対する観察対象がロボットであることに, 疑問を呈する. 事実, インタフェースの問題を度外視しても, 人間と現在のマルチエージェントシステムを用いたロボット群との意思伝達は容易ではない. そこで, 本研究では, 人間同士による協調行動を観察しロボットに応用することを考える. (Fig. 1 参照) 実ロボットの学習において, 課題として挙げられていた, 経験の偏りや、環境認識における状態空間からの入力ベクトル要素の選定について, ニューラルネットワークを用いて, 検討する. 本稿では, まず SOM を用いて人間の協調行動におけるポジションによる状況分析を行う. 人間の協調行動の観察対象として, 今回はフットサルの試合を用いる. 複数のエージェントの協調によって目的が達成され, 環境が動的に変化し, 多目的・多重制約問題を全て包括し, かつ実時間計画・推論を必要とする点において, 人間ロボット共生社会での課題と一致するからである. また, 人間同士による協調行動を評価するために, ロボット同士による協調行動についても, 観察を行う. 観察対象は, ロボカップサッカーにおける中型リーグの試合である. ロボカップについては次章に記す.

ニューラルネットワークへの入力ベクトルの要素は, 人間とロボットの行動との間の可能な最小のギャップを達成するように選択される.そのためには, まず人間の行動を理解し, そのような行動を理解して模倣できるロボットを開発することが重要である. 文献 [??] では, フットサルゲームにおける人間とロボットの行動の研究に焦点を当てている. なぜなら, これは動的な環境, いくつかの制約を伴う優れたテストベッドであり, リアルタイムの計画が必要だからである. これらは, ロボットが将来動作する可能性がある一般的な共生システムの特性となる. 使用されるアルゴリズムでは, スコア, コーナーキック, ペナルティキック, ボールを持っているチームなど, ゲーム内の特定のイベントに応じて, SOM を使用してプレーヤーのポジションが分析された. その結果, 人間の試合もロボットの試合も同様に, 試合における場面を評価することが可能であり, 学習する入力要素を増やすことによって, 環境と行動の写像関係を示す SOM の作成によってロボットによる自律的行動の発現が可能であること

が判明した。この論文では、その写像関係を求め、テンソル自己組織化マップによりロボット同士の試合を学習し、サッカーロボットの行動学習システムの開発を行う。

### 1.1.5 ロボカップ

ロボカップとは「2050年までに人間のW杯優勝チームにかつ自律ロボットチームを作る」という大きな目標を掲げた国際ランドマークプロジェクトである。サッカーというテーマは、動的な環境下で囚人に判断をする知能や、マクドナルドシステムによるチーム行動といった要素を有している。しかし、サッカーというテーマだけでなく、ロボカップでは、人間のせい買う空間であるキッチンやリビングを再現したフィールドで競技を行う「@ホームリーグ」や、災害現場で教授に役立つ自律ロボット開発を推進する「レスキューリーグ」、次世代の人材育成を目的とした「ジュニアリーグ」、2012年から追加された「インダスリアルリーグ」など、「サッカーリーグ」を含めて現在では5つのリーグが存在し、様々な分野の技術を発展させるプロジェクトとなっている。各リーグで共通する課題は、「複数ロボットの協調によって目的は達成され、かんきょうが動的に変化し、多目的/多重制約問題を包含し、実時間計画/推論を必要とすること」である。

筆者が所属するチーム「Hibikino-Musashi」は、サッカーリーグの中の、「ロボカップサッカー中型ロボットリーグ (Middle Size League)」(Fig. ) である。

### ロボカップサッカー中型ロボットリーグ (MSL)

中型ロボットリーグ（以下、MSLと表記する）では、年々フィールドのサイズが大きくなっているが、現状では  $12[m] \times 18[m]$  の大きさのフィールドで試合を行うリーグである。各チームの試合に参加可能なロボットは5台までとなっており、ロボットの大きさは  $52[cm] \times 52[cm]$ 、高さ  $80[cm]$  以下となっている。（ゴールキーパーロボットは1秒間のみ、 $60[cm] \times 60[cm]$ 、高さ  $90[cm]$  まで拡大可能）ロボットは搭載されたカメラなどのセンサ情報から環境認識し、自律行動を行う。試合は前後半15分の合計30分間行われる。

### 試合データ

ロボカップサッカー中型リーグでは、2017年から試合中のロボットの自己位置・障害物認識状態などの情報をリアルタイムに近い状態で、提出することが義務化されるようになった。これは、試合中のデバッグ用だけでなく統計分析のためのデータ標準化でのある。

提出された全参加チームデータは、インターネット情に公開され、研究・開発その他に誰でも利用可能である。ロボットや障害物、ボールに関する全ての座標データは、どのチームも

定義された 3 次元空間のデカルト座標系で提出しなければならない。

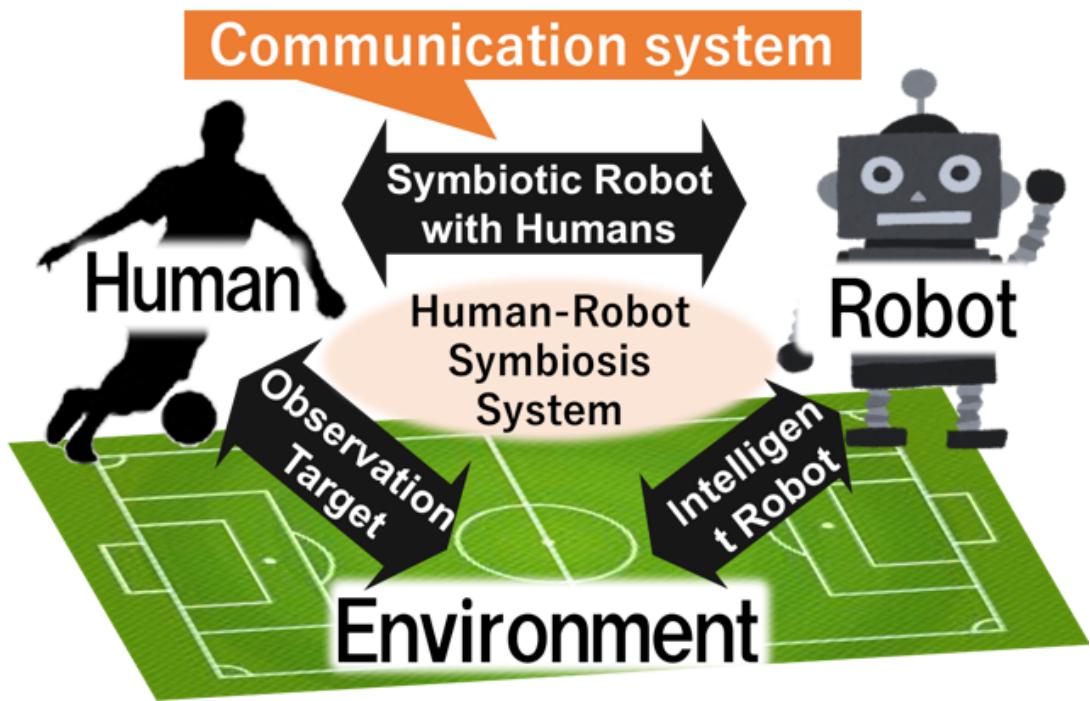


図 1: Human-Robot Symbiosis System



図 2: RoboCup Middle Size League 2016

## 1.2 本研究の目的

本研究の目的は、人間のチームとしての行動決定アルゴリズムを解析し、教師なし学習である自己組織化マップを用いて、近似されたモデルを用い、ロボットに環境人氏から行動決定の写像関係を学習させることを目的とする。

### 1.3 本論文の構成

本論文の構成は、第2章では、本研究において使用する自己組織化マップについての歴史とアルゴリズムを述べた上で人間のチーム行動の解析を行う。第3章では、第4章では、最後に第5章では、考察及びまとめを述べる。

#### 1.4 変数, 記号の約束

ここに表が入るかな ??

<sup>1</sup>

---

<sup>1</sup>補足があればここに...

## **第2章**

# **人間のチーム行動解析**

## 第2章 人間のチーム行動解析

### 2.1 緒言

本章では、人間のチーム行動の解析及びロボットへの適応について議論するため、ニューラルネットワークに代表される学習機能を有するアルゴリズムを取り上げる。ニューラルネットワークの特徴として、様々なシステムに対して寛容であり、学習を行うことで非線形的な信号に対しても適応可能であるようなロバスト性を有する。

## 2.2 階層型ニューラルネットワーク (MLP)

階層型ニューラルネットワークは、入出力間の写像関係を学習して構築され、信号が常に入力層から出力層方向に向かう脳型情報処理アルゴリズムである。入力データと出力目標データの対からなる訓練データを多数組与え、ネットワークがこれらの入出力関係を再現するよう結合荷重を修正していくことで学習が進む。

## 2.3 自己組織化マップ (SOM)

自己組織化とは、局所的な相互作用から地域的に順序づけられた構造が導かされることであり、複数の要素から構成されるシステムが他からの制御を得ることなく、時間とともに何らかの意味で自発的に秩序化する過程である。神経回路の構造は、脳内においてこの自己組織化の代表的な例として挙げられる。1956年にSperryによってはじめられた視覚神経通路の発達における位相的順序を保持した状態での網膜位相マップの構築を目的としたモデル化が、自己組織化の位相形成原理についてのモデル化の先駆けとなり、Von der Malsburg[?]が特徴選択における皮質細胞の局所的な順序づけを目的としたシナプス学習を含む自己組織化過程を導入するなど、創刊がある神経活動によって駆動される位相的なマップの形成に関係した自己組織化の要素が神経活動による位相マップの初期の形成には存在するという結論が主流となった。その後、1976年にWillshawとvon der Malsburg[?]によって、本質的には同様の幾何学的関係がある離散的な格子間での類似関係を検出する自己組織化処理モデルが、自己組織化マップ (Self-Organizing Map : SOM) として提案され、SOMは、教師なし学習であり、生命の神経回路による情報処理に近い競合学習として知られるようになった。

Kohonenは前述したモデルからさらに、類似関係があるものではない、かつ幾何学条性格である必要もない離散的格子上に連続した入力空間からマップ化することを目的として学習アルゴリズムを変更した[?].より強度な自己組織化効果を得るために、近傍関係を学習に導入するため荷重更新規則を修正し、これらの近傍関係によって広がった範囲が時間経過にともない縮まる要素も追加した。これにより、SOMのアルゴリズムとして、Kohonenモデルが一般的となり、本論文中にあいてもKohonenモデルを取り扱う。

この構造により、多次元入力ベクトルの低次元化が出来そうに可能であり、特徴が近似している入力ベクトルほど、近くに配置される。SOMの応用例としては、複雑なデータの二次元表示への可視化があり、クラスタリング技術と同様に、抽象化の創造が可能である。優れた補間性能を保つために、学習データの一般化が可能であり、未知な入力データに対しても妥当性の高い出力が期待される。

### 2.3.1 競合学習

競合学習は階層型ニューラルネットワークにおいて重要な懸念である。同一の入力に対し、各々のニューロンが活性度を競い合う側方相互作用が利用される。具体的に競合学習としてのSOMは、入力層と出力層からなり、入力層(入力空間V)の各入力サンプル(入力ベクトル $v(t)$ )は全ての出力層ニューロンと結合し、結合には結合荷重 $w_i$ が付加されている。ここ

で  $t$  は時間座標であり、 $N$  個の出力層ニューロンは、 $i = 1, 2, \dots, N$  とラベル付けされている。 (Fig3) そして、近くのニューロンとの結合は興奮性（正の結合荷重）があり、より遠方のニューロンとの結合には抑制性（負の結合荷重）がある状態で出力層のニューロンは活性化するために競合する。

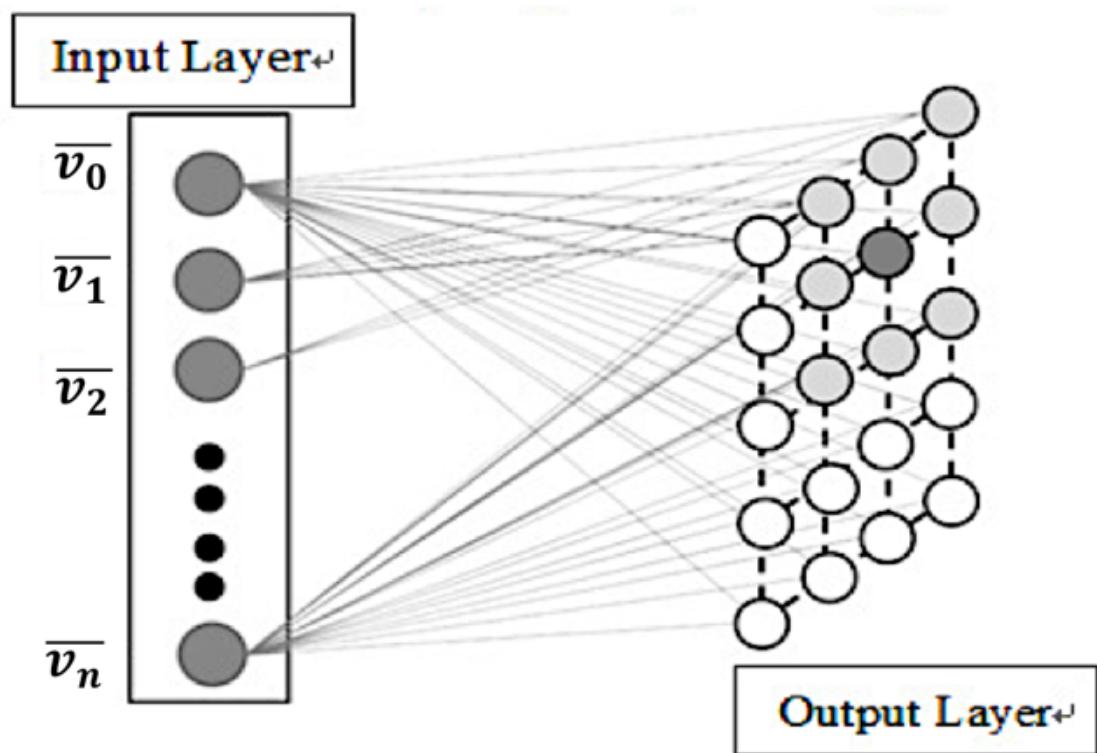


図 3: Observation of SOM algorithm

### 2.3.2 アルゴリズム

Kohonen の SOM アルゴリズム [?] は、大きく競合段階と協調段階の二つの改装で構成される。最初の段階ではまず、入力に対して、最整合ニューロン (Best Matching Unit : BMU) の選択すなわち勝者を選択する。次に、その最整合ニューロンおよびその直近の格子の荷重変更が行われる。以下に、これら二つの段階について述べるが、本研究では、新しい入力ベクトルの投入のたびに、荷重更新を行う増分学習 (オンライン学習) ではなく、全入力ベクトルの協業段階が行われた後に、生成できる全入力が考慮された荷重更を行う、バッチ型 SOM[?] を用いる。これにより、入力ベクトルの投入順番による出力への影響を考慮する必要がなくなる。よって、以降は本研究で用いたバッチ型 SOM について述べているため、バッチ型 SOM への変更に伴い、入力空間  $V$  を  $\mu = 1, \dots, M$  の  $M$  個のサンプルからなる固定の学習ベクトル集合  $\{v^\mu\}$  として定義する。

#### 競合段階

競合段階では、各入力ベクトル  $v^\mu$  に対する最整合ニューロンについて式 (1) を用いて選択する。BMU の決定方法としては、内積距離規則やユークリッド距離規則が考えられる。内積距離規則は生物学的モデルに類似しているため、多くその方面で用いられているが、今回は生物学的なモデルとして SOM を扱う必要がないため、ユークリッド距離規則による BMU 決定を採用する。ユークリッド距離を用いる場合、結合荷重改め、ベクトル量子化の観点から参照ベクトルと呼ぶこととする。ここでは、勝者  $j^*$  とラベル付けする。また、 $\|.\|$  はユークリッド距離を表している。

$$j^* = \operatorname{argmin} \|w_i - v^\mu\| \quad (1)$$

#### 協調段階

協調段階では、ニューロンが発火 (BMU が決定されること) することにより、その周囲のニューロンも協調的に発火しやすくなるように荷重更新に影響を与える段階である。各丹生論の参照ベクトル  $w_i$  を式 (2) により更新する。ここで、 $t$  は学習回数を示す変数であり、総学習回数を  $t_{max}$  とする。 $\eta$  は学習係数である。

$$w_i(t+1) = w_i(t) + \frac{\eta}{M} \cdot \Delta w_i \quad (2)$$

ただし、参照ベクトルの増減値  $\Delta w_i$  は、次式 (3) で示される。

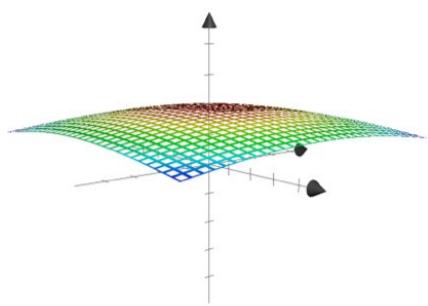
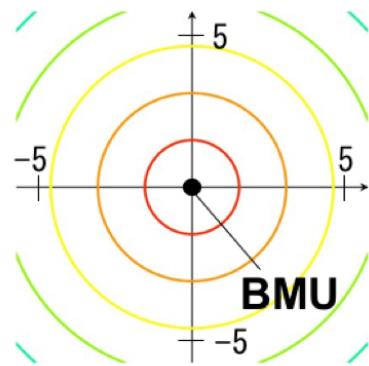
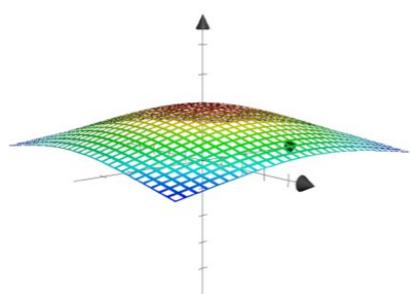
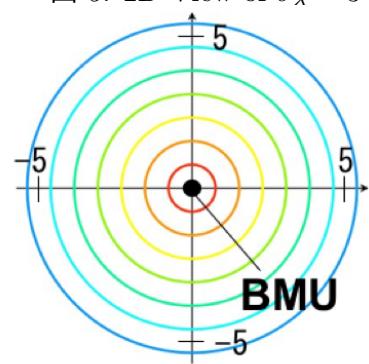
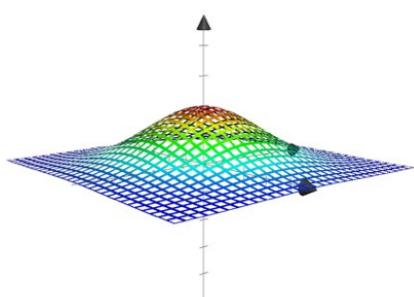
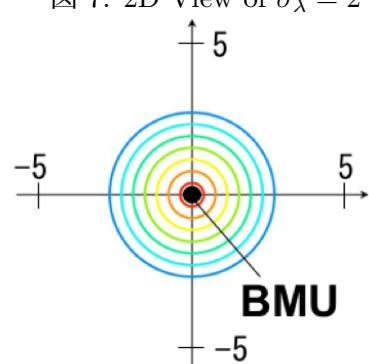
$$\Delta w_i(t+1) = \Delta w_i(t) + \Lambda(i, i^*, \sigma_\Lambda) \cdot (v^\mu - w_i) \quad (3)$$

ここで、 $\Lambda$  はニューロン  $i$  と  $i^*$  の格子座標  $r$  と  $r^*$  での尺度調整を行う近傍関数である。勝者の位置に関して左右対称であり、勝者からの格子距離が増加するにしたがって単調に減少する関数でなければならない。今回は、もっとも一般的なガウス型近傍関数(式 4 参照)を用いた。(図 4-9 参照)  $\sigma_\Lambda$  は、その範囲(標準偏差)である。

$$\Lambda(i, i^*) = \exp\left(-\frac{\|r_i - r_{i^*}\|^2}{2\sigma_\Lambda^2}\right) \quad (4)$$

先に述べた、近傍関数によって広がった範囲が時間経過とともに縮まる要素として、標準偏差、式(5)によって学習回数とともに小さな値をとる。(図 10 参照)

$$\sigma_\Lambda = \sigma_{\Lambda_0} \exp\left(-2\sigma_{\Lambda_0} \frac{t}{t_{max}}\right) \quad (5)$$

図 4: 3D View of  $\sigma_\lambda = 3$ 図 5: 2D View of  $\sigma_\lambda = 3$ 図 6: 3D View of  $\sigma_\lambda = 2$ 図 7: 2D View of  $\sigma_\lambda = 2$ 図 8: 3D View of  $\sigma_\lambda = 1$ 図 9: 2D View of  $\sigma_\lambda = 1$

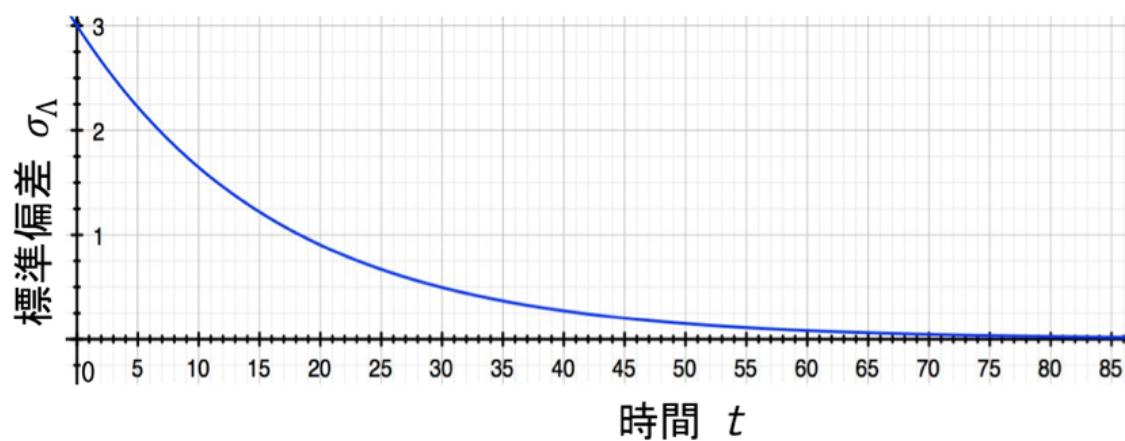


図 10: Time variation of standard deviation

## 2.4 解析実験

実験は、大きく分けて以下の3つの目的について行う。

1. 普段からチーム行動をしているチームとしていないチームで差がでるのか。
2. 人間によるサッカーにおいて同チームの仲間との位置関係は、SOMのクラスタリングとどう影響するか。
3. 人間によるサッカー(普段からチームプレーをしているチーム)とロボットによるサッカーの試合では、SOMにできるクラスタの観測はどう異なるか。

### 2.4.1 観察対象試合

今回観察対象とする試合は、Table1に示す3試合である。

表 1: Observation

試合名	プレイヤー	チーム行動
A	人間	慣れていない
B	人間	慣れている
C	ロボット	慣れている

試合Aについての観察対象は、Fig11に示しており、九州工業大学の大学生の即席チーム同士の試合である。また、試合Bについては、普段から、フットサルチームとして活動されているチーム同士の試合である。(Fig12参照)最後に、試合Cは2016年に行われたロボカップ世界大会中型リーグ決勝戦でのオランダチーム(Tech United Eindhoven)と中国チーム(Water)の試合である。(Fig13参照)

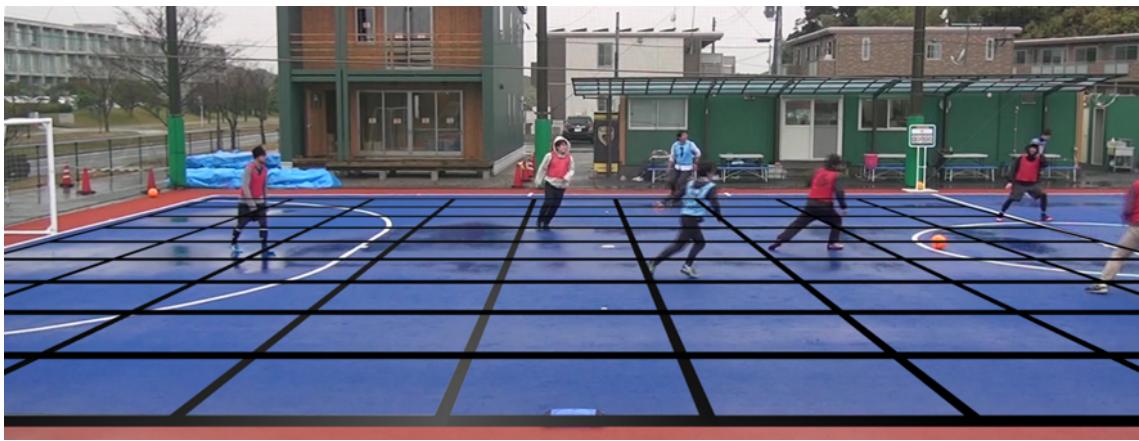


図 11: Observation target game A



図 12: Observation target game B



図 13: Observation target game C

### 2.4.2 状況分析における観察要素の設定

試合を観察するにあたり、まず、用意に観察可能な要素を対象として設定する。対象としては、選手やボールの座標や速度に加え、試合経過時間や、得失点差なども考えられる。今回は、人間による試合の観察方法として、試合フィールド全体を、単眼カメラを用いて撮影する方法を選択したため、フィールド内での様々な座標データをもとに、単眼カメラを用いて撮影する方法を選択したため、フィールド内での様々な座標データをもとに、状態空間を張る。人間による試合・ロボットによる試合どちらとも、2チームの選手同士5対5で行い、全選手とボールの意志座標をピクセル座標からフィールド座標に変換して取得した。(Fig??参照) また、観察ベクトル  $A$  を式(6)で示す。

$$A = [P_{k=1}, P_{k=2}, \dots, P_{k=10}, B] \quad (6)$$

ただし、フィールドでの選手位置座標  $P_k$ 、ボール座標  $B$  については(7), (8)に示す。ここで、 $k$ は選手識別のためのIDであり、 $k = 1, 2, 3, 4, 5$ の選手で構成されるチーム1と $k = 6, 7, 8, 9, 10$ の選手で構成されるチーム2での試合を観察している。なお、 $k = 5, 10$ は両チームのギールキーパーである。

$$P_k = [P_{x_k}, P_{y_k}] \quad (7)$$

$$B = [B_x, B_y] \quad (8)$$

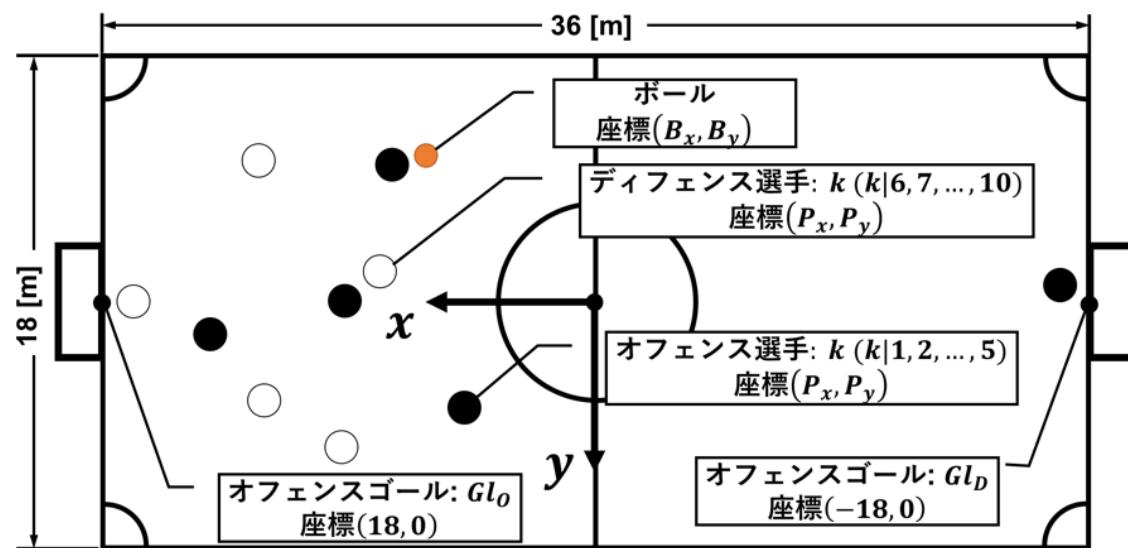


図 14: The Observation Environment of the Game

### 2.4.3 チーム行動の有無による変化

#### 実験方法

まず、人間同士の試合 A, B を比較して、チーム行動の慣れ具合により違いが生じるかを観測した。入力ベクトルを、式 (9) に示す。

$$v^\mu = [P_k, B] \quad (9)$$

評価は、近傍ユニットとのユークリッド距離によるクラスタ配置の違いと、Fig15 に示すように、ボールに一番近い選手からアルファベータガンマと割り振った際、学習が収束した後の各入力ベクトルに対する最整合ユニット (BMU) の配置に対して行う。また、クラスタが明確に生成可能かどうかも観察する。

#### 実験結果

まず、試合 A について、各パラメータを Table2 に示す。また、作成した SOM の特徴マップを Fig.16,17 に示す。ここで、色のグラデーションは、近傍のユニットとの距離が遠くなるにつれて赤色、近くなるにつれて、青色のグラデーションで示している。つまり、青いユニットの集団は類似している重みベクトルが配置されていることとなる。同様に表 3, Fig.18,19 に試合 B についてのパラメータおよび作成した SOM の特徴マップを示す。

表 2: SOM parameters for experiment (Observation target game A)

パラメータ名	変数	値
学習回数	$T$	10
入力データ数	$M$	300
マップサイズ	-	$15 \times 15$
初期近傍半径	$\sigma_{\Lambda_0}$	0.8

#### 考察

Fig.16, Fig.17 からわかるように、チームプレーに慣れていない選手群の試合では、ボールに対するチーム内での選手の相対的距離により、クラスタが分けられることはなかった。しかし、Fig.18, Fig.19 からわかるように、チームプレーに慣れている選手群の試合では、クラスタリングされた。図で赤いユニットによって大きく二つのクラスタに分けられ、そこでボールに近いアルファとベータの群と、遠いデルタ、ガンマの群は大きく分けられた。このこと

表 3: Role sharing by the relative distance between the player and the ball among the players in the team (Observation target game A)

パラメータ名	変数	値
学習回数	$T$	10
入力データ数	$M$	400
マップサイズ	-	$20 \times 20$
初期近傍半径	$\sigma_{\Lambda_0}$	0.8

から、チームプレーに慣れている選手から構成されているチームのほうが、チームプレーに慣れていない選手から構成されているチームよりもチーム内での自分のボールへの相対的近さに対する行動選択への影響が大きいことがわかる。

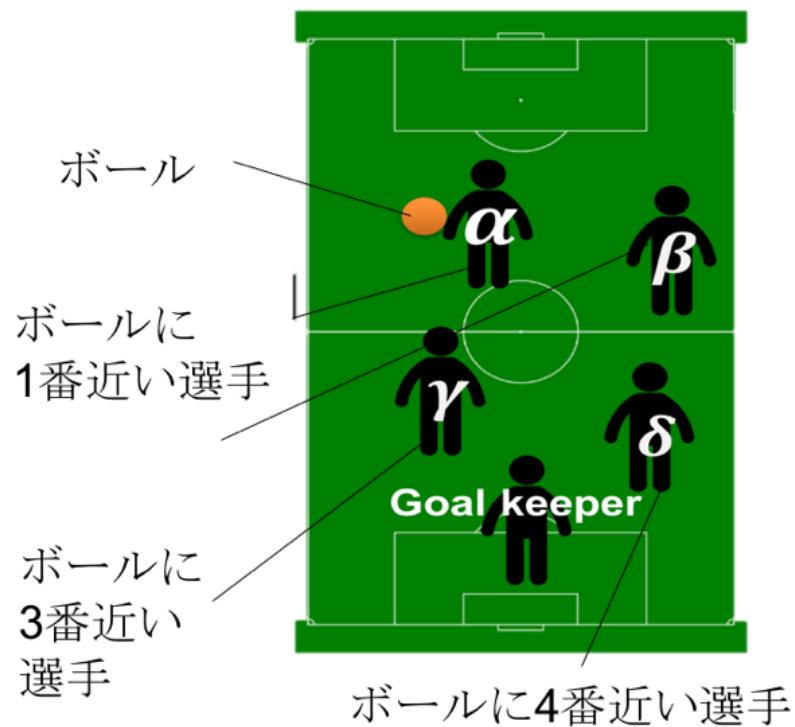


図 15: Defining the position by the relative distance within the team relative to the ball

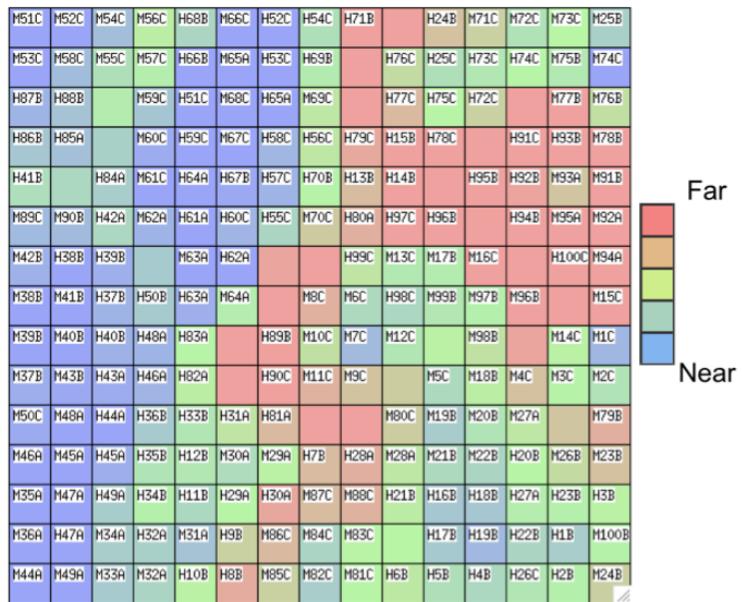


図 16: The feature map expressed by the gradation display (Observation target game A)

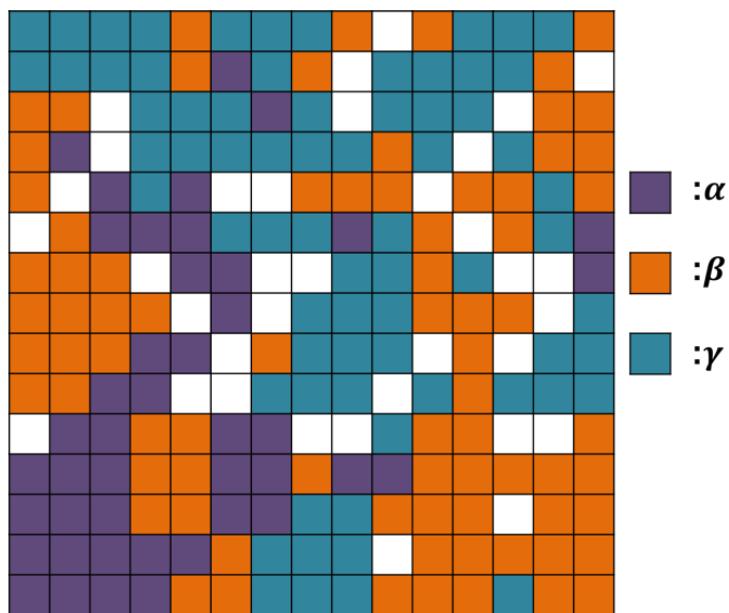


図 17: Role sharing by the relative distance between the player and the ball among the players in the team (Observation target game A)

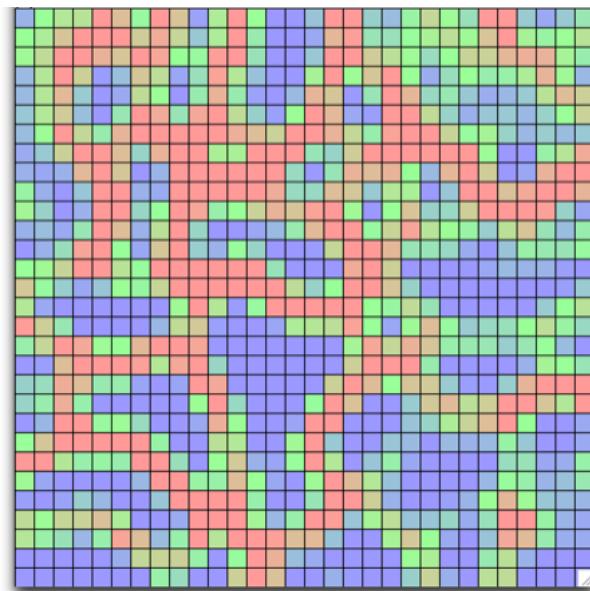


図 18: The feature map expressed by the gradation display (Observation target game B).eps

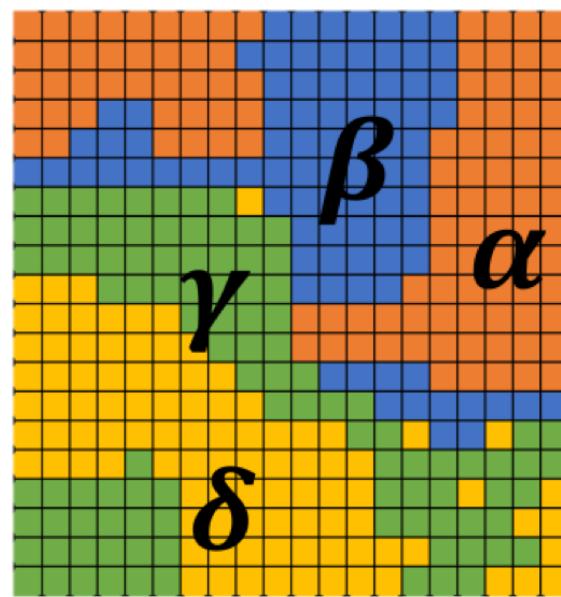


図 19: Role sharing by the relative distance between the player and the ball among the players in the team (Observation target game B)

表 4: Role sharing by the relative distance between the player and the ball among the players in the team (Observation target game A)

パラメータ名	変数	値
学習回数	$T$	10
入力データ数	$M$	400
マップサイズ	-	$20 \times 20$
初期近傍半径	$\sigma_{\Lambda_0}$	0.8

#### 2.4.4 チーム内での仲間との位置関係

##### 実験方法

次に、人間同士の試合 B について、チーム内での仲間の選手の行動を入力ベクトルに入れ、仲間との位置関係により、クラスタリングが可能かどうかを観察する。入力ベクトルを、式 (10) に示す。

$$v^\mu = [P_k, P_2, P_3, P_4, B] \quad (10)$$

評価は、近傍ユニットとのユークリッド距離によるクラスタリングと、ゴールキーパー以外の仲間の選手からできる四角形の面積の大きさを見ることで、選手の疎密具合による最整合ユニット (BMU) の最近ユニットとのユークリッド距離によるクラスタと一致しているかを確かめる。仲間の選手 4 人の位置座標と頂点とする四角形の面積を用いた評価関数は三斜法に基づき式 (11) により導出した。ここで  $d$  は、1 チームの 4 人の選手を x 座標値が小さい順に並べ直した時の各位置座標の ID である。

$$Y = \frac{1}{2} \Sigma (P_{x_d} - P_{x_{d+1}})(P_{y_d} + P_{y_{d+1}}) \quad (11)$$

各パラメータを Table4 に示す。また、作成した SOM の特徴マップを Fig.20-23 に示す。

##### 考察

まず、Fig ??を見ると、大きく 4 つのクラスタが観測できる。右上から左下にかけての協会と左右を分けるように上下に伸びる境界が見て取れる。しかし、Fig??, Fig??, Fig??を観察すると、どれもこのクラスタと合致するようにクラスタ生成がなされていない。次に Fig??, Fig??を観察すると、オフェンス、ディフェンス両方のチームで特徴マップ上部の色が赤と青とで対象的である。この部分は、Fig ??からボールがオフェンスチームが攻めるゴールから

遠い時を示している。したがって、オフェンスチームは自陣の攻めるゴールからボールが遠くなると密集し、対象雨滴にディフェンスチームは自陣が守るゴールからボールが遠くなると分散したポジショニングをしていることがわかる。



図 20: The feature map expressed by the gradation display (Observation target game B)

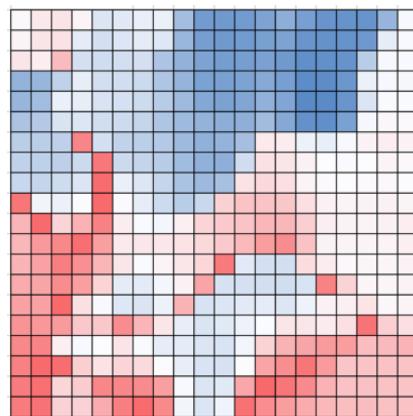


図 21: The area of the rectangle constituted with the player of the defense team as the apex (Observation target game B)

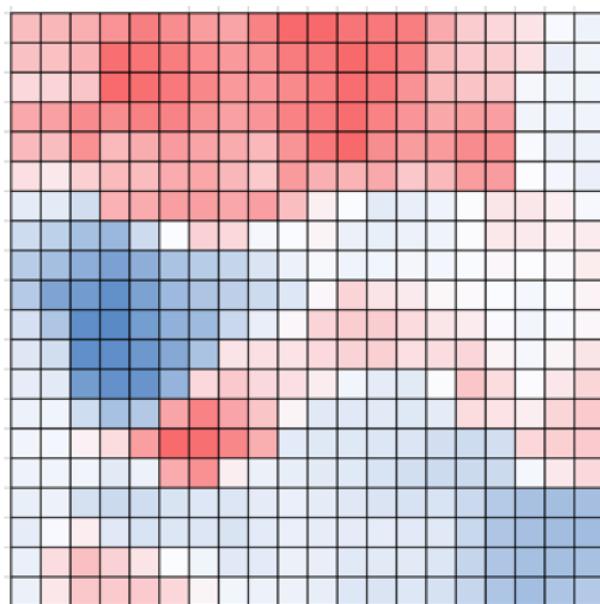


図 22: The area of the rectangle constituted with the player of the offense team as the apex  
(Observation target game B)

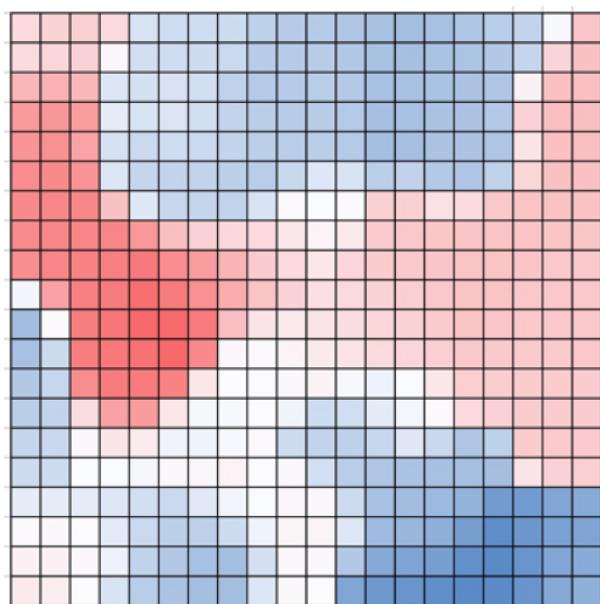


図 23: Distance of the ball against the goal attacked by the offense team(Observation target game B)

#### 2.4.5 人間同士によるフットサル試合とロボット同士によるサッカー試合の比較解析

全選手が参加している状態において、セットプレイからシュートまでの選手・ボール位置座標を用いた SOM によるクラスタリングを行い、試合状況の解析を行う。ここにおいて、セットプレイとは、キックオフ・フリーキック・コーナーキックなど、試開始に際しボールをセットして行うプレーを指す。SOM の学習において、入力ベクトル 9 次元であり、次式 (12) で示す。

$$v^\mu = [P_k, P_2, P_3, P_4, P_6, P_7, P_8, P_9, B] \quad (12)$$

ここで、今回はゴールキーパー以外を対象とするため、ゴールキーパーである  $k=5,10$  以外の  $k=1,2,3,4$  および  $k = 6,7,8,9$  選手位置座標を用いる。以降簡単のために、 $k=1,2,3,4$  の選手で構成されるチームを赤チーム。 $k = 6,7,8,9$  の選手で構成されるチームを青チームとする。また、以下の選定条件によりデータを取捨選択した。

1. セットプレイからシュートまでの一連のプレーであること
2. 1 トライで選手の交代が行われていないこと

ここで 1 トライとは、1 回のセットプレイからシュートまでのことである。今回入力データとして用いたトライは以下 Fig.?? の計 10 トライである。

#### 学習パラメータ

Table5 に本実験においての、SOM のパラメータを示す。

表 5: SOM parameters for experiment(Observation target game B,C)

パラメータ名	変数	値
学習回数	$T$	10
入力データ数	$M$	200
マップサイズ	-	$30 \times 30$
初期近傍半径	$\sigma_{\Lambda_0}$	0.8

#### 実験結果

実験の結果 1 つの特徴マップを作成した。ユークリッド距離による色のグラデーションを Fig.?? に示す。赤い部分がユークリッド距離の大きい部分となる。

また、クラスタリングされた学習終了後の特徴マップに、各入力ベクトルの最終的な最整合ニューロン (BMU) を付加したものを Fig ??に示す。黒く示した部分が、ユークリッド距離が大きいニューロンであり、赤や青で示した部分が、赤チームや青チームが最終的にシュートに漕ぎ着けたトライの入力ベクトルが選択した最整合ニューロンである。

### 考察

攻めるチームが、赤チームか青チームという違いで大きく選択したクラスが異なった。これは、攻めるときと守るときで、チーム内でのフォーメーションが大きく異なることを示している。この時は、セットプレーからシュートまでを1トライと換算し、入力した。人間の試合では結果、赤チームがセットプレーからシュートまでボールを奪われずにゴールした時、青チームがセットプレーからシュートまでボールを奪われずにゴールした時、赤チームがセットプレーを始めたが途中で青チームに奪われてしまった時、青チームがセットプレーを始めたが途中で赤チームに奪われてしまった時のすべての場合が、違うクラスタに分類された。

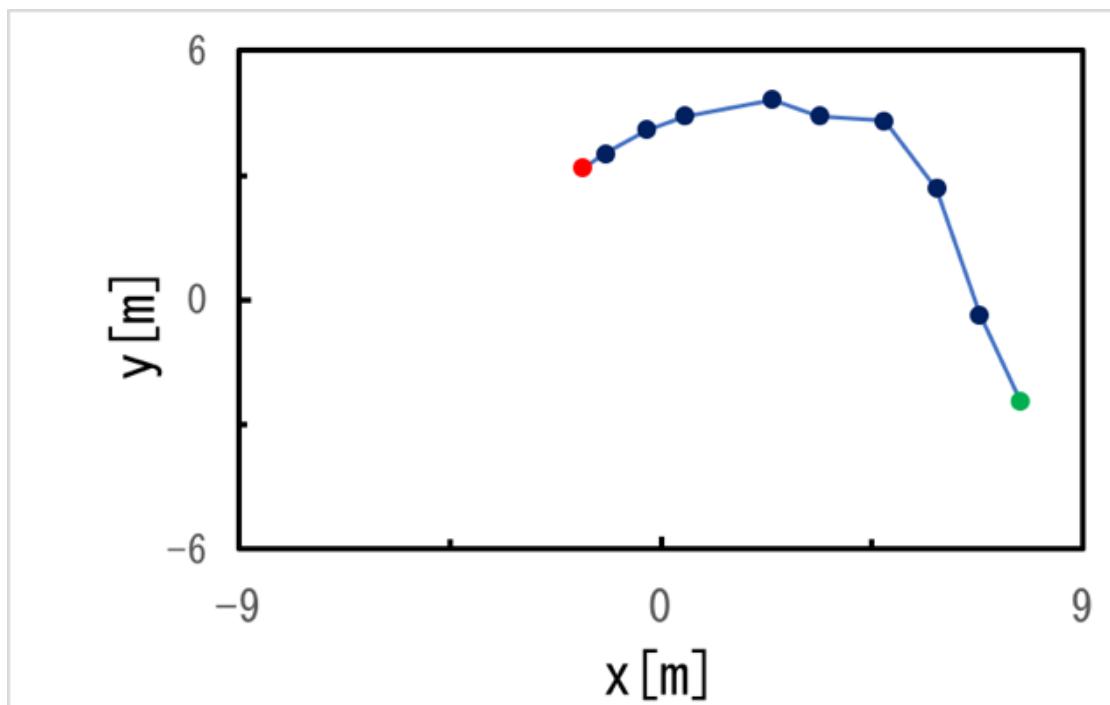


図 24: TRY01 red to red

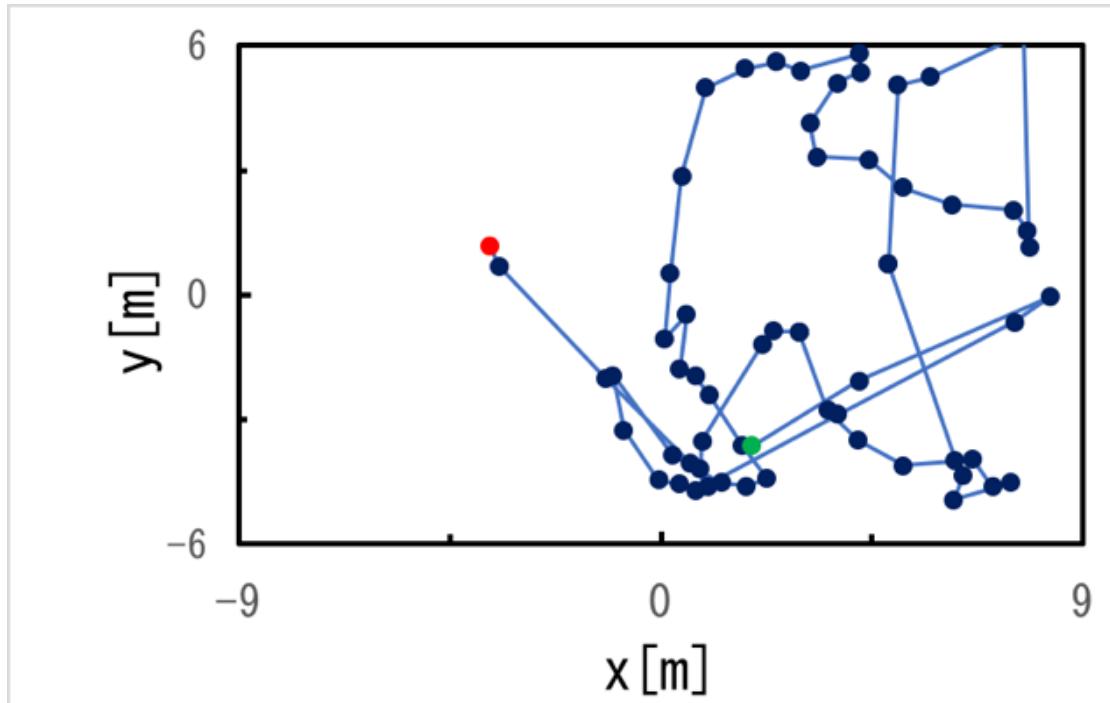


図 25: TRY02 red to red

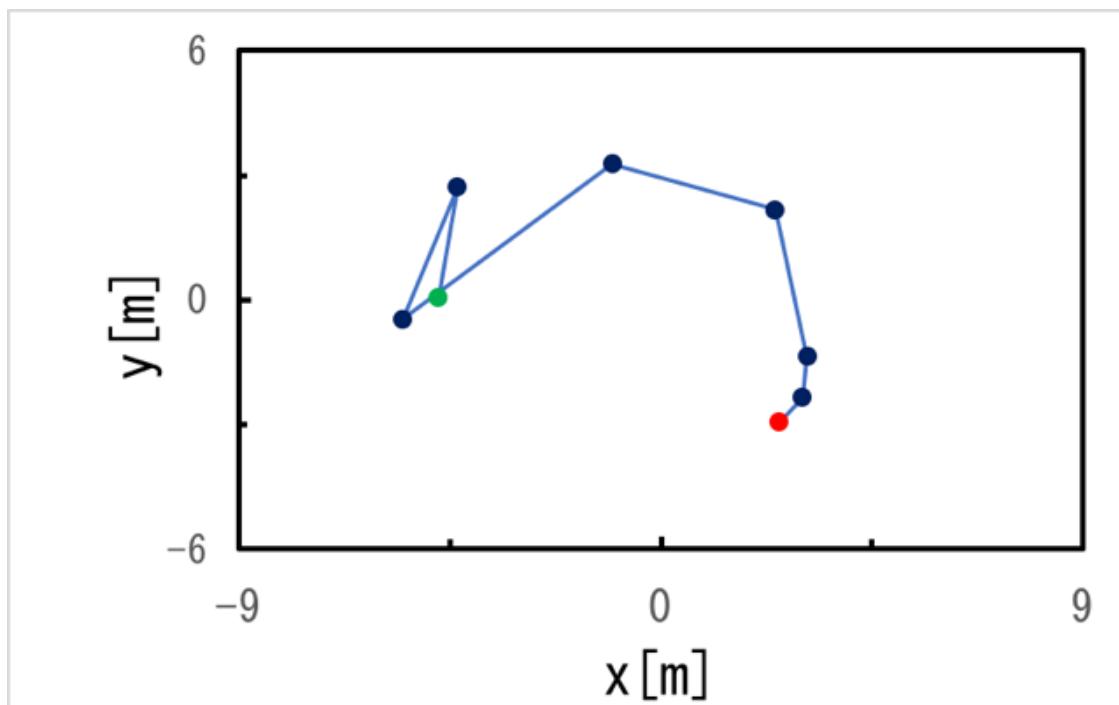


図 26: TRY03 red to blue

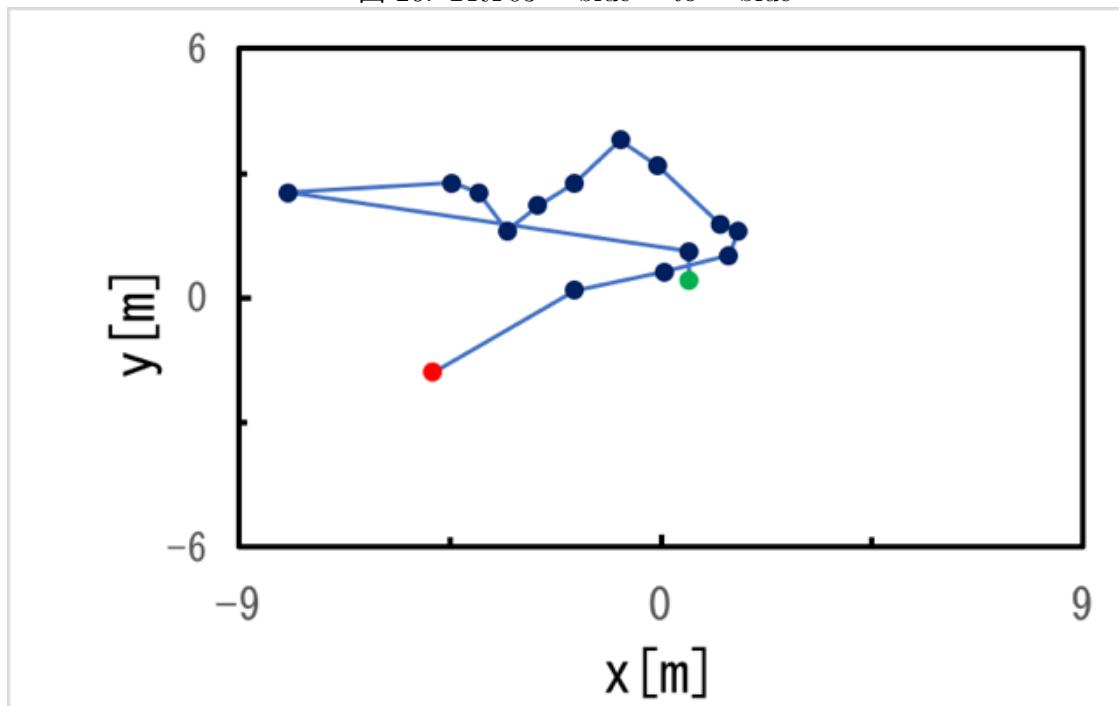


図 27: TRY04 red to blue

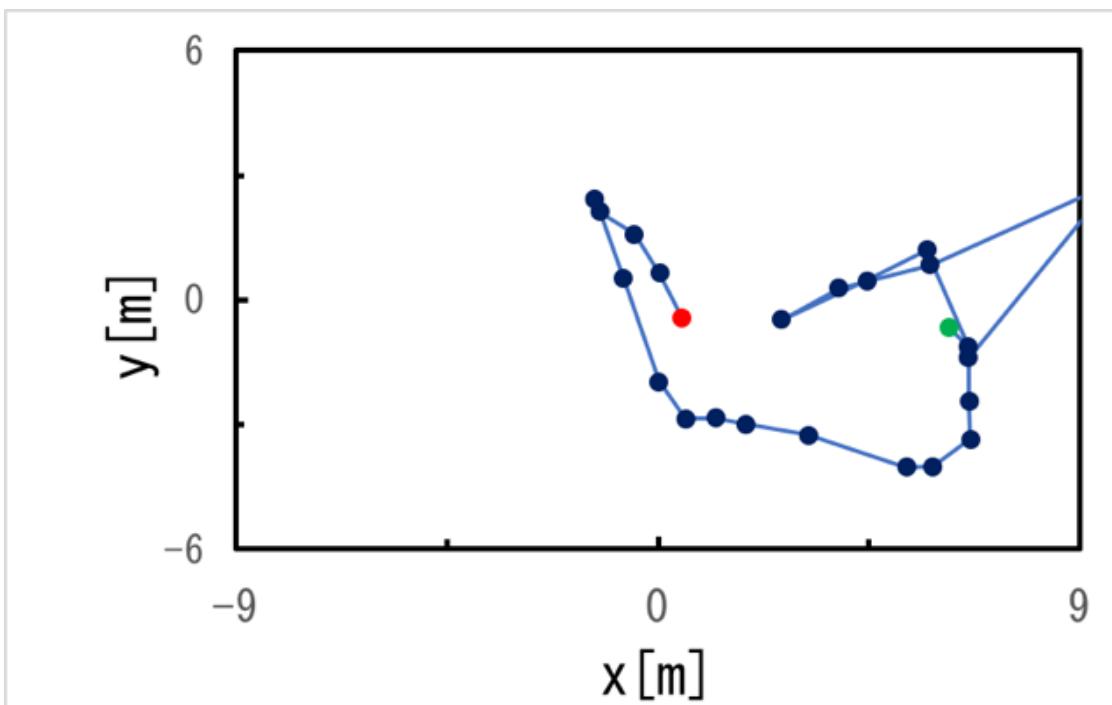


図 28: TRY05 red to red

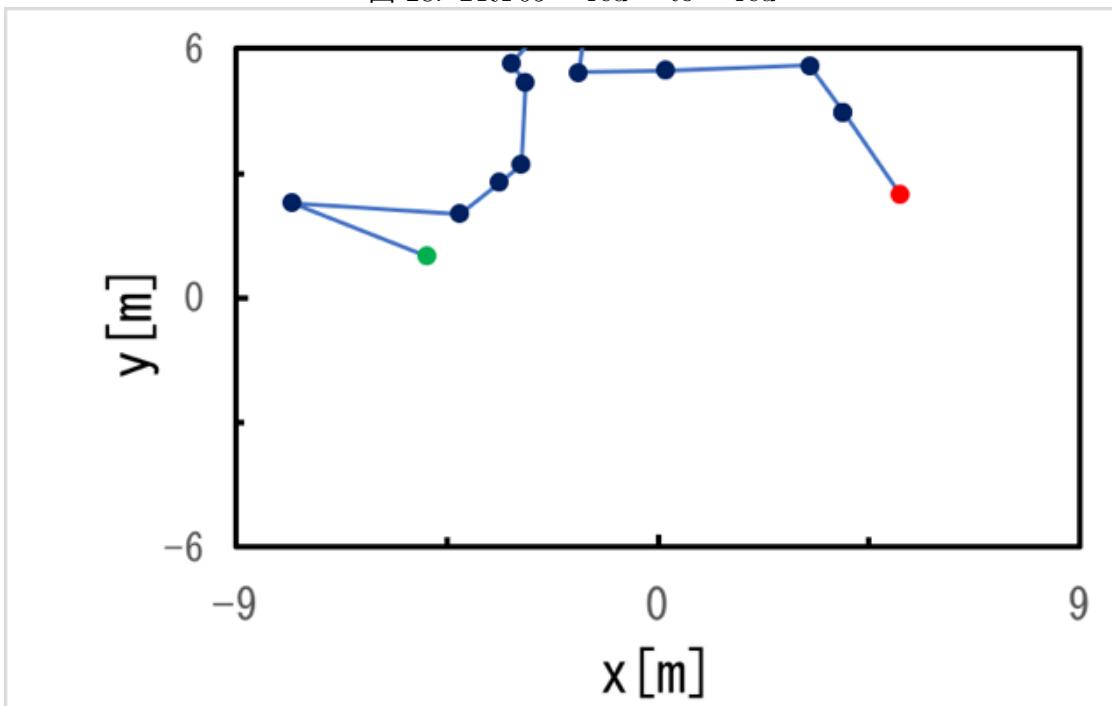


図 29: TRY06 blue to blue

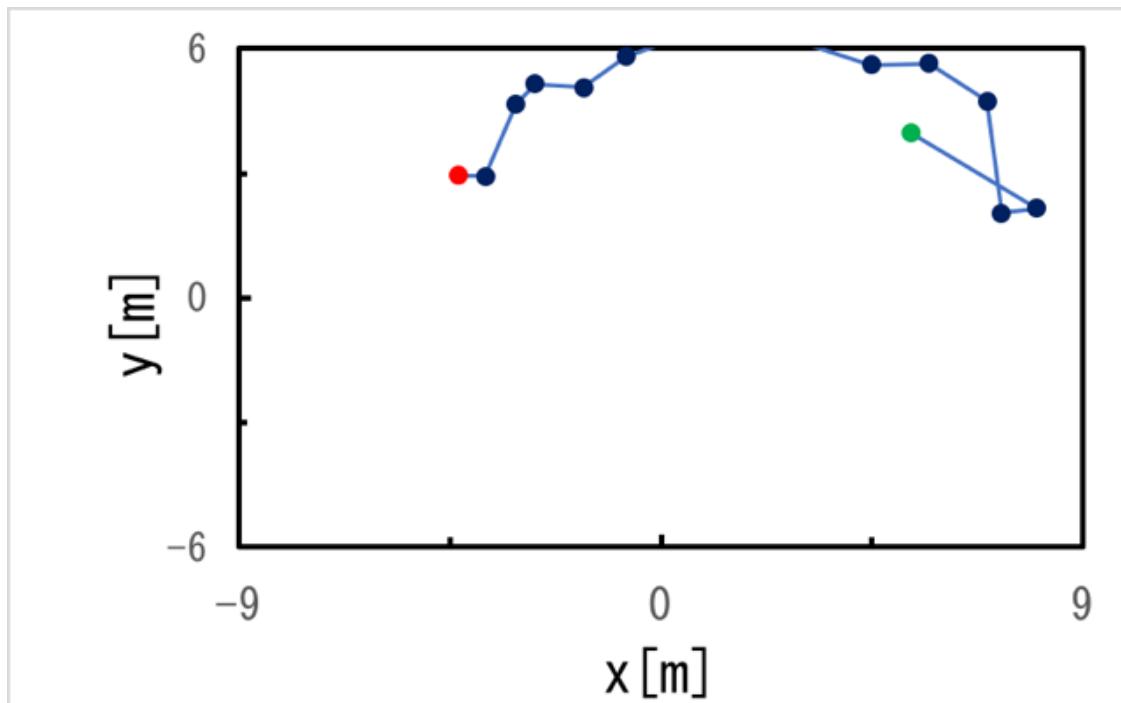


図 30: TRY07 red to red

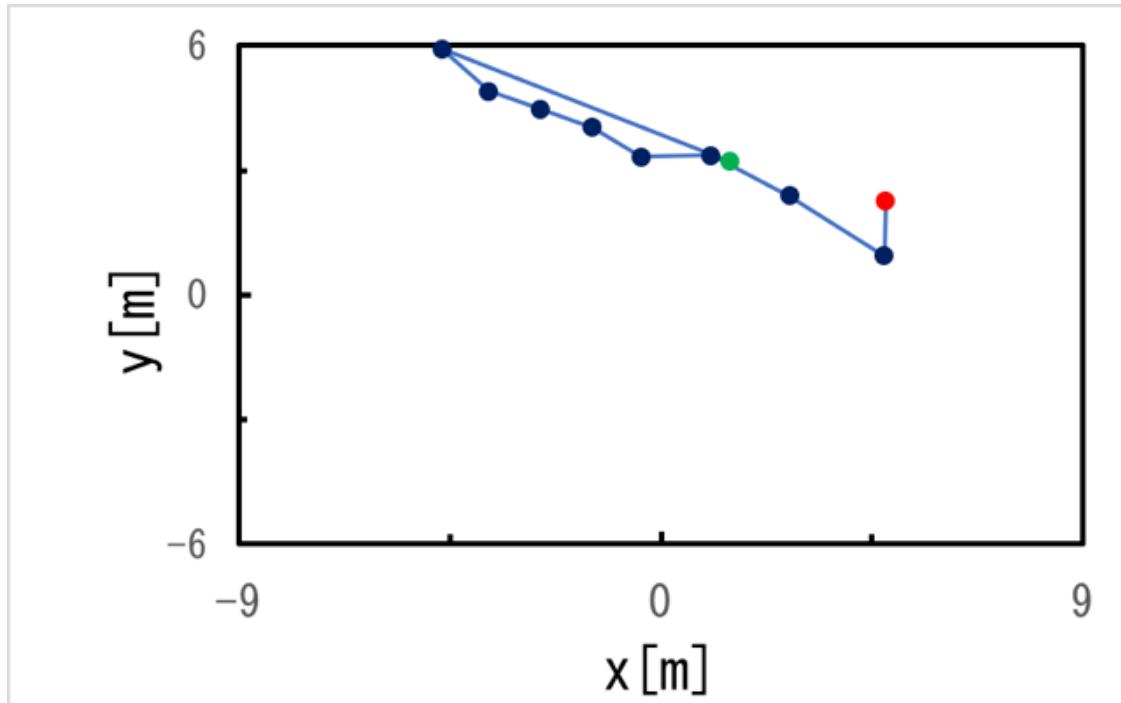


図 31: TRY08 blue to blue

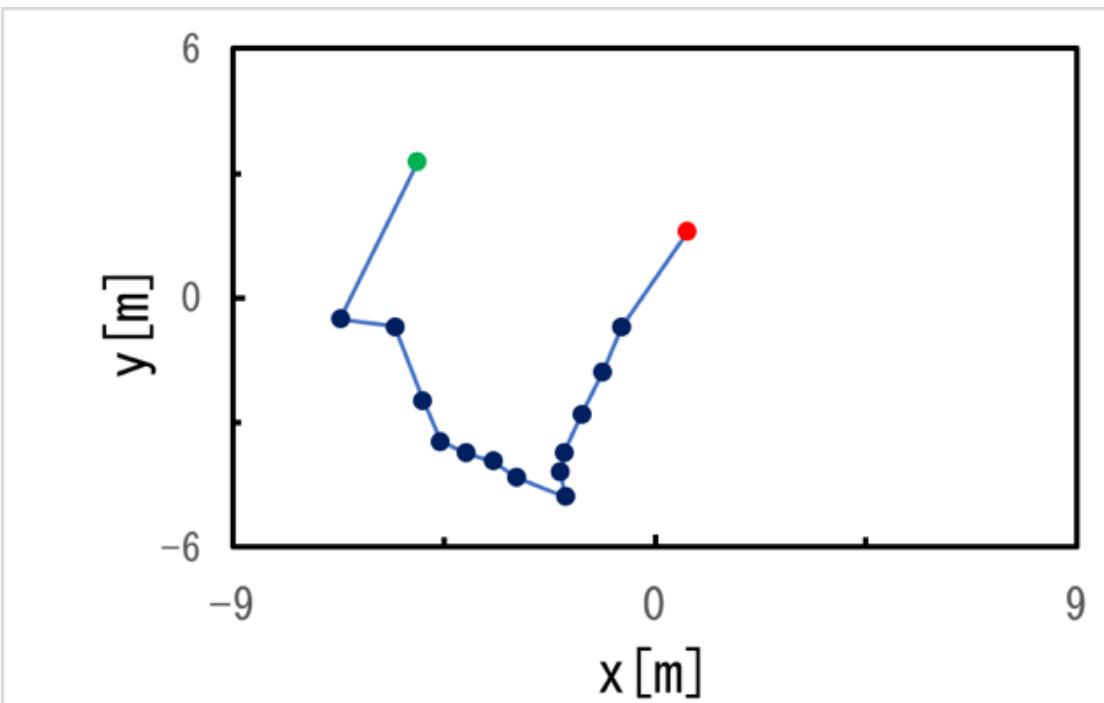


図 32: TRY09 blue to blue

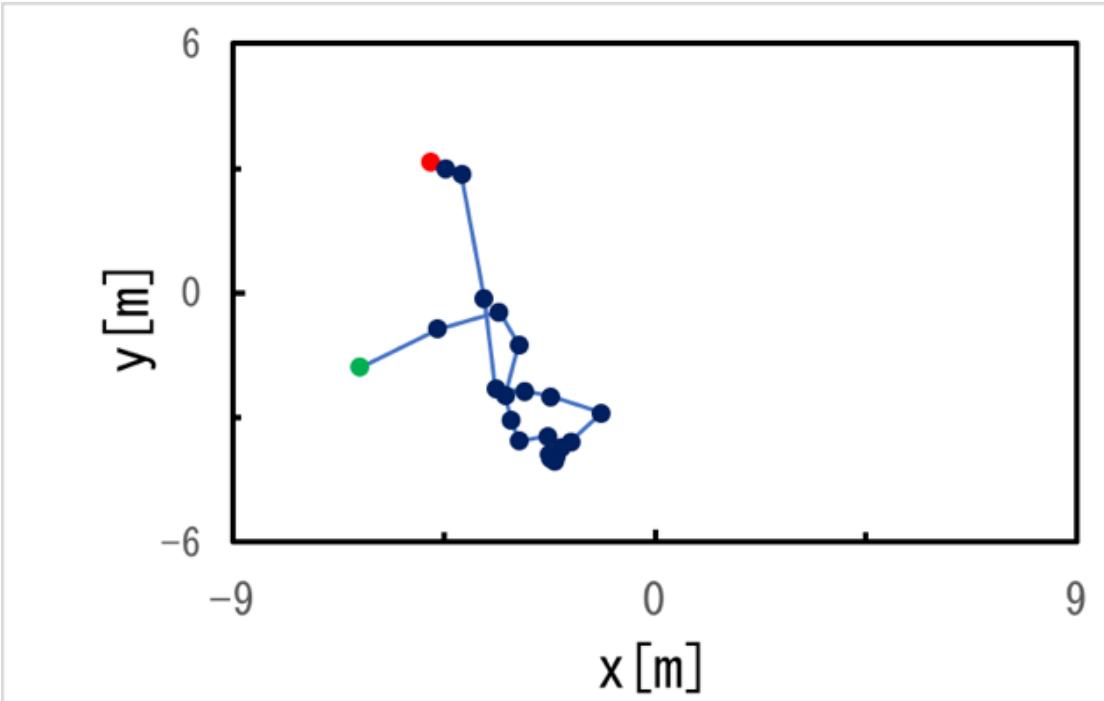


図 33: TRY08 red to blue

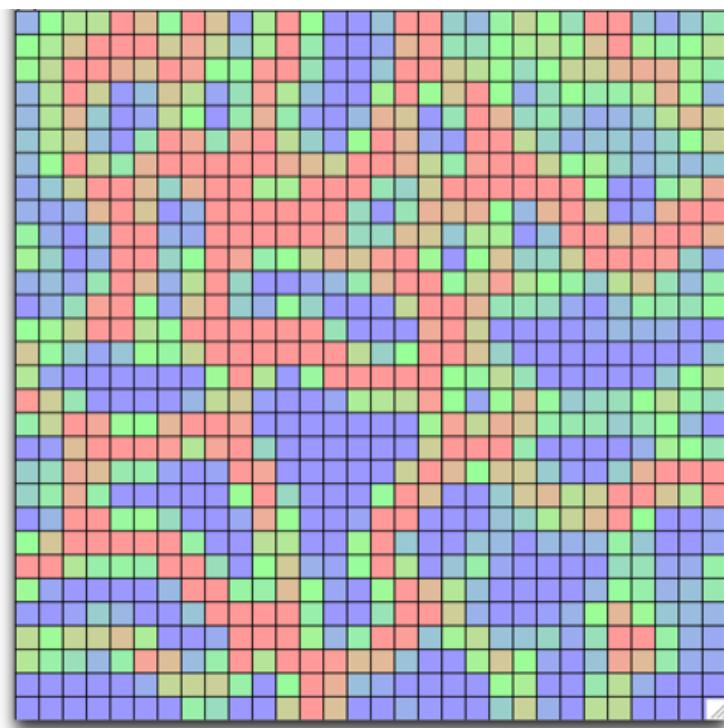


図 34: The feature map expressed by the gradation display (Observation target game B)

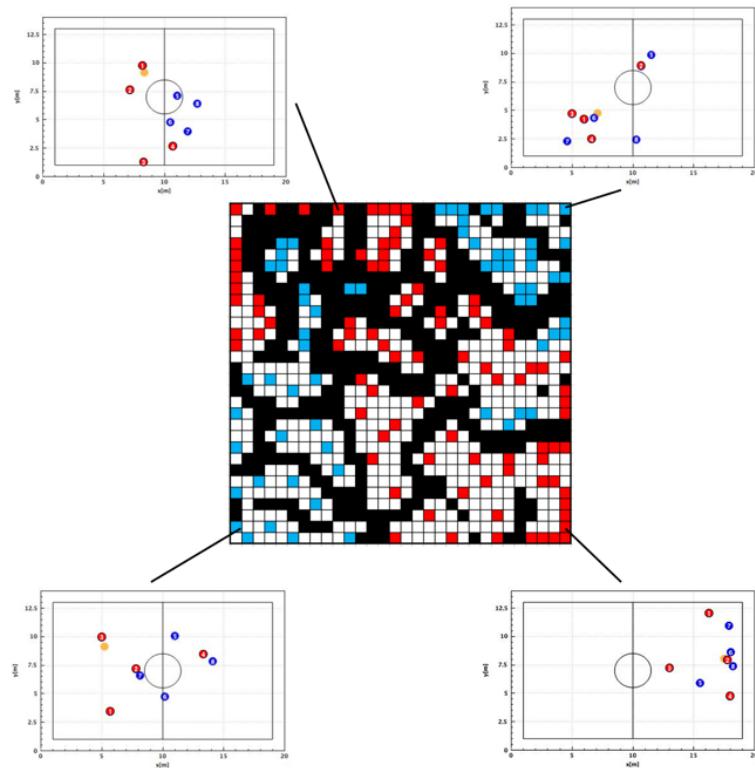


図 35: A color-coded figure with the team try finally chosen a shot(Black : Euclidean distance  $\geq 0.6$ )

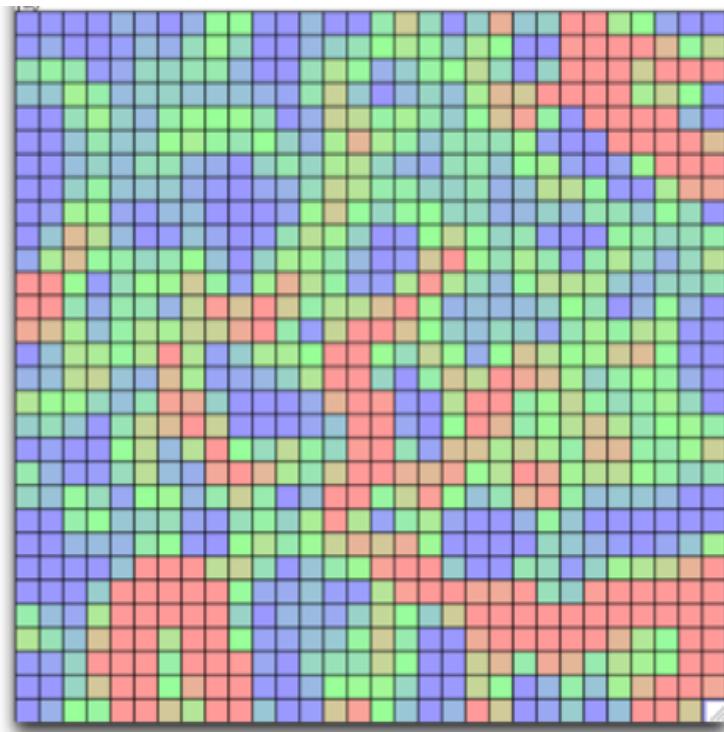


図 36: The feature map expressed by the gradation display (Observation target game C)

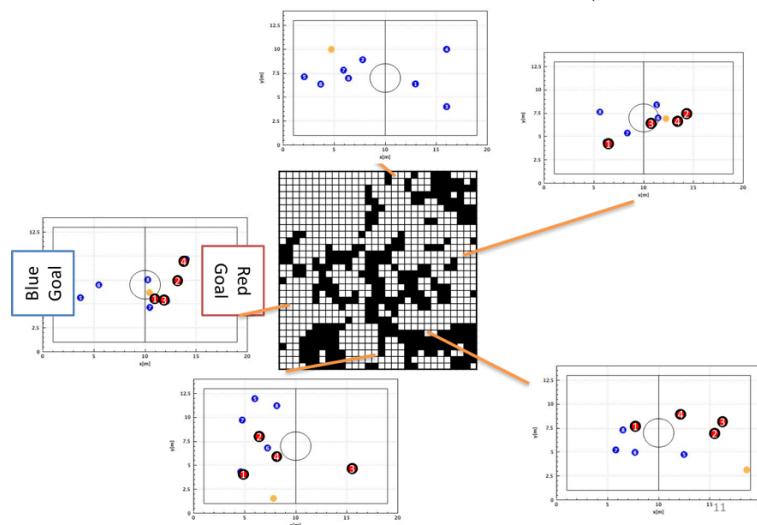


図 37: A color-coded figure with the team try (Black : Euclidean distance  $\geq 0.6$ )

#### 2.4.6 結論

本章では、人とロボットの協調にむけたマルチエージェントゲームの行動快解析を目標として、3つの実験を行なった。全て、人間同士の試合もしくはロボット同士の試合を観察することで、位置座標から得られる試合情報がどのくらいクラスタリング要素として有効かということを調べた。

まず、1つ目の実験では、各選手とボールの4次元での入力であったが、普段からチームとして行動に慣れている選手群と、慣れていない選手群では、大きく異なった。慣れている選手群には、SOMの特徴マップからユークリッド距離によりクラスタリングされた結果から見て取れる境界で、チーム内での選手のボールへの近さも分かっていた。これは、慣れていないチームでの試合との比較をすることにより、より明確化される。人間の試合でも、即席では「チームで協調」することができない場面があることを示している。

次に、同チームの選手群ができる四角形の面積を求めるこことにより、チーム内での選手の疎密度を観察した。すると、試合中において攻める局面と守る局面では、疎密度が真逆になることがわかった。ここまで、実験で、仲間の情報を含んだ場合の選手の行動選択を観察することができた。

最後に、相手選手も含めたSOMの特徴マップを作成することにより、試合全体の観察を行なった。この時は、セットプレーからシュートまでを1トライと換算し、入力した。人間の試合では結果、赤チームがセットプレーからシュートまでボールを奪われずにゴールした時、青チームがセットプレーからシュートまでボールを奪われずにゴールした時、赤チームがセットプレーを始めたが途中で青チームに奪われてしまった時、青チームがセットプレーを始めたが途中で赤チームに奪われてしまった時のすべての場合が、違うクラスタに分類された。これは、ユークリッド距離によるクラスタ配置とも一致しており、セットプレーからシュートまでの1トライが大きなクラスタの一つの中で動いており、試合状況のクラスタリングが可能であることを示している。

一方、ロボットの試合では、人間同様に解析を行ったが、まったくクラスタリングできなかつた。このことから、ロボカップ中型リーグでは、ロボットの協調が人間の協調と大きく異なることを示している。

今後の展望として、このように、クラスタリングできたSOMをロボットに導入することで、現在の試合状況の理解をロボットにされることを考える。

# **第3章**

## **自己組織的チーム行動獲得システム**

## 第3章 自己組織的チーム行動獲得システム

### 3.1 緒言

ロボットやその頭脳としてのコンピュータの優れた能力は、繰り返し処理等のある決まつた作業が容易に可能であることが挙げられる。しかし、状況に応じて変化する要素が含まれる際や、曖昧な情報を含有する際には、その処理能力は未だ人間には及ばない。ロボットに対して柔軟な処理の実装を試みる際には、生物の情報処理メカニズムからヒントを得た考え方、ソフトコンピューティング等の考え方があると考えられている。

自律型ロボットの特徴を生かして大いに活用するためには、システムの管理運営について、効率化を重視し、中枢系を一つに集中する集中管理方式も好ましいが、各エージェント自体の自律性の向上が第一に望まれる。人間と試合可能なゴミの実現のための課題には運動制御、環境認識、自己位置同定などが挙げられる。各エージェントとしてのロボットは得られたセンサ情報を基に、自身の状態の把握に加え、道徳性の同定を行い、環境の変化等も認識した上で、行動を決定することが望ましい。

サッカーロボットが環境を認識するためには、複数のセンサ情報から特徴を見つけ、状態や環境を認識するために必要な情報を抽出する必要がある。本節では自己組織化マップの低次元化及び汎化性能、特徴抽出能力に着目し、SOM を用いたサッカーロボットの環境認識から行動決定システムとして、サッカー試合中の人間による行動決定システムのロボットへの実装を提案する。

### 3.2 テンソル自己組織化マップ (TSOM)

本論文ではニューラルネットワークの中でもテンソル自己組織化マップ (テンソル SOM) を用いて行動学習システム及び行動選択システムの開発を行う。テンソル SOM とは SOM をテンソルデータに拡張したものである。テンソル SOM のアルゴリズムを勝者決定、近傍計算、モデルの更新の 3 つのステップにより簡潔に述べる。<sup>[?]</sup>(38 参照)

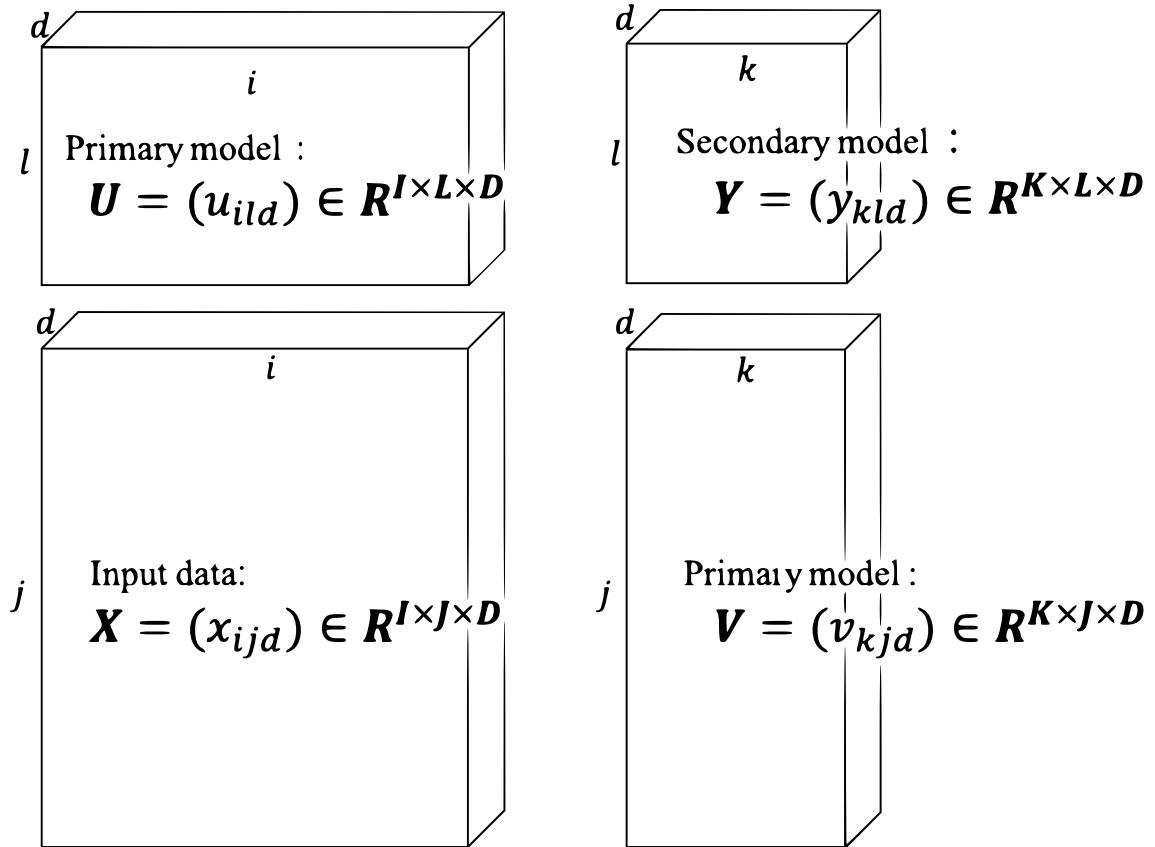


図 38: Data image of tensor SOM

#### 1. 勝者決定

13,14 を用いて各モードの勝者 (Best Matching Unit: BMU) を決定する。ここで  $k_i^*$  はモード 1 に対する BMU 座標であり、 $l_i^*$  はモード 2 に対する BMU 座標である。

$$k_{i^*} = \arg \min_k \sum_{l=1}^L \sum_{d=1}^D (u_{ild} - y_{kld})^2 \quad (13)$$

$$l_j^* = \arg \min_l \sum_{k=1}^K \sum_{d=1}^D (v_{kjd} - y_{kld})^2 \quad (14)$$

## 2. 近傍計算

$15, 16$  を用いて近傍を求める。 $\alpha_{ki}$  は、モード 1 に対する BMU 近傍ユニットの変化量であり、 $\beta_{lj}$  も同様にモード 2 に関する変化量を示している。

$$\alpha_{ki} = \exp\left(-\frac{1}{2\sigma_1^2}\|\zeta_{k_i^*}^{(1)} - \zeta_k^{(1)}\|^2\right) \quad (15)$$

$$\beta_{lj} = \exp\left(-\frac{1}{2\sigma_2^2}\|\zeta_{l_j^*}^{(2)} - \zeta_l^{(2)}\|^2\right) \quad (16)$$

## 3. モデルの更新

$17, 20$  を用いて 1 次モデル、2 次モデルとともに更新する。

- 2 次モデルの更新

$$y_{kld} = \frac{1}{g_k g_l} \sum_{i=1}^I \sum_{j=1}^J \alpha_{ki} \beta_{lj} x_{ijd} \quad (17)$$

$$g_k = \sum_{i=1}^I \alpha_{ki} \quad (18)$$

$$g_l = \sum_{j=1}^J \beta_{lj} \quad (19)$$

- 1 次モデルの更新

$$u_{ild} = \frac{1}{g_l} \sum_{j=1}^J \beta_{lj} x_{ijd} \quad (20)$$

$$v_{kjd} = \frac{1}{g_k} \sum_{i=1}^I \alpha_{ki} x_{ijd} \quad (21)$$

この 3 ステップを近傍半径である  $\sigma$  を時定数  $\tau$  に従って単調に減少させながら繰り返し行う。

### 3.3 TSOM を用いた学習方法検証

環境情報とそのときの取るべく望ましい行動は一種の入出力関係にある。このような入出力関係のある対の情報を一つのベクトルとして、TSOMへ入力することでその関係を得ることが可能である。ここでは、検証実験として、三次元空間のある平面状にある点の座標値を学習させ、 $x$  及び  $y$  の座標値から  $z$  の座標値を推定する実験を行った。

### 3.4 TSOM を用いた環境認識と行動決定

### 3.5 人間と同じ行動判断アルゴリズムの創発

人間と同じ行動決定創発を書き書きします

# **第4章**

## **ロボットの実装実験**

## 第4章 口ボットへの実装実験

### 4.1 緒言

TSOM を用いて行った解析をいかに示す

# **第5章**

## **考察及び結論**

## 第5章 考察及び結論

### 5.1 緒言

### 5.2 今後の展望

まとめるぜ!!!!