

移動体の意思決定のための情報量基準による観測戦略

Observation strategy for decision making based on information criterion

光永 法明 (阪大) 浅田 稔 (阪大)

Noriaki MITSUNAGA, Osaka University, 2-1, Yamadaoka, Suita, Osaka

Minoru ASADA, Osaka University

Abstract: This paper proposes a method of constructing a decision tree and prediction trees of the landmarks that enable a robot with a limited visual angle to localize itself in the environment. Since global positioning from the 3-D reconstruction of landmarks is generally time-consuming and prone to errors, the robot makes decisions depending on the appearance of landmarks. By using the decision and the prediction trees based on information criterion, the robot can achieve the goal efficiently.

Keywords: decision making, decision tree, information criterion, localization

1 はじめに

移動ロボットは様々な場所へ移動し、場所に応じた行動をとることが期待される。そのため意思決定には、場所の認識が重要な役割を果たすと考えられる。場所の表現としては定量幾何学表現あるいは位相幾何学表現(トポロジー)が広く用いられているが、一般には場所を識別可能な他の表現形態も可能と考えられる。実際にはロボットに利用可能なセンサの制約のため、過去の観測を含む複数回の観測の統合手法が必要となることが多い。さらに一般に自己位置の確度を高めるには、多観測を必要とするが、観測は少ないことが望ましい。そこで効率的な観測を行う手法が必要とされている。

一般に用いられるデッドレコニングのモデルは場所に依存せず、一定の行動に一定の移動量に対応させる。しかし歩行による移動の場合、移動距離がすべりなどのため歩行面に依存して大きく変化するので、予め移動量をモデル化することは困難である。

場所を定量幾何学的に表現した際の観測の統合については占有格子に Bayes 推定を適用した研究¹⁾や、拡張カルマンフィルタを用いた研究³⁾がある。前者では移動可能性のある空間全体での確率分布を記憶、計算するが、後者では逐次的に自己位置を推定する。そのため前者は複数の自己位置候補を表現しやすいが記憶計算のコストが高く、後者は記憶、計算のコストが低いが、事前の特徴の分節化を必要とする。

効率的な観測を行うために Moon et al.⁴⁾⁵⁾は障害物を回避するナビゲーション問題で高速に移動するための、場所の確認観測点 (view point) 計画を行っている。出来るだけデッドレコニングを用いて、視覚による場所の確認を減らしている。彼らの方法ではあらかじめ地図と経路が既知と仮定しており、反射的な行動などを行う場合には利用できない。Burgard et al.²⁾は、Bayes 推定を利用した占有格子を計算する自己位置推定の手法を基に、曖昧さを減少させる行動をとる手法を提案している。この手法は自己位置の同定が目的であり、それ以上のタスクは考慮されていない。

これらの手法は場所の記述を何らかの幾何学的表現に基づいて行っている。しかしながら、視覚を持つ移動ロボットを考えた場合、視覚情報から幾何学的な位置を推定するには、計算コストだけでなく、ロボット

自身や環境に関するモデルやパラメータなどの知識が必要となる。一方、視覚情報を直接利用する場合にはそのような知識は必要ない。また観測を位置の計算とそれ以外で区別することなく扱うことができる。そういった特長から、場所の幾何学的表現を利用せず観測の統合にリカレントニューラルネットワークによる予測を用いた研究^{8) 10)}が行われている。これらの研究では観測は受動的であったが、Tani et al.⁹⁾は観測の能動的な切り替えを行っている。部分視覚を持つ移動ロボットが視覚目標追従を行いながら、壁に沿った移動を行う際に、観測の統合はリカレントニューラルネットワークによる見え方の予測で行い、予測精度に応じて観測対象を視覚目標と壁との間で切り替える実験を行っている。この実験では場所によらず、ロボットが必要とする情報は壁と視覚目標物の見え方であり、意思決定に必要な情報は一定である。そのため、状況に応じて観測を効率的に行う手法とはなっていない。

本研究では、移動ロボットの行動決定に関して効率的な観測を行う行動決定法を提案する。情報量基準により生成した決定木による、視覚情報の予測と行動決定を行うことで、意思決定に必要な観測のみを行うことで観測を効率的に行っている。提案手法は、視野角の限られたセンサを持つ移動ロボットが、意思決定に必要な情報の量が変化する環境内で行動する際に特に有効である。

2 タスクと仮定

ロボットや環境、与えられるデータ等に関して以下を仮定する。1) ロボットの視野角が限られており、行動決定に必要な情報が瞬時には得られない。2) ランドマークが配置されており、カメラを振ることで視野を拡大することにより、行動決定に十分な情報が得られる。3) 行動決定に必要な情報は一定ではなく、状況に応じて変化する。4) 決定木作成のため、行動、視覚情報は離散化されており、決定木を生成するために十分なデータが用意されている。

本研究の実験では、歩行移動ロボットを利用し、決定木生成のためのデータを用意するため教示を用いた。以下では2種類の決定木を作成する。一つは行動を決定する木である。もう一つは、視覚情報の予測を行う予測木で各ランドマーク毎に作成するが、生成法は前

者と同じである．各木のノードは論理センサを，ノードをつなぐ枝は論理センサの値を示し，葉は枝をたどったセンサ値の場合の行動あるいは予測される論理センサの値である．また説明を簡単にするため，以下の決定木生成の説明ではセンサ値として各ランドマークの方位のみを扱うがそれ以外も同様に扱うことができる．

ロボットとしては，RoboCup99 SONY 脚式ロボットリーグのロボット（図 1）を用いた．カメラの画角は横 53 度，縦 41 度，画素数はそれぞれ 88, 59 である．脚は各 3 自由度，首は 3 自由度（パン，チルト，ロール）ある．ランドマークを見る際には脚，首のチルト，ロールを固定し，パン軸のみを利用した．パン軸はロボット正面に対して，-90 度から 90 度が可動範囲である．



Fig.1 The SONY legged robot for RoboCup 99 SONY legged robot league.

環境を図 2 に示す．ランドマークは 8 あり，ボールが一つある．それぞれ，敵ゴール (TG)，自陣ゴール (OG)，北西ボール (NW)，北東ボール (NE)，中央西ボール (CW)，中央東ボール (CE)，南西ボール (SW)，南東ボール (SE) とする．すべてのランドマークとボールは色により識別される．ロボットがボールを TG に入れることをタスクとする．これを実現するためには場所に依じたボールへの回り込み，ボールの探索などが必要となる．

3 手法

3.1 決定木，予測木の生成

行動の種類を r ，ランドマークの方位の分割数（見えない場合を含む）を q ，ランドマークの種類数を m ，トレーニングデータの数を n とする．まず各行動 $j = 1, \dots, r$ の生起確率 p_j を求める．行動 j をとった回数を n_j とすると，

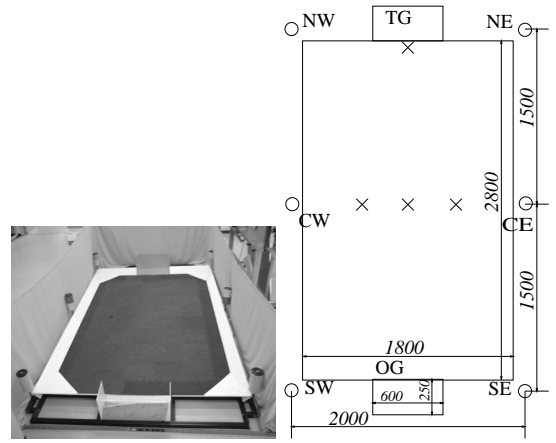
$$p_j = \frac{n_j}{n} \quad (1)$$

このときの p の情報量 I_0 は，

$$I_0 = - \sum_{j=1}^r p_j \log_2 p_j \quad (2)$$

である．次に，ランドマークあるいは視覚目標物の方位がわかった場合の事後生起確率を求める．ランドマーク i が方位 k に観測されたときに行動 j をとった回数を n_{ikj} とすると，事後生起確率は，

$$p_{ikj} = \frac{n_{ikj}}{\sum_{j=1}^r n_{ikj}} \quad (3)$$



(a) Photo of the field. (b) Size of the field

Fig.2 Experimental field (same as the one for RoboCup SONY legged robot league). Cross marks are for the first experiment.

これらを知ったときの，情報量期待値を計算すると，

$$I_i = - \sum_{k=1}^q \left(\frac{\sum_j n_{ikj}}{\sum_k \sum_j n_{ikj}} \right) \sum_{j=1}^r (p_{ikj} \log_2 p_{ikj}) \quad (4)$$

となる．ここで I_0 から最も情報量が減少するランドマークから順に木の上位に置き木を生成する．この情報量による基準は ID3⁶⁾ と同じである．同じ観測で異なる行動をとったトレーニングデータに関しては，それぞれの行動をとった確率を計算し別の葉として生成する．ランドマークの見え方予測木も同様に生成する．

3.2 確率分布の計算

時刻 t での各ランドマーク i がどの方向 k に観測されるかの確率を $p_{ik}^L(t)$ ($i = 1, \dots, m, k = 1, \dots, q$) とし，時刻 t で行動 j をとった確率を $p_j^a(t)$ ($j = 1, \dots, r$) とする．過去の経験による現在とるべき行動が j である確信度を $r_j(t)$ ($j = 1, \dots, r$) とする．

確率分布の計算は次のように行う．現在画像上で観測されているランドマーク i については，その方向 K の確率を $p_{iK}^L(t) = 1$ とし，それ以外を $p_{iK}^L(t) = 0$ ($k \neq K$) とする．1 時刻前の行動については，実際にとった行動 J を $p_J^a = 1$ とし，それ以外を $p_j^a = 0$ ($j \neq J$) とする．観測されていないランドマークについては，1 時刻前のそれぞれの確率分布 $p_{ik}^L(t-1)$ から予測木を使って計算する．カメラを振ることによりランドマーク探索した場合に見えない場合にのみ，見えていない確率を 1 とし，残りの方位を 0 とする．

予測木は次のように用いる．ランドマーク i の木の根から葉までたどると，時刻 $(t-1)$ の各ランドマークの見え方と行動の論理積を満たした場合の，ランドマーク i の時刻 t での見え方が記述されている．そこですべての葉について，論理積をその見え方であった（行動をとった）確率の積に置き換え，その葉に到達する確率を計算を行う．複数の葉に同じ見え方が現れる

act-1-LM0-+1-LM1-1-LM2-1-[LM0=1]
+2-LM1-1-LM2-1-[LM0=2]
+2-LM1-2-[LM0=0] (0.3)
+2-LM1-2-[LM0=2] (0.7)

Fig.3 Example of landmark prediction tree

ので、それらの和を時刻 t での、その見え方をとる確率と見なす。例えば、ランドマーク 0 の予測木が図 3 であれば、

$$\begin{aligned} p_{00}^L(t) &= p_1^a(t-1)p_{12}^L(t-1) \times 0.3 \\ p_{01}^L(t) &= p_1^a(t-1)p_{12}^L(t-1)p_{21}^L \\ p_{02}^L(t) &= p_1^a(t-1)p_{11}^L(t-1)p_{21}^L(t-1) \\ &\quad + p_1^a(t-1)p_{12}^L(t-1) \times 0.7 \end{aligned} \quad (5)$$

となる。

これらから得られた時刻 t でのランドマークの見え方の確率分布 $p_{ik}^L(t)$ をもちいて、行動決定木を同様にたどり、行動の確信度 $r_j^a(t)$ を計算する。

3.3 行動決定

行動の確信度分布の計算後、行動を決定する。確信度分布の中で、ある行動の確信度が、特に高い山になっていればその行動をとればよい。そうでなければ、行動確信度の山が十分に高くなるまで、行動決定木の上から順にランドマークの確率分布を調べ、山が低いランドマークについて再観測を繰り返す。すなわち情報量基準で順に再観測の必要なランドマークを調べる。また再観測の際、ランドマークの確率分布の山(方位)を優先的に調べることで、再観測時間を軽減できると期待される。

4 実験結果

各ランドマークの見え方は、ロボットに対して前方を 0 度として、 $(, -65)$, $[-65, -40)$, $[-40, -15)$, $[-15, 15)$, $[15, 40)$, $[40, 65)$, $[65,)$ の 7 方位に分割し、見えない場合を含めて見え方は 8 通りとした(図 4)。ボールの見え方は、 $(, -45)$, $[-45, -12)$, $[-12, 12)$, $[12, 45)$, $[45,)$ の 5 方位に分割し、さらにロボットから遠い近いの 2 通り(水平に対して下向 30 度で分割)に分け、見えない場合を含めて見え方は 11 通りとした。ボールは、ロボットの行動と一時刻前のボールの見え方だけに依存する特殊なランドマークとして扱い予測木を生成した。

行動決定木や各予測木をたどって確率分布、確信度分布を計算する際、各木を生成する時に含まれなかった見え方により、確率分布の合計 $\sum_{j=1}^N p[j]$ が 1 にならない場合がある。ここでは、合計が 1 となるよう $(1 - \sum_{j=1}^N p[j])/N$ を $p[j](j = 1, \dots, N)$ に加えた。 N は分布の大きさである。

教示再生の際には、行動の確率分布の最大値が 0.6 以上であればその行動をとり、そうでなければ首を振って見えていないランドマークとボールを再探索することとした。教示中、教示再生時ともにランドマーク再探索時以外は、ボールを追跡視あるいは探索するようにした。また、行動確率分布の山が平らなときには、ランドマークの方位の確率分布の山も平らであったため、再探索の際には予測した方位を利用せず、全体を探索した。

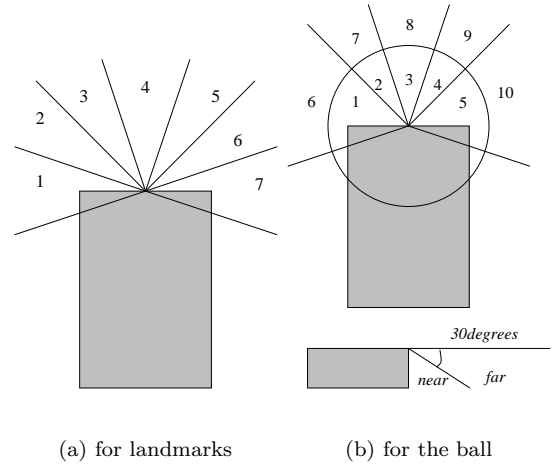


Fig.4 Quantization for landmarks and the ball.

4.1 Experiment 1

まず図 2 のフィールドにおいて、ゴール前(図の×印)にボールをおき、フィールドの中央 3 点(フィールド中央の×印)から開始してボールをゴールに入れるタスクを行った。行動は我々の開発した歩行プログラムを用い、前進、左右大回りの 3 つとした。行動の継続時間は、2.4 秒とした。これは 4 歩行周期で、直進の場合約 0.45m の移動となり、ほぼ一度の行動で見え方が変化する。教示は中央 3 点から各 5 回ずつ行い、80 のデータを得た。このデータから生成した決定木の一部を図 5 に、決定木と予測木の大きさや情報量による順序を表 1-4 に示す。

| # of leaves | min dep. | mean dep. | max dep. |
|-------------|----------|-----------|----------|
| 43 | 1 | 4.91 | 8 |

Table 1 Depth (minimum, mean and maximum) and size of the action decision tree (experiment 1).

| | # of leaves | min dep. | mean dep. | max dep. |
|------|-------------|----------|-----------|----------|
| ball | 52 | 2 | 2 | 2 |
| OG | 13 | 1 | 4.23 | 8 |
| TG | 44 | 1 | 5.39 | 8 |
| SE | 6 | 1 | 2 | 3 |
| SW | 1 | 0 | 0 | 0 |
| CE | 28 | 2 | 4.69 | 8 |
| CW | 11 | 1 | 3.91 | 8 |
| NE | 51 | 1 | 5.96 | 8 |
| NW | 54 | 2 | 5.91 | 8 |

Table 2 Depth of the prediction trees(experiment 1).

教示再生を行った場合の予測と実際にとった行動例を次に示す。まずフィールド中心から開始した場合に、

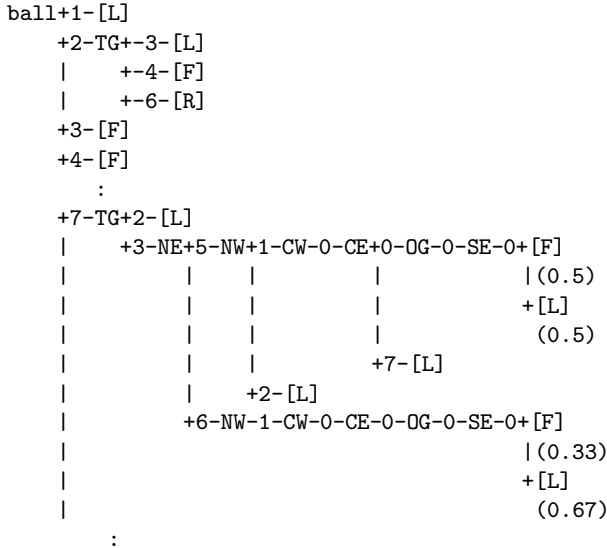


Fig.5 The action decision tree (experiment 1). F, L, and R mean forward, left forward, and right forward respectively.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------|----|----|----|----|----|----|----|
| ball | TG | NE | NW | CW | CE | OG | SE |

Table 3 The order of information for the action decision tree (experiment 1).

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------|------|-----|----|----|----|----|----|----|
| ball: | ball | act | | | | | | |
| OG: | act | NE | TG | NW | CW | CE | OG | SE |
| TG: | TG | act | NE | NW | CE | OG | CW | SE |
| SE: | act | CE | NE | OG | NW | TG | CW | SE |
| SW: | - | | | | | | | |
| CE: | act | NE | TG | CE | NW | CW | OG | SE |
| CW: | TG | act | NE | NW | CE | CW | OG | SE |
| NE: | NE | act | NW | TG | CE | CW | OG | SE |
| NW: | act | NE | TG | NW | CE | OG | SE | CW |

Table 4 The order of information for prediction trees (act means action, experiment 1).

ロボットは、前進 前進 前進 前進という行動をとった。このときボールと TG は常に見えており、それ以外のランドマークの予測木をつかった確率分布と過去にとった行動の確率分布は図 6 のようになった。

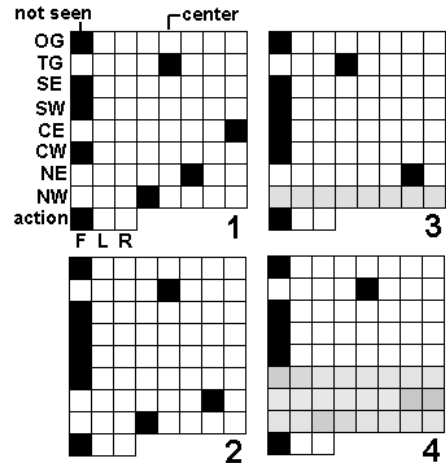


Fig.6 Probability distribution in experiment 1-1 (The height of the black box indicates probability).

同じフィールド中心から開始した場合でも、別の行動をとることもあった。これは、初期のロボットからのランドマークの見え方や歩行の結果が必ずしも同一ではないからである。この例では、初期のランドマークの見え方は一致しているが、歩行の結果が一致しなかった。ロボットは、前進 前進 ランドマーク確認 前進 ランドマーク確認 左回転 前進 前進という行動をとった。ボールと TG は常に見えており、それ以外のランドマークの予測木をつかった確率分布と行動決定木の分布は図 7 のようになった。

フィールドの右側から開始した場合の行動例を次に示す。ロボットは、左回転 ランドマーク確認 前進 前進 左回転という行動をとった。ボールと TG は常に見えており、それ以外のランドマークの予測木をつかった確率分布と行動決定木の分布は図 8 のようになった。

4.2 Experiment 2

同じフィールド上で RoboCup 99 の試合を考慮した教示を行った。教示の負担を減らすため、行動は、前進、左右大回り、左右小回り、ボール追跡の 6 つとした。行動の継続時間は、前の実験と同じ 2.4 秒とした。

この教示により、1364 のデータを得た。このうち不適切な教示を除き、856 を行動決定木の生成に、1364 すべてを予測木の生成に用いた。各木の大きさ、情報量による順序を表 5-8 に示す。

このデータを実際に、RoboCup 99 で用いたところ、ロボットは教示者の期待した行動を行った。しかしながらランドマークの確認が頻繁に行われた。これはランドマークの予測木生成に用いたデータが少なかったためと思われる。

5 まとめ

experiment 1 を見ると、行動確率の分布はほぼ 1 の高い山があるか、ほとんど平らな確率分布かのいずれかとなっている。行動確率分布が平らな場合、情報量

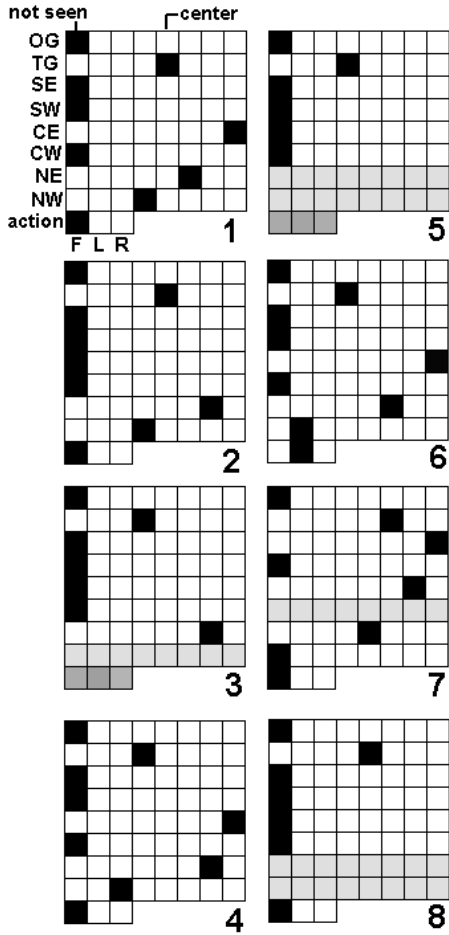


Fig.7 Probability distribution in experiment 1-2.

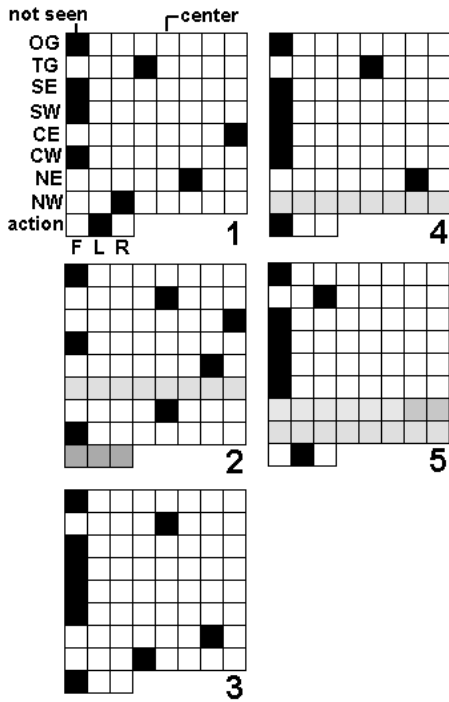


Fig.8 Probability distribution in experiment 1-3.

| # of leaves | min dep. | mean dep. | max dep. |
|-------------|----------|-----------|----------|
| 586 | 2 | 5.89 | 9 |

Table 5 Depth and size of the action decision tree (experiment 2).

| | # of leaves | min dep. | mean dep. | max dep. |
|------|-------------|----------|-----------|----------|
| ball | 403 | 2 | 2 | 2 |
| OG | 958 | 2 | 7.58 | 9 |
| TG | 1050 | 2 | 7.67 | 9 |
| SE | 845 | 2 | 7.35 | 9 |
| SW | 901 | 2 | 7.41 | 9 |
| CE | 901 | 2 | 7.13 | 9 |
| CW | 873 | 2 | 7.37 | 9 |
| NE | 1031 | 2 | 7.60 | 9 |
| NW | 980 | 2 | 7.55 | 9 |

Table 6 Depth and size of the prediction trees (experiment 2).

| | | | | | | | | |
|------|----|----|----|----|----|----|----|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| ball | TG | OG | SW | SE | NW | NE | CE | CW |

Table 7 The order of information for the action decision tree (experiment 2).

| | | | | | | | | |
|-------|------|-----|----|----|----|----|----|--------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| ball: | ball | act | | | | | | |
| OG: | OG | SE | SW | TG | NW | CW | NE | CE act |
| TG: | TG | OG | SE | SW | NW | NE | CW | CE act |
| SE: | SE | OG | TG | SW | CE | NE | NW | CW act |
| SW: | SW | OG | CW | SE | TG | NW | NE | CE act |
| CE: | CE | SE | OG | TG | NE | SW | NW | act CW |
| CW: | CW | SW | OG | TG | NW | SE | NE | CE act |
| NE: | TG | NE | OG | SE | CE | NW | SW | CW act |
| NW: | NW | TG | OG | SW | CW | SE | NE | CE act |

Table 8 The order of information for prediction trees (experiment 2).

の高いランドマークの確率分布も平らになっている。このためランドマークの山を優先的に見ることはできない。これは、予測木を使うのに用いたトレーニングデータが少なかったためと思われる。そのため、行動確率が低い際にはランドマークを見直すという戦略は正しかったと思われる。

experiment 1, experiment 2 のランドマーク予測木を比較すると, experiment 1 では行動が上位に来ており, experiment 2 では行動が下位になっている。experiment 1 では教示データが少ないため, 場所に依存して見え方の異なるランドマークの予測は, ランドマークから行った方がよいという結果がでなかったと思われる。

ここでは, 決定木の圧縮は積極的には行っていない。圧縮を行うとトレーニングデータにない状況への対応が期待される一方, ランドマークの再確認を行うべき状況において確認を行わないことが増加すると思われる。また各ランドマーク, ボールの見え方の離散化方法を, ここではあらかじめ決定している。しかし今回利用した離散化法が最適であるとはいえない。行動決定木の生成に C4.5⁷⁾ などの連続値を扱え情報量基準を用いた離散化を同時に行える手法を用いることで, 自律的な離散化を行える可能性がある。

一度見渡せば場所に関して十分な情報を得られると仮定しているが, 仮定の成立しない場合への対処が必要である。また, 行動中の注視対象をボールに限ったが, これも情報量基準で選択することが望ましい。またデータ収集は教示により行ったが, 教示の負担も大きく, 負担の軽いデータ収集法が望まれる。

参考文献

- [1] W. Burgard, A. Derr, D. Fox, and A. B. Cremers. Integrating global position estimation and position tracking for mobile robots: The dynamic mrkov localization approach. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 730–735, 1998.
- [2] W. Burgard, D. Fox, and S. Thrun. Active mobile robot localization. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI)*. Morgan Kaufmann, San Mateo, CA, 1997.
- [3] M. W. M. G. Dissanayake, P. Newman, H. F. Durrant-Whyte, S. Clark, and M. Csorba. An experimental and theoretical investigation into simultaneous localisation and map building. In *Preprints for 6th International Symposium on Experimental Robotics*, pp. 171–180, 1999.
- [4] I. H. Moon, J. Miura, Y. Yanagi, and Y. Shirai. Planning of vision-based navigation for mobile robot under uncertainty. In *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 1, pp. 1202–1207, 1997.
- [5] I. Moon, J. Miura, and Y. Shirai. Dynamic motion planning for efficient visual navigation under uncertainty. In Y. Kakazu, M. Wada, and T. Sato eds., *In Proc. of the Intelligent Autonomous Systems 5*, pp. 172–179, 1998.
- [6] J. R. Quinlan. Discovering rules from large collections of examples: a case study. In D. Michie ed., *Expert Systems in the Microelectronic Age*. University Press, Edinburgh, Scotland, 1979.

- [7] J. R. Quinlan. *C4.5: PROGRAMS FOR MACHINE LEARNING*. Morgan Kaufmann Publishers, 1993.
- [8] J. Tani. Model-based learning for mobile robot navigation from the dynamical systems perspective. *IEEE Trans. on System, Man and Cybernetics Part B (Special Issue on Robot Learning)*, 26(3):421–436, 1996.
- [9] J. Tani, J. Yamamoto, and H. Nishi. Dynamical interactions between learning, visual attention, and behavior: An experiment with a vision-based mobile robot. In P. Husbands and I. Harvey eds., *Fourth European Conference on Artificial Life*, pp. 309–317. The MIT Press, 1997.
- [10] 鈴木, 横井, 嘉数. 連想記憶を用いたロボットナビゲーションシステムに関する基礎研究. 第17回日本ロボット学会学術講演会講演論文集, pp. 613–614. 日本ロボット学会, 1999.