

移動体の意思決定のための情報量基準に基づく観測対象選択戦略

光永 法明* 浅田 稔*

Observation strategy of landmark selection for decision making of legged robot based on information criterion

Noriaki Mitsunaga* and Minoru Asada*

This paper proposes a method for constructing a decision tree and prediction ones of the landmarks that enable a robot with a limited visual angle to make decisions without self-localization 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 task efficiently.

Key Words: 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的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、ロボットに利用可能なセンサの制約のため、複数回の観測の統合が必要となることが多い。定量幾何学的な表現の場合の観測の統合には占有格子に Bayes 推定を適用した研究 [1] や、カルマンフィルタを用いた研究 [2] 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、事前の特徴の分節化を必要とする。

一方、場所を区別する必要はあるが、必ずしも通常の幾何学的位置の表現形態をとる必要がない場合がある。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位置を再構成するには、計算コストだけでなく、ロボット自身と環境についてのモデルやパラメータといった知識が必要となる。幾何学的表現を用いずに、視覚情報をそのまま用いる場合には、そういった知識は必要ない。また位置の計算のための観測とそれ以外の意思決定に必要な観測を区別することなく扱うことができる。そのような特長から、状態の表現としてセンサ情報そのまま利用し、センサ情報とコンテキストベクトルからなるリカレントニューラルネットワークを用いて、次のセンサ情報(場所)を予測する研究 [5] [6] などが行われている。

一般に観測が多いほど自己位置の確度は高くなるが、必要となる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ための手法が提案されている。幾何学的な場所の表現を用いた手法としては、文ら [3], Burgard et al. [4] の研究があげられる。文ら [3] は障害物を回避するナビゲーションの問題で、高速に移動するため、場所の確認観測点 (view point) 計画を行っている。できるだけデッドレコニングを用い、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がわかっている必要があり、反射的な行動などを行う場合には利用できない。Burgard et al. [4] 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、曖昧さの減少する行動をとる手法を提案している。この手法は占有格子の計算コストが高い。また、自己位置の同定が目的であり、それ以降のタスクは考慮されていない。

幾何学的表現を用いない手法では、Tani et al. [7] が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実験で、予測精度に応じて視覚目標と壁との間で観測対象を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。

また、これらの研究でも用いられているデッドレコニングのモデルは、一般に場所に依存せず、一定の行動には一定の移動量に対応させるものである。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へ依存して大きく変化し、移動量のモデルは場所に応じたものが必要である。そのため、あらかじめ移動量のモデルを用意することは困難である。

原稿受付 2000年9月21日

*大阪大学大学院工学研究科

*Graduate School of Engineering, Osaka University



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

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い、観測を効率化する。幾何学的な場所の表現は用いず、あらかじめデッドレコニングモデルを用意することは行わない。提案手法は、視野角の限られたセンサを持ち、意思決定に必要な情報の量が変化する環境内を行動する移動ロボットに有効である。

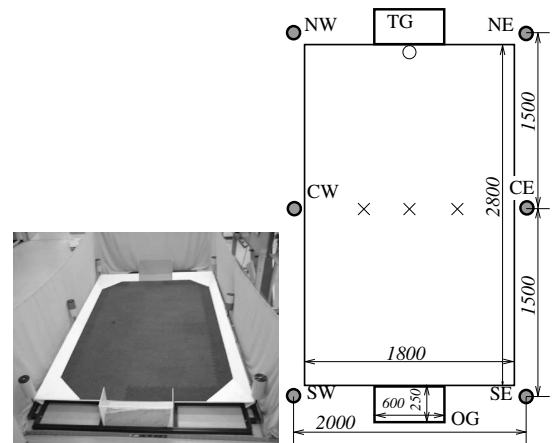
以下では、まずタスクと仮定について述べ、提案手法について説明する。そして実機を利用した実験結果を示し、最後にまとめと今後の課題を述べる。

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ることが少ない。2) ランドマークが配置されており、カメラを振り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、情報量の計算と決定木の生成に十分なデータが用意されている。本研究の実験では、4) のデータを用意するため教示を用いた。以下、このデータをトレーニングデータと呼ぶ。また説明を簡単にするため、手法の説明においてセンサ情報としてランドマークの方位のみを扱うが、それ以外のセンサ情報も同様に扱うことができる。

ロボットとしては、RoboCup 99 SONY 脚式ロボットリーグ [9] のロボット (Fig. 1) を用いた。カメラはロボットの頭部の鼻に当たる位置に設置されており、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、カメラは 3 自由度 (パン, チルト, 視線周り) あるが、ランドマークを観測する際には脚の角度と、カメラのチルト, 視線周りの角度を固定し、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。

環境を Fig. 2 に示す。ランドマークは 8 個あり、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



(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 and circle marks are for the first experiment.

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応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

以下では、まずランドマークの観測基準としての情報量を用いることにより、より早く行動決定の曖昧さを減少することができることを示す。次にトレーニングデータを圧縮して記憶するために、情報量を基準とした分類木の手法を用いて行動決定木を生成すれば、木の上でのランドマークの順序と情報量を基準にした観測順序が一致することを示す。そして観測戦略に必要な確率計算を示し、行動決定法を述べる。

観測戦略は次のようになる。各行動をとるべき確率を行動確率と呼び、いずれかの行動確率が十分に高くなるまで、情報量が大きいランドマークから観測する。さらに、各ランドマークがある方位に観測される確率を観測確率と呼び、一時刻前の観測確率と、行動から計算する。計算には行動決定木の生成と同じ手法を用いて生成した予測木を用いる。

3.1 観測基準としての情報量

行動の種類を r , ランドマークの方位の分割数 (観測されない場合を含む) 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。各トレーニングデータは、カメラを振った際に観測される各ランドマークの方位と、そのときにとるべき行動を含む。まず各行動 $k = 1, \dots, r$ のトレーニングデータ中での生起確率 p_k を求める。行動 k をとった回数を n_k とすると、

$$p_k = \frac{n_k}{n} \quad (1)$$

となる．このときの各行動の生起確率の曖昧さを表す量であるエントロピー H_0 は，

$$H_0 = - \sum_k p_k \log_2 p_k \quad (2)$$

である．次に，ランドマーク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 j に観測されたときに行動 k をとった回数を n_{ijk} とすると，

$$p_{ijk} = \frac{n_{ijk}}{\sum_k n_{ijk}} \quad (3)$$

となる．このときの行動に関するエントロピー h_{ij} は，

$$h_{ij} = \sum_k (p_{ijk} \log_2 p_{ijk}) \quad (4)$$

である．ランドマーク i が方位 j に観測される統計的確率 p_{ij}^l は n_{ijk} を使って，

$$p_{ij}^l = \frac{\sum_k n_{ijk}}{\sum_j \sum_k n_{ijk}} \quad (5)$$

と表せる．したがって，ランドマーク i の方位を知ることにより期待されるエントロピー H_i と情報量 I_i を計算すると，

$$H_i = - \sum_j p_{ij}^l h_{ij} \quad (6)$$

$$I_i = H_0 - H_i \quad (7)$$

となる．すなわち，期待される情報量が大きいランドマーク ($\arg \max_i I_i$) から観測することで，より早くとるべき行動の曖昧さを減少させ行動を決定することができる．

3.2 決定木，予測木の生成

与えられたデータをいくつかのクラスに分類するためのコンパクトな分類木を生成する方法として ID3 [8] がある．分類木を生成するためには，トレーニングデータセットが必要であり，各データは分類先のクラスと分類のための属性の離散値からなる．属性 i について知ったときのクラスに関する情報量 I_i をすべての属性について計算する．最も情報量の大きい属性について，その属性値によってトレーニングデータセットを分割する．分類木としては，この属性をノードとし，属性値によって枝分かれさせる．新しく分割されたそれぞれのデータセットについて，情報量を計算し，全ての属性について情報量が0になるまで，データセットの分割を繰り返すことにより分類木を生成する．その結果，分割後のデータセットに一つのクラスのデータのみが含まれ，その枝の先には，そのクラスを示す葉をおく．

例えば，トレーニングデータセットが Table1 の場合には，クラスに関する情報量を計算すると属性 A は 0.5, 属性 B は 1.2, 属性 C は 1.0 となる．そこで属性 A の値 a, b によってトレーニングセットを No. 1,2 と No. 3, 4 に分割する．No. 1,2 のデータセットはクラス x のみであるので，情報量は0である．No. 3,4 のデータセットにおいて，クラスに関する情報量を計算すると属性 B は 1, 属性 C は 0 となるので，属性 C によりデータセットを分割し，木の生成を終える．出来上がった木は Fig.3 となる．

Table 1 Example training data of ID3.

No.	attribute A	attribute B	attribute C	class
1	a	b	a	x
2	a	a	b	x
3	b	a	a	y
4	b	a	b	z

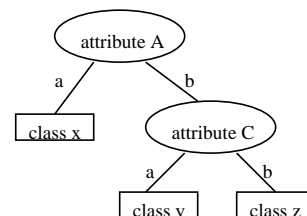


Fig. 3 Example classificatoin tree by ID3.

Table 2 An example of teaching data.

Landmark A	Landmark B	Landmark C	action
a	a	a	x
b	a	a	y
c	b	a	y
a	c	a	z
a	b	b	z

ここで，分類先のクラスをとるべき行動に，属性を各ランドマークの方位とすることにより，行動決定木を生成する．このように生成された決定木では，木の根から葉に向けてノードは，行動決定に関する情報量の大きいランドマークから順に並んでいる．したがって，観測を情報量の順に行うことは，木の根に近いランドマークから観測することに等しくなる．ランドマーク i の予測木は，一時刻後のランドマーク i の方位を分類先クラス，現時刻での各ランドマークの方位と，行動を属性とすることにより生成する．

実際のトレーニングデータには，センサ値を離散化しているため，同じ観測で異なる行動をとったデータが含まれる場合がある．その場合に関しては，それぞれのトレーニングデータ中で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．ランドマーク予測木の生成に関しても同様である．

Table 2のトレーニングデータが得られた場合の計算は次のようになる．まず， $p_x = 0.2, p_y = 0.4, p_z = 0.4$ より， $H_0 = 1.52$ となる． p_{ijk} を計算し，情報量を計算すると $I_A = 0.97, I_B = 0.72, I_C = 0.32$ となるので，決定木でのランドマークの順位を A, B, C とし，木を生成すると Fig.4 となる．

3.3 確率の計算

時刻 t での各ランドマーク i がどの方位 j に観測されるかの確率(観測確率)を $p_{ij}^L(t)$ ($i = 1, \dots, m, j = 1, \dots, q$) とし，時刻 t で行動 k をとった確率を $p_k^a(t)$ ($k = 1, \dots, r$) とする．過去の経験から時刻 t でのとるべき行動が k である確率(行動確率)を $\hat{p}_k^a(t)$ ($k = 1, \dots, r$) とする．

確率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，その方位 J の確率を $p_{iJ}^L(t) = 1$

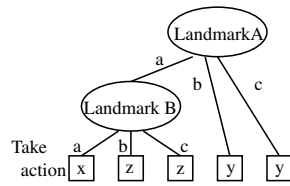


Fig. 4 An example of an action decision tree.

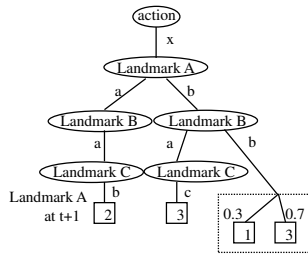


Fig. 5 An example of a prediction tree for landmark A.

とし、それ以外を $p_{ij}^L(t) = 0 (j \neq J)$ とする。1 時刻前の行動については、実際にとった行動 K を $p_K^a = 1$ とし、それ以外を $p_k^a = 0 (k \neq K)$ 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 $p_{ij}^L(t-1)$ 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も観測されない場合にのみ、観測されていない確率を 1 とし、残りの方位に観測される確率を 0 とする。

予測木からの観測確率計算は次のように行う。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方位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 での方位が記述されている。そこですべての葉について、論理積をその方位であった (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の計算を行う。複数の葉に同じ方位が現れるので、それらの和を時刻 t での、その方位となる確率と見なす。例えば、ランドマーク A の予測木が Fig.5 であれば、

$$\begin{aligned} p_{Aa}^L(t) &= p_x^a(t-1)p_{Ab}^L(t-1)p_{Bb}^L(t-1) \times 0.3 \\ p_{Ab}^L(t) &= p_x^a(t-1)p_{Aa}^L(t-1)p_{Ba}^L(t-1)p_{Cb}^L(t-1) \\ p_{Ac}^L(t) &= p_x^a(t-1)p_{Ab}^L(t-1)p_{Ba}^L(t-1)p_{Cc}^L(t-1) \\ &\quad + p_x^a(t-1)p_{Ab}^L(t-1)p_{Bb}^L(t-1) \times 0.7 \quad (8) \end{aligned}$$

となる。

これらから得られた時刻 t でのランドマークの方位の観測確率 $p_{ij}^L(t)$ を用いて、行動決定木を同様にたどり、行動確率 $\hat{p}_k^a(t)$ を計算する。

3.4 行動決定

行動確率の計算後、行動を決定する。ある行動の確率が、特に高ければその行動をとればよい。そうでなければ、特定の行動確率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観測確率を調べ、方位の曖昧な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、ランドマークの観測

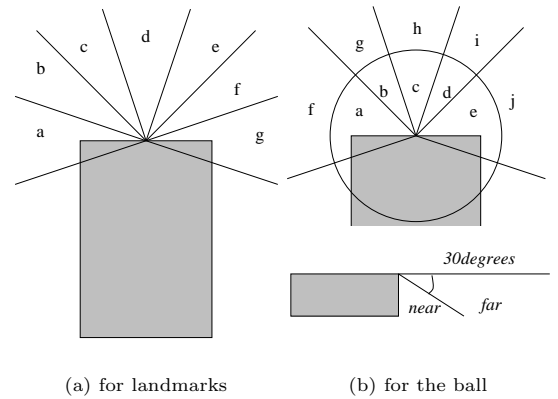


Fig. 6 Quantization for landmarks and the ball.

確率の高い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

各ランドマークの方位は、ロボットに対して前方を 0 度として、 (-65°) , $[-65^\circ, -40^\circ)$, $[-40^\circ, -15^\circ)$, $[-15^\circ, 15^\circ)$, $[15^\circ, 40^\circ)$, $[40^\circ, 65^\circ)$, $[65^\circ,)$ の 7 方位に分割し、観測されない場合 (o で表す) を含めて方位は 8 通りとした (Fig. 6 参照)。ボールの方向は、 (-45°) , $[-45^\circ, -12^\circ)$, $[-12^\circ, 12^\circ)$, $[12^\circ, 45^\circ)$, $[45^\circ,)$ 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、観測されない場合 (o で表す) を含めて方向は 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方向のみ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を計算する際、各木を生成する時に含まれなかった方位により、確率の合計 $\sum_{i=1}^N p_i$ 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 $(1 - \sum_{i=1}^N p_i)/N$ を $p_i (i = 1, \dots, N)$ に加えた。 N は行動の種類数あるいは方位の分割数である。

生成された木を使って行動を決定する (以下教示再生と呼ぶ) 際には、行動確率の最大値が 0.6 以上であればその行動をとり、そうでなければカメラを振って観測され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。

4.1 実験 1

まず Fig.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 3 つの × 印) から開始してボールをゴールにい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、4.8 秒とした。これは 8 歩行周期で、直進の場合約 0.45[m] の移動となり、ほぼ一度の行動で観測が変化するからである。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 Fig.7 に示す。ボールが左前方 (b の方向) に観測された場合に

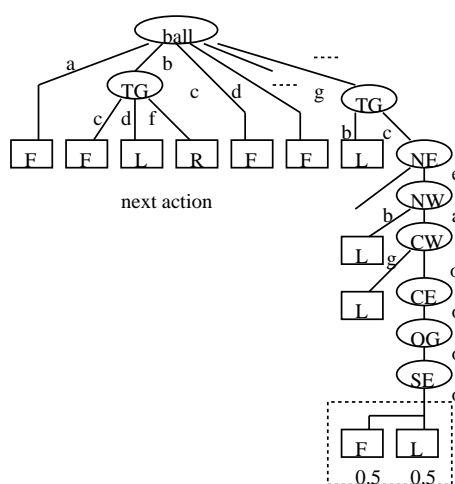


Fig. 7 A part of the action decision tree (Experiment 1). F, L, and R mean the action forward, left forward, and right forward respectively.

Table 3 Depth of the prediction trees(Experiment 1).

tree for	# 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

は、回り込みのため敵ゴールの方位(c,d,f)に応じてとるべき行動が変化することが分かる。決定木と予測木の大きさや情報量による順序を Table 3 から 5 に示す。行動決定木の大きさは、葉の数が 43、最小深さは 1、平均深さは 4.91、最大深さは 8 となった。Table 3は予測木の大きさや深さを示し、表中 # of leaves は葉の数を表し、min dep., mean dep., max dep. は、それぞれ木の最小深さ、平均深さ、最大深さを表す。進行方向と逆であり、小さく観測されることの少ない SW, SE などのランドマークの予測木は小さくなっている。Table 4 (Table 5) は行動決定木(各予測木)における、ランドマークなどの情報量による順位を表す。1が最も木の根に近く情報量が大い、8が最も根から遠く情報量が小さい。Table 4, 5 でball は一時刻前のボールの方向を、act は一時刻前の行動を、TG, OG, NW, NE, CW, CE, SW, SE はそれぞれのランドマークの一時刻前の方位である。行動決定木の最上位にボールが、予測木の比較的上位に行動が現われている。これらの木を使った教示再生中の動きを Fig.8に示す。

次に教示再生を行った場合の予測と実際にとった行動例を示す。まずフィールド中心から開始した場合に、ロボットは、1) 前進 2) 前進 3) 前進 4) 前進という行動をとった。ボールと TG は常に観測されており、各行動決定時の観測確率と行動

Table 4 The order of information for the action decision tree (Experiment 1).

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

Table 5 The order of information for prediction trees (Experiment 1).

tree for	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

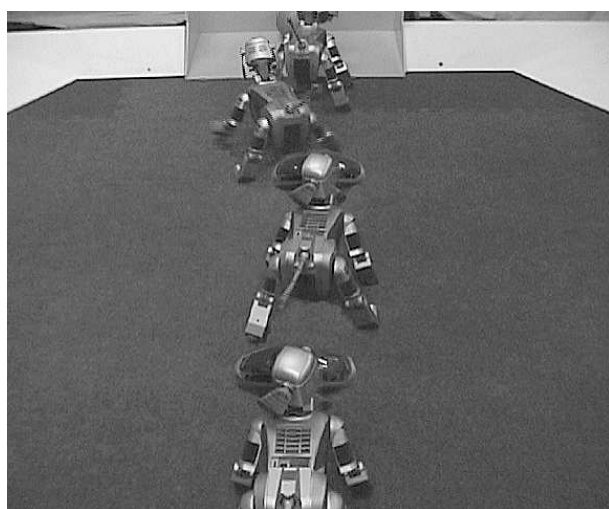


Fig. 8 Robot movements with the action and decision trees.

確率は Fig.9のようになった。各マス目の濃度は確率を表し、白が 0、黒が 1 である。ball はボールの方向の、OG, TG, SE, SW, CE, CW, NE, NW はそれぞれのランドマークの方位の予測木あるいは実際の観測による観測確率であり、not seen はカメラを振っても観測されない場合の確率を、それ以外はボールは下段の左から a,b,...,e の上段の左から f,g,...,j の確率を、ランドマークは左から順に方位 a,b,...,g の確率を表している。action は行動決定木による行動確率であり、F, L, R はそれぞれ。前進、左大回り、右大回りをとるべき確率を表している。行動開始時はカメラを振って観測を行っている。Fig.9から、時刻 3, 4 ではいくつかのランドマークの観測確率が平均して低いものになっているが、カメラを振ることなく行動決定を行っている。

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、初期位置のロボットからのランドマークの観測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観測された方位は一致してい

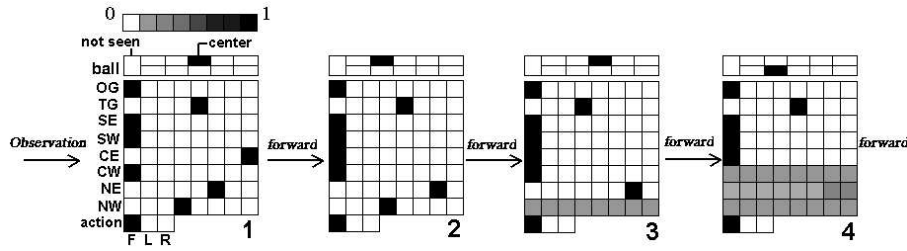


Fig. 9 Probability distribution in Experiment 1-1 (The gray level of each box indicates the probability, black 1 and white 0) .

Table 6 The number of needed observation (Experiment 1).

begin from	# of trials	# of total steps	# of re-observations	rate of re-observation
center	12	35	18	.51
left	12	31	15	.48
right	16	64	38	.59

るが、歩行の結果が一致しなかった。ロボットは、1) 前進 2) 前進 3) ランドマーク確認 4) 前進 5) ランドマーク確認 6) 左回転 7) 前進 8) 前進という行動をとった。ボールとTGは常に観測されており、観測確率と行動確率はFig.10のようになった。時刻3, 5では行動決定ができず、首を振った観測を行う必要が出ている。一方、時刻7, 8では観測確率が曖昧なランドマークがあっても行動決定が行われている。

次にフィールドの右側から開始した場合の行動例を示す。ロボットは、1) 左回転 2) ランドマーク確認 3) 前進 4) 前進 5) 左回転という行動をとった。ボールとTGは常に観測されており、観測確率と行動確率はFig.11のようになった。時刻2では行動確率が曖昧なため行動決定ができず、時刻4, 5では観測確率が曖昧なランドマークがあっても行動決定が行われている。

教示再生中の再観測回数を Table 6 に示す。左から順に試行開始場所、試行回数、合計行動回数、合計再観測回数、再観測率である。再観測回数が半分程度に減少していることが分かる。

4.2 実験 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4.8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 Table 7から 10に示す。ランドマーク毎の予測木の平均深さの差が少なくなっていること、行動決定木の上位二つは変わらないこと、予測木の最上位がほぼ予測対象のランドマークとなっていることなどが分かる。

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がうまく働かなかったためと思われる。

Table 7 Depth and size of the action decision tree (Experiment 2).

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

Table 8 Depth and size of the prediction trees (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 9 The order of information for the action decision tree(Experiment 2).

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

Table 10 The order of information for prediction trees (Experiment 2).

tree for	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

5. 討 論

実験 1 の結果 (Fig.9-11) から、行動確率はほぼ 1 の行動があるか、ほぼ一様な確率となるかのいずれかとなっている。行動確率が一様な場合、観測確率の低いランドマークの確率も平均して低くなっている。このため観測確率の高い方位を優先的

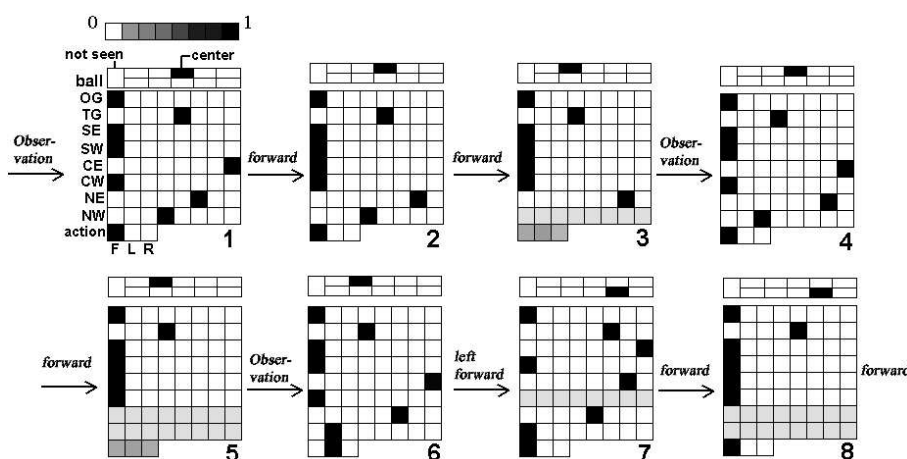


Fig. 10 Probability distribution in Experiment 1-2.

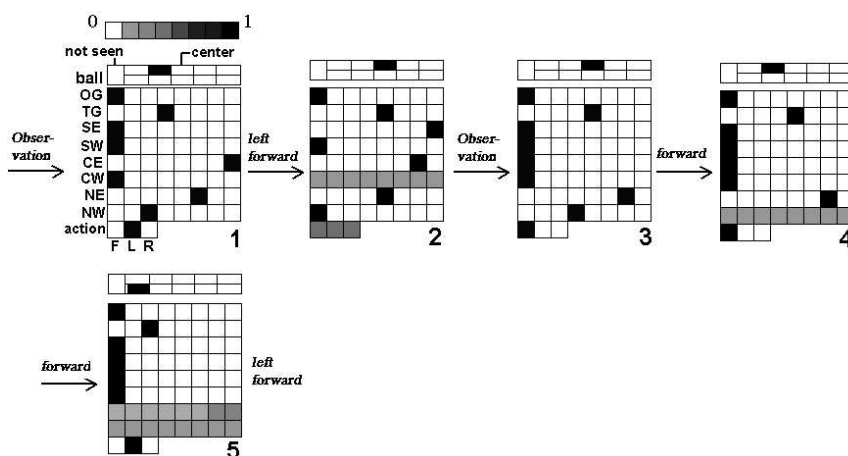


Fig. 11 Probability distribution in Experiment 1-3.

に観測することはできていない。また、行動確率が低かったため、カメラを振り観測を行った前後の観測確率を比較すると、確率が平均して低かったランドマークの確率が変化だけでなく、より情報量の大きいランドマークの確率も変化している。原因としては、予測木の作成に用いたトレーニングデータが少なかった、あるいは予測木の汎化能力が不足している、またランドマークの予測には分類木は向かないといった問題が考えられる。

実験 1, 実験 2 のランドマーク予測木を比較すると、実験 1 では行動が上位に来ており、実験 2 では行動が下位になっている。実験 2 の結果のように場所に依存して観測される方位の異なるランドマークの予測は、ランドマークにより大まかな位置が判明してから行った方がよい。実験 1 でこれが見られないのは、トレーニングデータが少ないためであると考えられる。

ここでは、決定木の圧縮はさほど行っていない。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考えられる。また各ランドマーク、ボールの

方向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。連続な属性値を情報量基準により閾値によって2分していくことにより扱うC4.5 [10] のような分類木生成法を行動決定木の生成に用いることで決定木の生成と自律的な離散化を同時に行える可能性がある。

この手法は、to look or to move また、what to look の解にはなると思われるが、行動の時間的区切り方(ここでは 4.8秒で固定)、when to look の問題は残っている。さらに、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。ここでの実験では行動確率から行動を決定する際の閾値は実験的に決定したが、行動確率からの行動決定法とともに閾値の適切な決定法は今後の課題である。データ収集は教示により行ったが、教示者の負担も大きく、負担の軽いデータ収集法が望まれる。

6. 結 言

本研究では、移動体の意思決定のための効率的な観測のための情報量を利用した観測戦略を提案した。さらに行動決定木を情報量をもとに作成することにより、トレーニングデータの圧縮と観測順序の記述が統一的に扱えることを示した。そして実機実験により本手法の有効性を検証した。提案した手法は、視覚センサによるランドマークの観測のみでなく、他の能動センサによる観測も統一して扱うことが可能である。

参 考 文 献

- [1] Wolfram Burgard, Andreas Derr, Dieter Fox, and Armin 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] Akio Kosaka and Avi Kak. Fast vision-guided mobile robot navigation using model-based reasoning and prediction of uncertainties. Vol. 3, pp. 2177–2186, 1992.
- [3] 文仁赫, 三浦純, 白井良明. 不確かさを考慮した観測位置と移動のオンライン計画手法. 日本ロボット学会誌, Vol. 17, No. 8, pp. 1107–1113, 1999.
- [4] 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.
- [5] Jun 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)*, Vol. 26, No. 3, pp. 421–436, 1996.
- [6] 鈴木育男, 横井浩史, 嘉数侑昇. 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.
- [7] Jun Tani, Jun Yamamoto, and Hiro Nishi. Dynamical interactions between learning, visual attention, and behavior: An experiment with a vision-based mobile robot. In Phil Husbands and Inman Harvey, editors, *Fourth European Conference on Artificial Life*, pp. 309–317. The MIT Press, 1997.
- [8] J. Ross Quinlan. Discovering rules from large collections of examples: a case study. In D. Michie, editor, *Expert Systems in the Microelectronic Age*. University Press, Edinburgh, Scotland, 1979.
- [9] Manuela Veloso, Hiroaki Kitano, Enrico Pagello, Gerhard Kraetzshmar, Peter Stone, Tucker Balch, Minoru Asada, Silvia Coradeschi, Lars Karlsson, and Masahiro Fujita. Overview of robocup-99. In Manuela Veloso, Enrico Pagello, and Hiroaki Kitano, editors, *RoboCup-99: Robot Soccer World Cup III*, pp. 1–34. Springer, Lecture Note in Artificial Intelligence (1856), 2000.
- [10] J. Ross Quinlan. *C4.5: PROGRAMS FOR MACHINE LEARNING*. Morgan Kaufmann Publishers, 1993.

光永 法明 (Noriaki Mitsunaga)

1974年1月5日生まれ。1997年大阪大学大学院工学研究科電子制御機械工学専攻博士前期課程修了。同年同大学院工学研究科知能機能創成工学専攻博士後期課程進学。現在に到る。ロボットの知能に関連した問題に興味を持つ。

(日本ロボット学会学生会員)

浅田 稔 (Minoru Asada)

1953年10月1日生まれ。1982年大阪大学大学院基

礎工学研究科博士後期課程修了。同年同大学基礎工学部助手。1989年同年同大学助教授, 1995年同教授。1997年同大学大学院工学研究科教授となり現在に到る。この間, 1986年から1年間米国メリーランド大学客員研究員。1989年情報処理学会研究賞, 1992年IEEE/RSJ IROS'92 Best Paper Award, 1996年日本ロボット学会第10回論文賞受賞。知能ロボットの研究に従事。工学博士。電子情報通信学会, 情報処理学会, 計測自動制御学会, システム制御情報学会, IEEEなどの会員。(日本ロボット学会正会員)