

情報量に基づく移動ロボットの注視制御のための センサ空間構成

光 永 法 明* 浅 田 稔*

Sensor Space Segmentation for Visual Attention Control of a Mobile Robot based on Information Criterion

Noriaki Mitsunaga* and Minoru Asada*

Visual attention is one of the most important issues for a mobile robot to accomplish a given task in complicated environments since the vision sensors bring a huge amount of data. This paper proposes a method of sensor space segmentation for visual attention control that enables efficient observation taking the time needed for observation into account. The efficiency is considered from a viewpoint of not geometrical reconstruction but unique action selection based on information criterion regardless of localization uncertainty. The method is applied to four legged robot that tries to shoot a ball into the goal. To build a decision tree, a training set is given by the designer, and a kind of off-line learning is performed on the given data set. Discussion on the visual attention control in the method is given and the future issues are shown.

Key Words: active perception, attention control, information criterion, sensor space segmentation

1. は じ め に

移動ロボットには視覚センサが搭載されることが多いが、視覚センサを効率よく使うには能動視覚が重要となる。能動視覚の一つである注視によりシーンの理解など視覚の問題が解きやすくなることは多く報告されており、タスクに応じた選択的注視法が求められている [1]。注視には、ある視覚上の目標を見続ける注視制御、環境地図の獲得など計測を簡単にするための注視、ロボットの行動決定のための注視などの問題がある。また選択対象としては、視覚センサ自身、視覚の解像度、観測方向、観測対象物などが挙げられる。ここではロボットの行動決定のための注視における観測方向と対象物の選択問題を扱う。

移動ロボットの自己位置同定を目的として観測位置計画 [2] や、移動ロボットの能動観測 [3] などが提案されているが、注視選択は考慮されていない。Davison and Murray [4] は観測特徴量の選択を、城殿ら [5] は観測方向の決定を、阪口 [6] は注視の逐次計画を提案しているが、いずれも自己位置あるいは状態の推定を目的とした手法である。我々は、自己位置の同定や状態推定ではなく、行動決定のために効率的に観測する手法を提案している [7]。この手法では情報量により注視対象を選択しているが、観測方向についてはセンサ値の予測に依存していた。またセンサ値が前もって離散化されており、注視や観測にかかる時

間は一定であると仮定していた。しかし、さらに効率的に観測するには、注視対象と観測方向を考慮した注視制御のためのセンサ空間の自律的分割が必要となる。

強化学習の分野では、学習時間が状態空間の大きさに対して指数関数的に大きくなる [8] ことから、センサ空間の自律的分割による状態空間構成法が提案されている [9]~[12]。しかし、観測にかかる時間や、能動的視覚システムは考慮されていない。Kamiharako et al. [11] は、注視の粗密制御が結果として得られたことを示しているが、用いたセンサは周囲の観測に能動制御を必要としない全方位視覚センサであった。

そこで我々は、注視制御のためのセンサ空間の分割と行動決定木の生成法を提案する。自己位置同定のための観測ではなく、観測にかかる時間を考慮した、行動決定のための効率的な観測を実現する。また履歴を考慮するため行動決定木を生成し直すことで、さらに観測が効率化される。

2. 情報量に基づく効率的観測と行動決定

ロボットや環境、与えられるデータ等に関して以下のように設定する。

- (1) ロボットは視覚センサを持つが視野角が限られており、受動的な観測のみでは行動決定に必要な情報が得られるとは限らない。
- (2) 環境中にランドマークが配置されており、視覚センサの観測方向を変えることにより、視野を拡大し行動決定に十分な情報が得られる。

原稿受付

*大阪大学大学院工学研究科

*Graduate School of Engineering, Osaka University

(3) 教示などにより、視野を拡大した際に得られる視覚センサの値と、その際にとるべき行動のシーケンスが与えられる。また行動決定に必要な観測情報は一定ではなく、状況に応じて変化する。

情報量によりコンパクトな分類木を生成する手法として ID3, C4.5 [13] がある。本研究では、C4.5 を元にロボットのセンサの特性を考慮し拡張した行動決定木生成手法を用いる。以下に簡単に分類木の生成法を説明する。分類木の生成には、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値からなる。ID3 の場合には属性値として離散値のみを扱う。センサ i について知ったときの行動に関する情報量 I_i をすべてのセンサについて計算し、最も情報量の大きいセンサの値によってトレーニングデータセットを分割し、全てのセンサについて情報量が 0 になる (データセット中の行動が 1 種類となる) まで、データセットの分割を繰り返す。行動決定木は、ノードがデータセットを分割するセンサ、枝がセンサの値によって伸びた形となり、葉で行動を表す。C4.5 では離散値属性については ID3 と同じであるが、連続値をとる属性の場合には、センサ値と閾値の大小関係でデータセットを 2 分割した際に、情報量が最も大きくなる閾値を求め、閾値でデータセットを 2 分割することで、連続な属性値を扱う。

上記の設定で分類木を行動決定木として生成する場合には、クラスは行動に、属性はセンサに相当する。視野の限られた視覚センサを用いて、ランドマークの方位と閾値の大小関係を知るためには、ランドマークが観測されている場合を除き、視覚センサの方向を変え、観測する必要がある。これに対し、ランドマークがある範囲に観測されるか否かを知るためには、その範囲を含む方向の一度の観測で済む。そこで、データセットの分割をランドマークがある範囲 (以下、注視窓と呼ぶ) に観測されるか否かでを行い、分割後の情報量を比較する。

2.1 観測による情報量

行動の種類を r 、トレーニングデータの数を n とする。トレーニングデータ中の行動 j をとった回数を n_j とすると、各行動 $j = 1, \dots, r$ の生起確率 p_j は、 $p_j = n_j/n$ である。このとき行動に関するエントロピー H_0 は、次のようになる。

$$H_0 = - \sum_{j=1}^r p_j \log_2 p_j \quad (1)$$

ランドマーク i が注視窓 $[\theta_{Lk}, \theta_{Uk})$ の範囲に観測されるか否かが分かったときの事後生起確率を求める。ランドマーク i が $[\theta_{Lk}, \theta_{Uk})$ に観測された回数を n_{ik}^I 、観測されたときに行動 j をとった回数を n_{ikj}^I 、 $n_{ik}^I = \sum_{j=1}^r n_{ikj}^I$ とすると、範囲内に観測される場合の事後生起確率は、 $p_{ikj}^I = n_{ikj}^I/n_{ik}^I$ となる。同様に、観測されなかった場合の行動 j をとった回数を n_{ikj}^O 、 $n_{ik}^O = \sum_{j=1}^r n_{ikj}^O$ とし、観測されない場合の事後生起確率 p_{ikj}^O を求める。 $n_{ik} = n_{ik}^I + n_{ik}^O$ として、このときのエントロピーを計算すると、

$$H_{ik} = - \sum_{x=\{I,O\}} \frac{n_{ik}^x}{n_{ik}} \sum_{j=1}^r (p_{ikj}^x \log_2 p_{ikj}^x) \quad (2)$$

となり、それぞれの観測による情報量は、 $I_{ik} = H_0 - H_{ik}$ であ

る。情報量が大きいランドマーク i 、注視窓 k ほど、行動に関する曖昧さが減少する。観測範囲の上下限 θ_{Lk}, θ_{Uk} は、各ランドマーク i についてトレーニングデータ中に含まれる隣接する観測された方位の中間点などを候補とする。

2.2 観測時間の考慮

観測に要する時間が観測対象 (ランドマークとその観測される範囲) によらず一定の場合には、情報量の最も大きい観測対象により、行動決定木を生成する。この決定木はコンパクトであり、木のノードの観測を繰り返すことで最短観測時間で行動を決定できる。しかし、観測時間が観測対象により異なる場合には、最短観測時間となるとは限らない。また観測対象決定時に、得られる情報量と観測時間のトレードオフを計算するのは、観測時の計算コストが高く、決定木を使うメリットが減少する。

そこで決定木の生成の際に、観測時間を考慮した指標を用いる。現在の観測対象の次にその観測対象を観測する場合にかかる時間を T とし、単位時間当りに得られる情報量 (情報量速度) i_{ik} を次のように求める。

$$i_{ik} = \frac{I_{ik}}{T + T_C} \quad (3)$$

ここで T_C は 0 で割らないための正の定数である。現在、観測していない場合には、行動決定時の平均的なカメラの方向からの観測時間を T として用いる。すでに視覚センサが観測した方向である場合には $T = 0$ とする。

2.3 行動決定木の生成

行動決定木を次のように生成する。

- (1) 情報量速度 i_{ik} を全てのランドマーク i 、注視窓候補 k について計算する。
- (2) i_{ik} を最も大きくする i, k の組により、トレーニングデータを、ランドマーク i が注視窓 $[\theta_{Lk}, \theta_{Uk})$ の範囲に観測された場合と、観測されなかった場合に分ける。
- (3) 行動が決定できるまで情報量速度の計算とトレーニングデータの分割を繰り返す。

トレーニングデータの分割が木の枝分かれとなる。トレーニングデータに対し最後まで分割を繰り返しても、行動が確定しない場合には、各行動の頻度確率を記しておく。

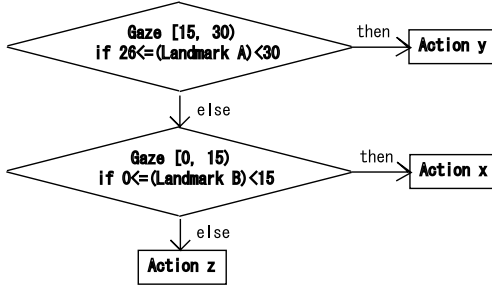
たとえば、Table 1 のトレーニングデータが与えられたとする。表中の数字は各ランドマークが観測された方向である。視野が限られており、視覚センサによる観測可能な方向が $[0, 15)$, $[15, 30)$, $[30, 45)$ の 3 つである。視覚センサは観測開始時に $[15, 30)$ を向いており、観測方向を変えるのに 1 時刻必要であるとする ($T_C = 1$)。まず、 H_0 を計算すると、 $p_x = 2/4$, $p_y = 1/4$, $p_z = 1/4$ から、 $H_0 = 1.5$ となる。ランドマーク A, B がある範囲に観測されるか否かがわかったときの、情報量 I_{ik} と単位時間当りに得られる情報量 (情報量速度) i_{ik} を計算すると Table 2 となる。ここで注視窓の上下限の候補には、各ランドマークの観測された方位の中間点と、観測可能な範囲の上下限を用いている。これは注視窓が各方向を観測した場合の視野を越えないためである。そこで、最も情報量速度が大きいランドマーク A が注視窓 $[26, 30)$ に観測されるか否かを確認する。観測された場合のトレーニングデータはデータ番号 3 のみで、行動は y が決定できる。観測されない場合には、データ

Table 1 Example training data

Data number	Landmark A	Landmark B	Action
1	5	5	x
2	25	15	x
3	27	10	y
4	40	30	z

Table 2 Information gain and information gain per time calculated from example training data (Info., Info./time and Lm indicate information gain, information gain per time and landmark respectively).

Observation	Info. (I_{ik})	Info./time (i_{ik})
$0 \leq (LmA) < 15$	.31	.15
$15 \leq (LmA) < 26$	.31	.31
$15 \leq (LmA) < 30$	.50	.50
$26 \leq (LmA) < 30$	1.4	1.4
$30 \leq (LmA) < 45$	1.4	.70
$0 \leq (LmB) < 7$	.31	.15
$0 \leq (LmB) < 15$	.50	.25
$7 \leq (LmB) < 15$	1.4	.70
$15 \leq (LmB) < 30$	.31	.31
$30 \leq (LmB) < 45$	1.4	.70

**Fig. 1** The action decision tree constructed from example training data

番号 1,2,4 が含まれ、行動は決定できない。このトレーニングデータのサブセット (データ番号 1,2,4) について、ランドマーク A, B がある範囲に観測されるか否かがわかったときの情報量速度を計算すると、ランドマーク B が $[0, 15]$ に観測されるか否かが分かったとき、ランドマーク A が $[30, 40]$ に観測されるか否かが分かったときが最も大きく、共に 0.05 である。左側を優先して観測すると行動決定木は、Fig. 1 となる。

2.4 行動決定

行動を次のように決定する。

- (1) 観測予測モデルが使える場合には、モデルを使い各注視窓にランドマークが入る確率を計算する[†]。モデルを持たない場合には、各注視窓に入る確率 (観測確率) は 0.5 とする。
- (2) 現在、視覚センサが向いている方向について、各注視窓にランドマークが入っているか、否かがわかるので、観測確率を 1 または 0 とする。
- (3) 観測確率を用いて行動決定木の各葉への到達確率を計算す

[†]例えば、観測予測モデルによりランドマーク i の観測方向が確率分布 $f(\mathbf{x})$ に従うと予測されたとする。このとき注視窓 k 内にランドマーク i が観測される確率 $\hat{p}_{ik}$ は、 $\hat{p}_{ik} = \int_{W_{ik}} f(\mathbf{x}) d\mathbf{x}$ となる。ここで W_{ik} は注視窓内の座標を表す。ただし、本論文における実験では観測予測モデルを使用せず、観測前の確率は 0.5 とした。

る。同じ行動を示す複数の葉への到達確率の和を、その行動をとるべき確率とする。

- (4) 特定の行動が閾値を越えていれば、その行動をとる。そうでなければ、根に近い観測範囲を含むまだ観測していない方向から順に観測し、観測確率を更新する ((2) に戻る)。
- 更新した確率分布により、特定の行動をとるべき確率が閾値を越えるまで、観測と確率分布の更新を繰り返す。

例えば、Fig. 2(a) の行動決定木により行動決定する場合には、次のようになる。行動確率の閾値は 0.8 とする。

- (1) 観測モデルを持たないので各注視窓にはいる確率は 0.5 であり、葉への到達確率は Fig. 2(a) となる。行動確率は、行動 x が 0.5, 行動 y が 0.25, 行動 z が 0.25 である。
- (2) 行動確率が閾値を越えていないので、根の注視窓 $[26, 30]$ を含む範囲 $[15, 30]$ を観測し、観測確率を更新する。注視窓内にランドマーク A が観測されたかとなると、葉への到達確率は Fig. 2(b) となる。行動確率は、行動 x が 0.5, 行動 y が 0.5, 行動 z が 0.0 である。
- (3) 行動確率が閾値を越えていないので、次に注視窓 $[35, 40]$ を含む範囲 $[30, 45]$ を観測し、観測確率を更新する。注視窓内にランドマーク B が観測されなかったかとなると、葉への到達確率は Fig. 2(c) となる。行動確率は、行動 x が 1.0, 行動 y が 0.0, 行動 z が 0.0 となるので、行動 x をとり行動決定を終了する。

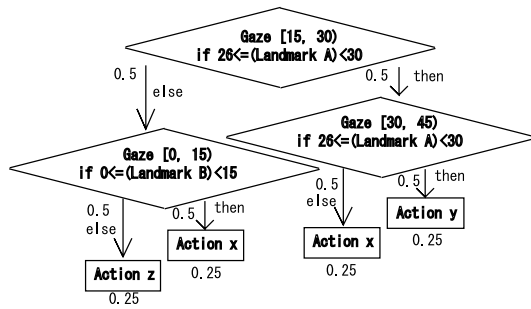
2.5 履歴を考慮した行動決定木の再生成

過去の観測や行動の履歴を利用することで行動決定までの観測の減少が期待される。しかし行動決定の際に決定に必要な観測しか行わない手法では、過去に十分に観測されているとは限らない。過去の観測を直接利用し、そのような場合を網羅した行動決定木を生成すると木が極端に大きく、実用に適さないと考えられる。

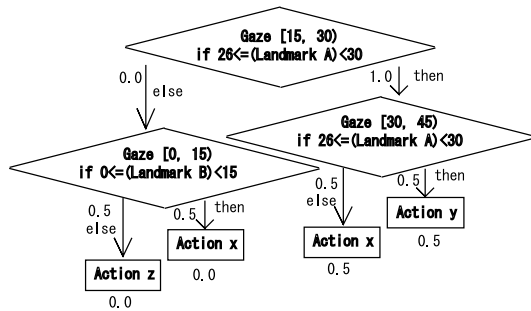
そこで過去の行動決定過程に関係なく得られる情報のみを用いることを考える。生成された行動決定木の葉に注目すると、葉はある条件の元でとるべき行動を表しており、根から葉までの経路は、その条件を満たすセンサ空間を表している。このセンサ空間は同一の行動をとる状態を表していると考えられる。そこで以下では、この空間を同一行動状態と呼び、これを用いて行動決定木を再生成し、観測時間の短縮を図る。

行動決定木は次のように再生成する。まず、各トレーニングデータの属している同一行動状態と、一時刻前の同一行動状態を調べておく。各シーケンスの先頭については、一時刻前の同一行動状態は不明としておく。不明も一つの状態である。情報量速度は、行動に関するものではなく同一行動状態に関する情報量を計算する。

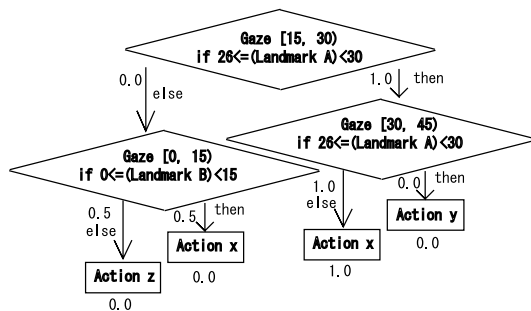
- (1) 同一行動状態に関する情報量速度 i_{ik} の最も大きい i, k の組を求める。
- (2) 一時刻前の同一行動状態によりトレーニングデータを 2 組に分けた場合に、最も情報量の大きくなる分け方を求める。同一行動状態を知るのに必要な時間は 0 であるから、この情報量を T_C で割った値を情報量速度とする。
- (3) (1),(2) の情報量速度を比較し、大きい方の分割法でトレーニングデータを分ける。



(a) Without observation



(b) Gaze [15, 30]



(c) Gaze [30, 45]

Fig. 2 Calculation example of action probabilities

(4) 同一行動状態が決定できるまで (1) から繰り返す。

ここで、行動決定シーケンスの開始状態が、トレーニングデータ中のいずれかのシーケンス開始状態に一致する場合には問題が起きないが、一致しない場合には正しく行動決定されない可能性がある。例えば、Fig. 3 のようにトレーニングシーケンスから行動決定木が再生成されたとする。図のトレーニングシーケンスは同一行動状態で表しており、最初の同一行動状態は p 、その一時刻前は不明である。行動決定の開始時は直前の同一行動状態は不明である。同一行動状態が r である状態から行動を開始した場合には、トレーニングシーケンスでは直前の同一行動状態は p または q であるのに対し、木を使った行動決定時は「不明」となる。そのため、行動決定を誤り、 x をとることとなる。

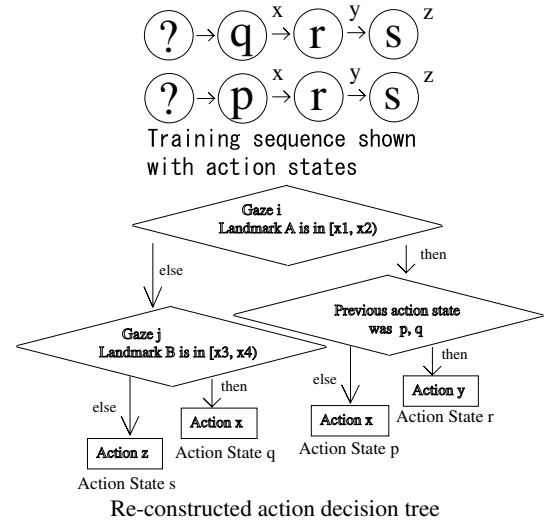


Fig. 3 An example training sequence and re-constructed action decision tree

そこで、トレーニングシーケンス途中から行動を開始する可能性のある場合には、シーケンス途中のトレーニングデータについて、一時刻前の同一行動状態を不明としたデータを追加し、行動決定木を再生成する。これにより、行動決定木は一時刻前の同一行動状態が不明であっても正しく行動決定できる。

行動決定の際には、観測確率と一時刻前の各葉に到達した確率 (同一行動状態確率) から、同一行動状態確率を計算し、行動確率を求める。このとき、行動決定が可能であれば、同一行動状態が特定できなくても、それ以上、観測しない。

3. 実 験

ロボットとしては、RoboCup SONY 脚式ロボットリーグのロボット (Fig. 4) を用いた。カメラの画角は横 52 度、縦 48 度、画素数はそれぞれ 88, 71 である。脚は各 3 自由度、首は 3 自由度 (パン、チルト、ロール) がある。ランドマークを観測する際には脚、首のロールを固定し、パン、チルト軸のみを利用した。パン軸はロボット正面に対して、-88 度から 88 度、チルト軸は、-80 度から 43 度が可動範囲である。能動的に観測する方向としては、パン軸を 44 度毎の 5 方向、チルト軸を 40 度毎の 2 方向の 10 方向にわけた。パン軸の最大角速度は $6[\text{rad/s}]$ 、チルト軸の最大角速度は $4[\text{rad/s}]$ である。また軸が目標値に達した後で画像が安定するまで待ってから処理をしており、カメラの方向を変えて観測する場合に少なくとも $0.36[\text{s}]$ かかることから、 $T_C = 0.36[\text{s}]$ とした。行動としては、ボール接近、前進、後退、左 (右) 前進回転、左 (右) その場回転、左 (右) 横移動、左 (右) 横移動回転、行動終了の 12 を用意した (Fig. 5)。

実験環境を Fig. 6 に示す。RoboCup SONY 脚式ロボットリーグのフィールドである。ランドマークは 6、ゴールが 2 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情報により識別さ



Fig. 4 A robot for the RoboCup SONY legged robot league

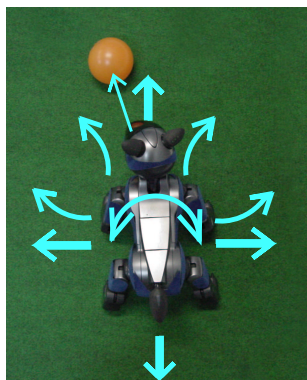


Fig. 5 Prepared moving actions

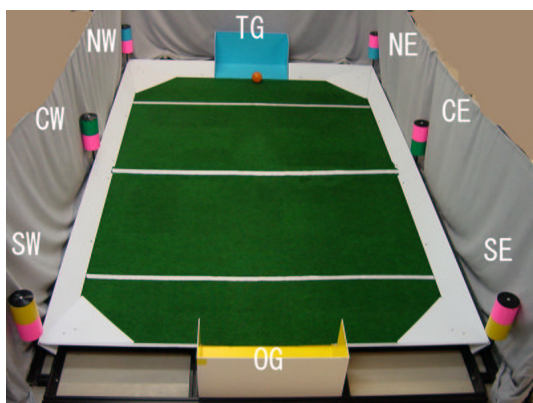


Fig. 6 The experimental environment (the same configuration for RoboCup SONY legged robot league)

れる。

視覚センサとしては、各ランドマークとボールの図心座標、両ゴール (TG, OG) の画像上での (x, y) 座標が最小/最大となる座標 4 つを用いた。 x, y の直積 (長方形領域) により注視窓を構成し、トレーニングデータを分割した。これは、例えば x 座標の範囲が -10 度から 10 度にあるか否かで分割すると、 y 座標に関しては指定されていないので y 軸についてカメラを回転して観測する必要があるためである。

4. 実験結果

ロボットがボールと TG を正面に観測する位置へ移動することをタスクとした。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。ボールの位置を 3 箇所、ロボットの行動開始位置を 5 箇所とし、各点から

Table 3 Comparison of the size of trees, expected number of gaze directions, and time to make a decision.

	# of nodes	Max. depth	# of leaves	# of dirs	time[s]
Pre-quantized	41	7	42	4.4	2.5
Info. gain	19	7	20	3.7	2.2
Info./time	29	9	30	1.5	0.84
Re-construct	34	9	35	1.4	0.60
Re-construct(M)	29	7	30	1.5	0.82

目標位置姿勢までの行動を教示した。

予めセンサ値を 20 画素毎に離散化しておいた場合と、情報量を用いた離散化をした場合、情報量速度を用いて離散化した場合、さらに同一行動状態を用いて行動決定木の再生成した場合 (シーケンス途中からの行動開始への考慮なし/あり)、を比較した。教示した行動ステップ数 97 回の時点でのトレーニングデータから、それぞれの手法により得られた注視窓を Fig. 7~Fig. 9 に示す。図中数字の単位は画素である。横軸が x 座標、縦軸が y 座標で、マス目は各観測方向にカメラが向いた際の観測可能範囲を示している。マス目の中の長方形が注視窓である。予め離散化した場合 (Fig. 7) と情報量による離散化 (Fig. 8) を比較すると、適切な離散化は均等に分けることではないことが分かる。情報量による離散化 (Fig. 8) と情報量速度による離散化 (Fig. 9) を比較すると、観測時間の考慮により注視窓のある観測方向が 6 から 3 に減少している。また注視窓を正面の方向に集中させることで、情報量の大きい注視窓ではなく、情報量速度が大きく、すでに観測した方向にある注視窓から行動を決定することにより、観測時間を減少しようということが分かる。履歴なし (Fig. 9) とあり (Fig. 10, Fig. 11) を比較すると、注視窓のある観測方向はほとんど同じであり、注視窓の傾向も似ていることが分かる。

Table 3 に、予めセンサ値を離散化した場合 (Pre-quantized)、情報量による離散化をした場合 (Info. gain)、情報量速度による離散化をした場合 (Info./time)、木の再生成をした場合 (Re-construct)、シーケンス途中からの行動開始を考慮した木の再生成をした場合 (Re-construct(M)) について、ノード数 (# of nodes)、木の最大深さ (Max. depth)、葉数 (# of leaves)、観測方向数の期待値 (# of dirs)、観測時間の期待値 (Time) の比較を示す。観測方向数と観測時間の期待値は観測予測なしの場合のものである。提案手法によるセンサ値の離散化が木の大きさ、観測方向数、観測時間の減少に有効であることが分かる。また情報量速度の利用により、木は少し大きくなるが、観測時間は半分以下に改善されている。一時刻前の同一行動状態を用いると、木は大きくなるが必要な観測が減少し観測時間がさらに短くなっている。シーケンス途中からの行動開始を考慮すると、履歴なしの場合とほぼ同一であるのは、再生成された木で同一行動状態が使われていないためと考えられる。Fig. 9 と Fig. 11 で注視窓がよく似ているのも同じ理由である。一時刻前の同一行動状態が使われていないのに、観測時間が減少している理由は、センサ空間のある程度まとまった同一行動状態についての決定木を生成しているためと考えられる。

Table 4 に、生成された木を使って実機で行動を決定した場

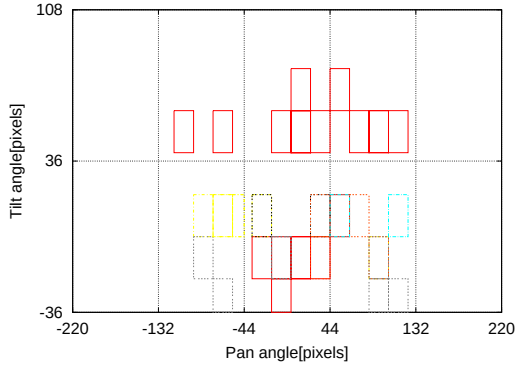


Fig. 7 Created attention windows with pre-quantized sensor values

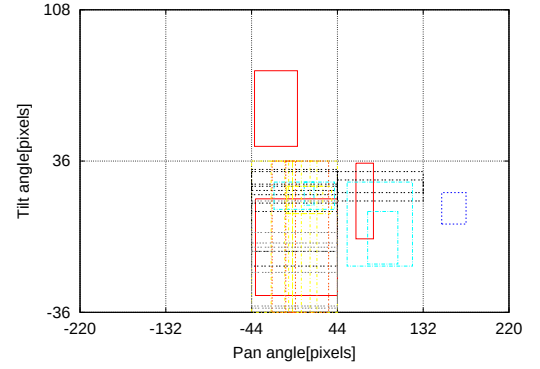


Fig. 10 Created attention windows of re-constructed tree. Quantization was done by information gain per time.

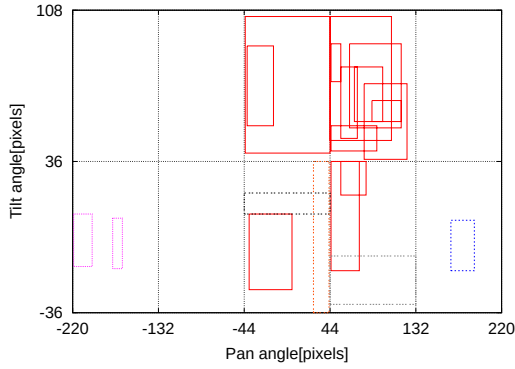


Fig. 8 Created attention windows with quantization by information gain

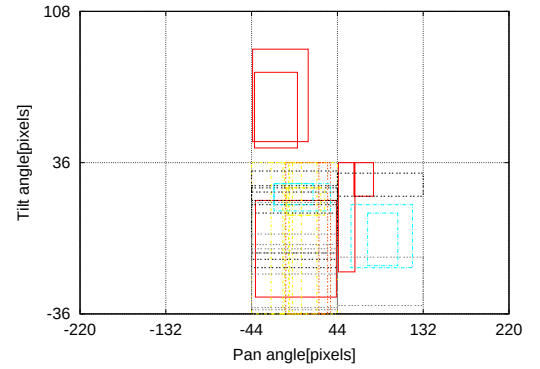


Fig. 11 Created attention windows of re-constructed tree. Quantization was done by information gain per time. Action decisions started from middle of the training sequence were considered

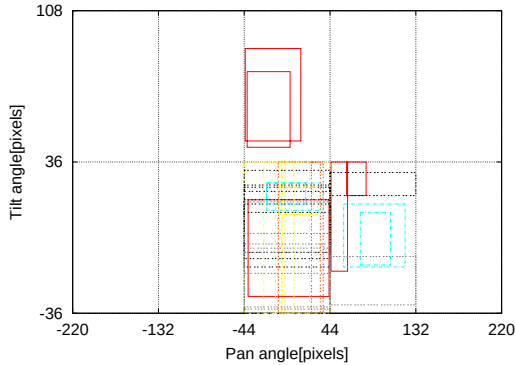


Fig. 9 Created attention windows with quantization by information gain per time

合の、平均観測方向数 (Average gaze directions) と平均観測時間 (Average time) を示す。期待値とほぼ同じ値が実現されている。Fig. 12～Fig. 15 に、行動決定時の各葉への到達確率、同一行動状態の確率、行動確率を示す。各マスは確率を表し、白が 0、黒が 1 である。上から下に時刻が進んでおり、観測方向が変わる毎に確率が計算され段が変わっている。空白行は行動決定されたことを示す。

実験ではセンターライン上の、ランドマーク CE 近傍の位置から行動開始した。情報量による離散化では行動決定までに観測を繰り返していることが分かる (Fig. 12, Table 4)。情報量

Table 4 Comparison of average number of gaze directions, and time to make a decision in experiments.

	# of data	Average gaze directions	Average time[s]
Pre-quantized	34	3.1	3.3
Info. gain	43	3.5	3.3
Info./time	35	1.3	0.85
Re-construct	45	1.0	0.62
Re-construct(M)	42	1.4	0.97

速度を用いた決定木により、1 または 2 方向の観測で行動が決定されていることが分かる (Fig. 13, Table 4)。この例では情報量速度により十分に観測時間が短縮されているため、同一行動状態による木の再生成の効果は現れていない (Fig. 14)。

教示した行動ステップ数が 53 回の時点での情報量速度を用いた決定木の場合には、異なる行動決定がされた (Fig. 15)。最初に決定した行動が異なるのは、初期位置姿勢の少しの差により異なった行動が選ばれているためである。これはトレーニングデータに同じ位置で異なる姿勢の場合に別の行動を与えたものを、微妙な姿勢差で識別しようとしたためと考えられる。行動ステップ数が 97 回の場合 (Fig. 13) と比較してシーケンスが短いのは、ロボットがボールの前にたどり着く前に行動終了を選択したためである。これらは、トレーニングデータの増加により解消されたと考えられる。

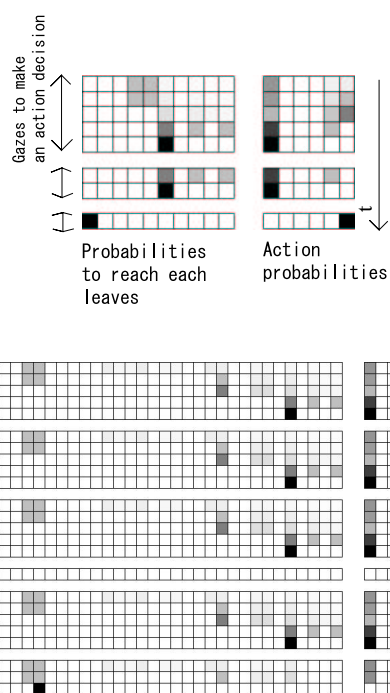


Fig. 12 An example action decision sequence by the tree constructed by information gain. The probabilities to reach each leaves (left) and the action probabilities (right) are shown. The intensity of each small box indicates the probability, white means zero and black means one.

同一行動状態を用いて行動決定木を再生成した場合 (シーケンス途中からの行動開始への考慮あり) により得られた行動決定木の一部を Fig. 16 に示す。図中数字の単位は画素である。まずゴール TG の位置と大きさによって行動決定しようとしている。教示された条件ではボールとゴールの位置が限られているため、行動決定にボールを利用しない場合も多くみられる。

5. 結 言

自己位置の同定ではなく行動決定のための、視覚センサと観測時間を考慮した、効率的な観測をする行動決定法を提案した。実験により提案手法が、適切な離散化による行動決定木の圧縮と、観測時間の短縮に有効であることを確認した。

行動決定木により決定された行動が期待したものとはならない場合がある。一つは、行動確率の閾値が低い場合である。これは観測が少ないときに行動確率の高い行動と、十分な観測を得た後に確率が高い行動が必ずしも一致しないためである。これには行動確率について信頼度の導入による解決が考えられる。もう一つは、トレーニングデータが不足している場合である。情報量速度の高い観測を優先するため、データが少ないときには、トレーニングデータに含まれたノイズを行動決定に利用した決定木を生成してしまうことがある。行動決定木のトレーニングデータに対する汎化能力は、トレーニングデータに含まれるランダムマーク方位間の内挿および、それらから注視窓内外を決め

る閾値 (ここでは方位間の中点あるいは、カメラの画角) までの外挿が期待出来る。汎化の程度は、センサ空間のタスクとの親和性に依存する。タスクに適切なセンサ空間、必要なトレーニングデータの見積もり、トレーニングデータの追加時の部分的な決定木の修正の実現と観測時間への影響などは今後の研究課題である。

謝辞 本研究の一部は、科学技術振興事業団戦略的基礎研究推進事業 (CREST) 領域「脳を創る」の補助を受けた。

参 考 文 献

- [1] 石黒. 注視に基づくロボットの視覚. 日本人工知能学会誌, 10(4):500–506, 1995.
- [2] 文, 三浦, 白井. 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, 17(8):1107–1113, 1999.
- [3] 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.
- [4] A. J. Davison and D. W. Murray. Mobile robot localisation using active vision. In *Proc. 5th European Conf. on Computer Vision*, pp. 809–825, 1998.
- [5] 城殿, 三浦, 白井. 誘導による移動経験に基づく視覚移動ロボットの自律走行. 日本ロボット学会誌, 19(8):1003–1009, 2001.
- [6] 阪口. 動きの予測を伴う能動的認識のアルゴリズム. 日本ロボット学会誌, 12(4):708–714, 1994.
- [7] 光永, 浅田. 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. 日本ロボット学会誌, 19(6):793–800, 2001.
- [8] S. D. Whitehead. A complexity analysis of cooperative mechanisms in reinforcement learning. In *Proceedings of AAAI-91*, pp. 607–613, 1991.
- [9] Y. Takahashi, M. Asada, and K. Hosoda. Reasonable performance in less learning time by real robot based on incremental state space segmentation. In *Proceedings of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1518–1524, 1996.
- [10] 矢入, 中須賀, 堀. 異種冗長なセンサー情報に基づく自律的状态抽象化法. 日本人工知能学会誌, 14(4):667–678, 1999.
- [11] M. KAMIHARAKO, H. ISHIGURO, and T. ISHIDA. Attention control for state space construction. In Y. Kakazu, M. Wada, and T. Sato eds., *In Proc. of the Intelligent Autonomous Systems 5*, pp. 258–265, 1998.
- [12] M. Asada, S. Noda, and K. Hosoda. Action based sensor space segmentation for soccer robot learning. *Applied Artificial Intelligence*, 12(2-3):149–164, 1998.
- [13] J. R. 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. 2002 年単位取得の上, 退学. 同年, 文部科学省・科学振興調整費・戦略的研究拠点「阪大フロンティア研究機構」研究プロジェクト・科学技術振興特任教員. ロボットの知能に関連した問題に興味を持つ. (日本ロボット学会正会員)

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

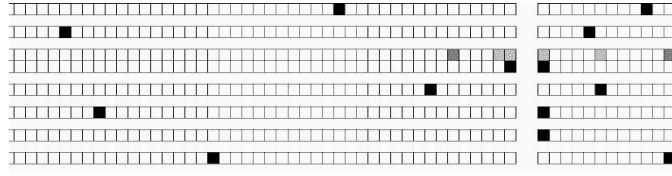


Fig. 13 An example action decision sequence by the tree constructed by information gain per time. The probabilities to reach each leaves (left) and the action probabilities (right) are shown.

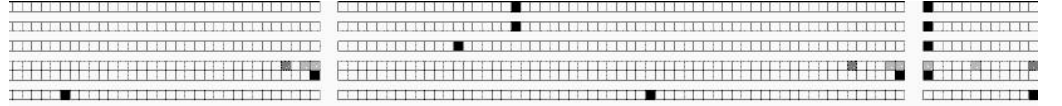


Fig. 14 An example action decision sequence by the re-constructed tree with action states. The original tree was constructed by information gain per time. The probabilities to reach each leaves (left), the action state probabilities (center), and the action probabilities (right) are shown.



Fig. 15 Another example action decision sequence by the tree constructed by information gain per time. We used 53 training data for this tree. The probabilities to reach each leaves (left) and the action probabilities (right) are shown.

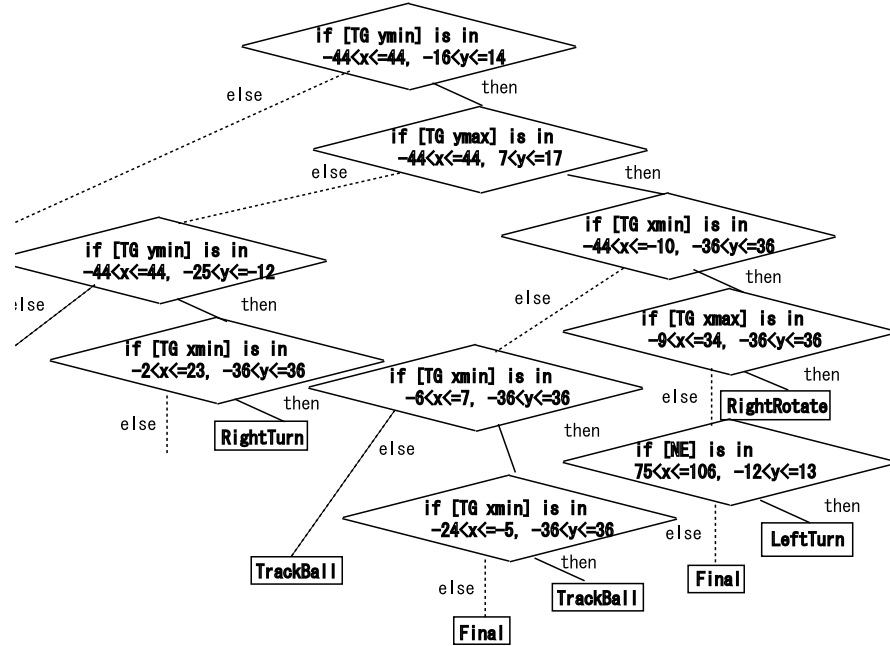


Fig. 16 Part of the action decision tree constructed by the proposed method

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）