

# 特別研究報告

題目

複数のセンサ・アクチュエータを用いた人物探索における  
能動推論の有効性の評価

指導教員

村田 正幸 教授

報告者

小菅 瑛斗

2024年2月7日

大阪大学 基礎工学部 情報科学科

## 内容梗概

近年、物理空間の情報を把握し、その情報を活用することが期待されている。Cyber-Physical System (CPS) では物理空間の情報を集積し、仮想空間で解析することによって、さまざまな産業への活用がなされている。また Society5.0 では、都市や地域単位で CPS を利用することで、そこに住むすべての人の生活をよりよくするための様々な活用方法が考案、研究されている。CPS においては物理空間を正確に把握し、仮想空間に多くの情報を集積することが求められる。しかしながら、物理空間の情報は過去情報を含めると膨大な情報量となり、仮想空間上において情報が欠損することが不可避である。そのため欠損した情報を CPS 上に存在する情報などの利用によって補う欠損情報の補間や、情報が不足している領域を能動的にセンシングすることによる欠損補填を行うことが考えられる。

情報の不確実性およびあいまいさを捉え、対象物の観測や自身の行動を推論する制御フレームワークに自由エネルギー原理をもとにした能動推論がある。能動推論では、エージェントは環境の状態に対する事前の信念と観測結果を組み合わせることで、環境の状態を確率的に推論し、かつ、自身の行動による確率の変化を把握することができる。すなわち情報が不確実であるなかで、環境の状態の推定とそうした不確実性を減少させる行動制御を1つのフレームワークで行うことが可能になる。

そこで本報告では、仮想空間上における欠損情報の補間や欠損補填を行うために物理空間にセンシングを適切に働きかけることを目標として、複数の可動式カメラによる人物探索に能動推論を適用する。探索の対象となるターゲットは移動し続けるため、常にセンシングし続けることは困難であり、ターゲットに関する情報が欠損する。そのため、ある可動式カメラでターゲットをセンシングできないときに、その位置を推定し、他の可動式カメラの捕捉範囲を制御することで、ターゲットをより効率的に捕捉することが期待できる。また、ターゲットの移動には傾向があると考えられ、その傾向を把握しながら探索を行うことで、ターゲットの捕捉が行いやすくなると考えられる。よってエージェントはターゲットの傾向の学習も行いながら、人物探索を行っていく。シミュレーションの結果、能動推論を利用しない場合に比べて、ターゲットの捕捉割合が向上することを示した。

## 主な用語

Cyber-Physical System、自由エネルギー原理、能動推論、ターゲット捕捉、人物探索、可動式カメラ

# 目次

<b>1</b>	<b>はじめに</b>	<b>6</b>
<b>2</b>	<b>関連研究</b>	<b>8</b>
2.1	自由エネルギー原理と生成モデル . . . . .	8
2.2	能動推論 . . . . .	9
2.3	能動推論を用いたマッピングシステム . . . . .	12
<b>3</b>	<b>能動推論を利用した人物探索システム</b>	<b>14</b>
3.1	システム概要 . . . . .	14
3.2	システム詳細 . . . . .	15
3.3	ポリシー推論の高速化 . . . . .	18
<b>4</b>	<b>シミュレーション評価</b>	<b>20</b>
4.1	評価環境 . . . . .	20
4.2	評価方法 . . . . .	20
4.3	評価結果 . . . . .	22
<b>5</b>	<b>おわりに</b>	<b>26</b>
	<b>謝辞</b>	<b>27</b>
	<b>参考文献</b>	<b>28</b>

## 図目次

1	自由エネルギー原理の概念図 . . . . .	9
2	能動推論を行うエージェントが保持する生成モデル . . . . .	10
3	システムの概要図 . . . . .	16
4	複数の観測を利用した能動推論のシステム図 . . . . .	18
5	ポリシー推論を高速化したシステム図 . . . . .	19
6	使用したマップの一例 . . . . .	21
7	可動式カメラの位置と観測範囲 . . . . .	22
8	直進確率による可動式カメラのターゲット捕捉割合の変化 . . . . .	23
9	事前知識が直進確率 0.33 であるときの生成モデルの直進確率の更新 . . . . .	24
10	事前知識が直進確率 0.5 であるときの生成モデルの直進確率の更新 . . . . .	25
11	事前知識が直進確率 0.1 であるときの生成モデルの直進確率の更新 . . . . .	25

## 表 目 次

1	生成モデルにおける変数・記号説明 . . . . .	9
2	分解された生成モデルにおける変数・記号説明 . . . . .	11
3	変分自由エネルギーにおける変数・記号説明 . . . . .	11
4	期待自由エネルギーにおける変数・記号説明 . . . . .	12
5	ポリシー推論における変数・記号説明 . . . . .	12

## 1 はじめに

近年、物理空間の情報をセンシングすることによって状況を把握し、その情報を利活用することが期待されている。その一例として、Cyber-Physical System (CPS) が挙げられる。CPS は、物理空間のデータをセンシングデバイスで収集し、仮想空間上のビッグデータを解析したうえで、その結果を物理空間の制御やサービスにフィードバックするシステムである [1]。これまでは人間が主体となって情報を解析していたことに対して、人間が扱えないようなビッグデータを計算機の力を使って解析することで、これまでには出来なかった新たな価値が物理空間にもたらされることが期待される。現在は医療におけるデータ分析や工場や農場の状況把握など、主に工業分野で活用されている [2]。さらにここ数年、Society5.0などの効果的なデジタルトランスフォーメーションの実現が期待されている。Society5.0とはCPSにより、経済発展と社会的課題の解決を両立する、人間中心の社会 (Society) である [3]。都市や地域単位でセンシングを行い、集まった膨大なデータを解析、フィードバックすることにより、交通情報推定や落とし物探索、迷子 / 徘徊老人 / ペット探索など、より人の生活に即したユースケースが考えられている。

CPSにおいて重要となるタスクに、物理空間の状況の把握がある。例えば、都市を対象にする場合、街中がどのように構成されているか、人や車などがどこにあるのか、など、物理空間の状況を正確に把握することが重要である。しかしながら、センシングデバイスの数は有限であり、物理空間におけるすべての場所の情報をセンシングすることはできないため、物理空間の情報が欠損することは明らかである。そのため、情報の欠損を考慮した仮想空間における情報の解析や、物理空間に対する追加のセンシングが必要となる。そのような情報の欠損に対応する方法として、欠損情報の補間や欠損補填が考えられる。欠損情報の補間は、CPS上の過去情報などの利用によって欠損情報を補間することが考えられる [4,5]。また、物理空間に対する追加のセンシングとしては、ロボットやドローンおよび可動式カメラなどアクチュエータを伴うものを利用して欠損補填を行うことが挙げられる。しかし、アクチュエータを利用したセンシングを行ったとしても、物理空間は膨大であり情報の欠損は不可避である。従って、ユースケースに応じて、欠損情報を補間するためのアルゴリズムと連動し、欠損情報の補間が適切に行われる領域を決定する必要がある。

情報の不確実性を捉え、対象物の観測や自身の行動を推論する制御フレームワークに能動推論がある [6]。能動推論では、エージェントは環境の状態に対する事前信念を生成することのできる生成モデルを保持している。生成モデルから出力された環境の状態の事前信念と実際の観測結果を組み合わせることで、環境の状態の事後信念を推定することができる。また、推定した状態をもとに、自身の行動による状態の変化を把握することもできる。これによってエージェントは、外界の状態の信念にしたがって、不確実性を減少させるための最適

な行動を選択することが可能になる。さらにエージェントは環境の状態の信念をより正確に生成するために、生成モデルを更新することが可能である。生成モデルを更新しながら環境の推定および能動推論を行っていくことで、環境の変化にも対応することができる。

本報告の関連研究として、能動推論を使用して、ロボットのナビゲーションとマッピングを行っている研究 [7] がある。探索範囲は倉庫の内部として、1台のロボットをエージェントとみなし、ロボットにカメラを1つ搭載して、カメラの画像を用いて倉庫の内部の地図と目標地点までの経路を把握することが目的である。倉庫の内部には視覚的に類似している場所が多く、倉庫の内部の環境や自身の位置があいまいになりやすいため、能動推論を利用しない場合には一回の探索でマップを作ることはできなかった。しかし能動推論を利用することにより、自身の位置の状態を推定しながら探索することができるため、一度の探索でマップを構成することができ、これによって最短経路の導出にも成功している。一方でこの研究において、制御しているのは単一のセンサ・アクチュエータであり、複数台のセンサ・アクチュエータを利用することができれば、より多くの応用への適用が可能となると考えられる。

そこで本報告では CPS における物理空間の情報の欠損を考慮しながら、仮想空間に存在する情報を分析し、物理空間に存在する複数のセンサ・アクチュエータを制御することを目的として、能動推論を適用することの有効性を評価する。ユースケースとして迷子 / 徘徊老人 などの人物探索を想定し、街中に配置された複数の可動式カメラの観測方向に関する個々の制御を、能動推論を用いて行う。能動推論を用いることで、あるカメラが対象者を観測できた場合ならびに観測できなかった場合に、隣接する可動式カメラにおいて対象者が観察される不確実性を捉えることが可能となり、アクチュエータを効率的に使用することが期待できる。また、街中を移動する人物には、移動先の傾向があると考えられる。例えば狭い道路よりも広い道路を選びやすかったり、暗い道路よりも明るい道路を好んだり、見通しの悪い道路よりも見晴らしの良い道路が使われやすかったりする。このような傾向を把握しながら人物探索を行うことで、ターゲットをより効率的に捉えることが可能ではないかと考えられる。能動推論を行うエージェントはこの傾向も推定しながら人物探索を行っていくものとする。

本報告の構成は以下の通りである。まず2章では自由エネルギー原理および能動推論について説明し、さらに能動推論を利用したマッピングに関する関連研究を述べる。3章では能動推論を利用した本報告における人物探索システムについて説明する。4章では人物探索システムについての評価を行う。最後に5章では本報告のまとめと今後の課題を述べる。

## 2 関連研究

本章では関連研究として、自由エネルギー原理の研究、及び自由エネルギー原理をもとにした能動推論を利用したマッピングについての研究について記述する。

### 2.1 自由エネルギー原理と生成モデル

自由エネルギー原理の概要について説明する [6,8]。自由エネルギー原理とは、フリストンが提唱した脳の情報理論で、生物の知覚、学習が変分自由エネルギーと呼ばれるコスト関数を最小化するように決定され、その結果生物は外界に適応できるとされている。

生物の脳は、外界や自身の状態およびその状態から得られる観測を確率的に表現する生成モデルを保持していると考えられている。生成モデルは、外界の状態から観測が生成される仕組みを表すモデルであり、未来の外界の状態と観測の信念として表現される。この生成モデルに基づき、観測から真の外界の状態を推論し、将来の観測や状態を予測することができる。

また、生成モデルによって生成された観測と、実際に得られた観測の誤差は、変分自由エネルギーとして計算される。この変分自由エネルギーを最小化することで、生成モデルによって生成された外界の状態の信念を、真の外界の状態に近づけることができると考えられる。また、状態の信念だけでなく、生成モデルを更新することもできる。これによって環境そのものが変わり、外界の状態の生成過程が変化したとしても、状態の推論を継続することができる。さらに自由エネルギー原理は、外界の状態の推論や未来の観測の予測などの知覚のみならず、行動・意思決定の最適化についても説明可能である。生物には目的があり、その目的を達成するために状態の推定や行動を行うと考えられている。その目的は事前の選好として表現され、望ましい観測を定義することができる。そうした事前の選好を得るためや信念における不確実性を減少させるための、変分ベイズ推定に基づく行動制御・意思決定の最適化は能動推論と呼ばれている。

図1は、自由エネルギー原理における処理内容を表した概念図である。ただし、生物はエージェントとして表現される。ここでは、外界が何か観測を生成すると、エージェントは生成モデルの出力と得られた観測から外界の状態を推定し、その事後信念を脳内で表現する。このとき、自由エネルギーを最小化するように事後確率や生成モデルを更新することで、変分ベイズ推論を最適に行うことができる。さらに、期待自由エネルギーと呼ばれる、将来の変分自由エネルギーを最小化すると予測される行動を能動的に推論し選択することで、外界の情報や事前の選好を得られる確率を最大化することができる。その後、外界の状態は時間経過で更新され、また観測の生成から繰り返される。この繰り返しによって、エージェン

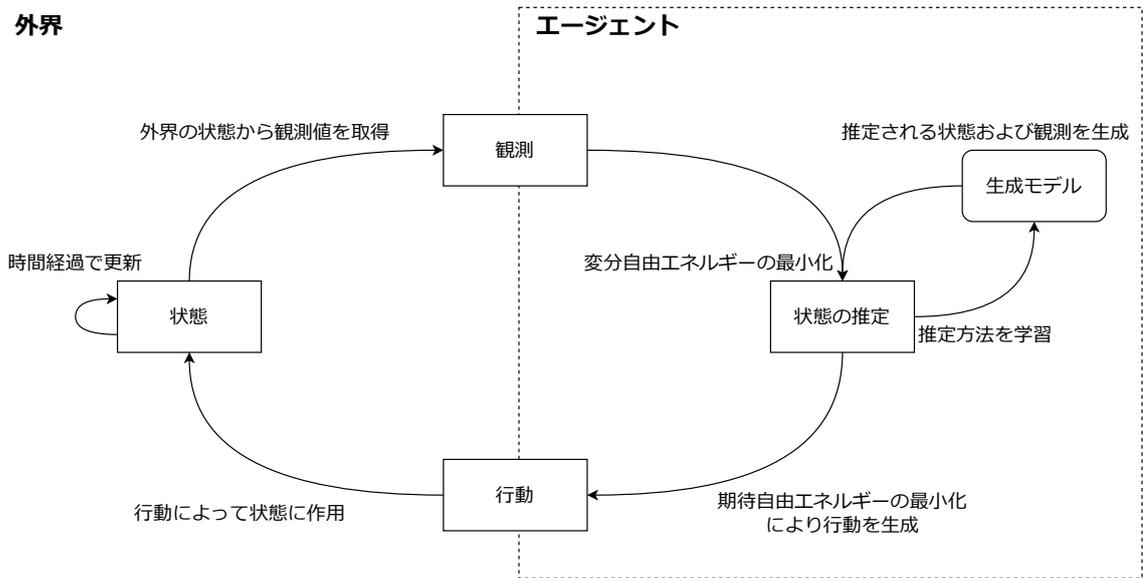


図 1: 自由エネルギー原理の概念図

表 1: 生成モデルにおける変数・記号説明

$t$	時刻のステップ
$T$	現在の時刻ステップ
$s_t$	時刻 $t$ における行動のベクトル
$o_t$	時刻 $t$ における観測値のベクトル
$s_t$	時刻 $t$ における状態のベクトル
$L$	行動を決定するうえでの考慮する未来の長さ
$\pi$	時刻 $T$ から $T + L$ までの行動のシーケンス (行列)

トは外界の不確実性を表現し、より外界を正確に表現するための生成モデルの学習や、不確実性を減少させるような行動計画を取ることができる。

## 2.2 能動推論

能動推論を利用するための定式化について説明する [9]。能動推論では、エージェントは外界の状態や観測の信念を生成する、生成モデルを保持しているとみなす。この時の生成モデルは図 2 で示すようなグラフモデルとして構成される。ただし、生成モデルにおける変数・記号説明は表 1 に示す。

グラフモデルのノードは確率変数を表し、灰色のノードは観測された変数、白色のノ

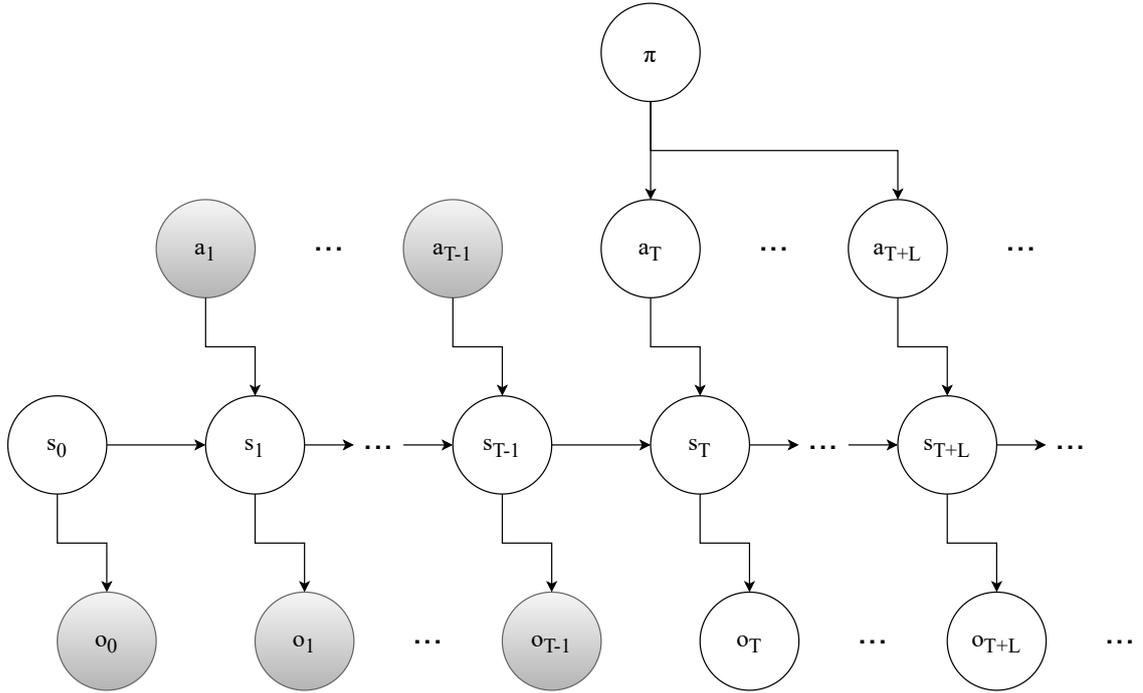


図 2: 能動推論を行うエージェントが保持する生成モデル

ドは観測されていない変数を示している。現在の時刻ステップにおける観測  $o_T$  は、状態  $s_T$  によって生成される。また、状態  $s_T$  は、1つ前の時刻ステップ  $T-1$  での状態  $s_{T-1}$  と行動  $a_{T-1}$  によって生成される。そして、将来の状態および観測は、ポリシー  $\pi$  によって決定される将来の行動によって生成される。ここで重要な点としては、エージェントは行動ではなくポリシーを推論するという点である。これは、未来の状態を推定する際に、一つ先の未来だけでなく、より遠い未来を考慮した上で、最適な行動を決定するために必要である。

現在の時刻  $T$  までの生成モデルは図2に従って分解すると、式1のようになる。なお、式1で使われている変数・記号の意味は表2に示すとおりである。

$$P(\tilde{o}, \tilde{s}|\pi) = P(s_0)P(\pi) \prod_{t=1}^T P(s_t|s_{t-1}, \pi) \prod_{t=0}^T P(o_t|s_t) \quad (1)$$

生成モデルは  $P(s_0), P(\pi), P(s_t|s_{t-1}, \pi), P(o_t|s_t)$  で構成されることが分かる。これらのベクトル、行列を適切に定めることによって生成モデルを構築することができる。ただし、これらのベクトル、行列は環境を真に表現しているものとは限らない。そのためエージェントは、生成モデルを必要に応じて更新していく必要がある。

また、ポリシー  $\pi$  における変分自由エネルギーは式2によって計算することが可能である。ただし式2で初出の変数・記号とその説明は表3に示す。

表 2: 分解された生成モデルにおける変数・記号説明

$\tilde{o}$	時刻 0 から $T$ までの観測値のシーケンス (行列)
$\tilde{s}$	時刻 0 から $T$ までの外界の状態のシーケンス (行列)
$P(\tilde{o}, \tilde{s} \pi)$	生成モデルが予測するポリシーの下での状態および観測の同時信念
$P(s_0)$	生成モデルに与えられる初期状態の確率 (ベクトル)
$P(\pi)$	生成モデルに与えられるポリシーの事前確率 (行列)
$P(s_t s_{t-1}, \pi)$	生成モデルに与えられる $\pi$ のもとでの状態の遷移確率 (行列)
$P(o_t s_t)$	生成モデルに与えられる $s_t$ のもとでの $o_t$ の生成確率 (行列)

表 3: 変分自由エネルギーにおける変数・記号説明

$F(\pi)$	$\pi$ における変分自由エネルギー
$Q(\tilde{s} \pi)$	$\pi$ のもとで推測した状態の近似事後信念

$$F(\pi) = \mathbb{E}_{Q(\tilde{s}|\pi)}[\log Q(\tilde{s}|\pi) - \log P(\tilde{o}, \tilde{s}|\pi)] \quad (2)$$

エージェントは真の状態を知ることが出来ないため、 $Q(\tilde{s}|\pi)$  を導入して、近似として事後信念を推定する。 $Q(\tilde{s}|\pi)$  は  $F(\pi)$  を最小化するように、変分ベイズ推論に基づいて計算される。生成モデルは、初期状態のベクトルや前の時間ステップにおける近似の事後信念や、ポリシーの事前信念を保持している。そのため、観測が与えられた際、各ポリシーについての変分自由エネルギーを計算し変分を行うことで、各ポリシーを行った結果の状態の事後信念を求めることが可能となる。

そして時間ステップ  $t$  の時の、ポリシー  $\pi$  における期待自由エネルギーは式 3 によって計算することが可能である。ただし式 3 で初出の変数・記号とその説明は表 4 に示す。

$$G(\pi, t) = \mathbb{E}_{Q(s_t, o_t|\pi)}[\log Q(s_t|\pi) - \log P(o_t, s_t|\pi)] \quad (3)$$

$$\simeq -\mathbb{E}_{Q(s_t, o_t|\pi)}[D_{KL}[Q(s_t|o_t, \pi)||Q(s_t|\pi)]] - \mathbb{E}_{Q(s_t, o_t|\pi)}[\log p(o_t)]$$

第一項は情報利得と呼ばれる項である。これは観測を得た時の、予測される近似事後信念の変化の度合いを表し、これが大きければ状態の不確実性を減少させることができる可能性が高いことを表す。そのためこの項はエージェントが探索を行うポリシーを選択するための項である。また、第二項は事前の選好と呼ばれる項である。エージェントの目的として、事前の選好を定める必要がある。事前の選好は観測についての報酬を設定するための項である。そのためこの項はエージェントが目的を達成するためのポリシーを選択するための項

表 4: 期待自由エネルギーにおける変数・記号説明

$G(\pi, t)$	$t, \pi$ における期待自由エネルギー
$\mathbb{E}_P[Q]$	$P$ における $Q$ の期待値
$D_{KL}[P  Q]$	$P$ の $Q$ に対するカルバック・ライブラー情報量
$Q(s_t o_t, \pi)$	$o_t$ を得た時点で $\pi$ のもとで推測した $t$ における状態の近似事後信念
$Q(s_t \pi)$	$\pi$ のもとで推測した $t$ における状態の近似事後信念

表 5: ポリシー推論における変数・記号説明

$P(\pi)$	$\pi$ のカテゴリ分布
$\omega(z)$	$z$ におけるソフトマックス関数
$\gamma$	ソフトマックス関数における温度

である。これら二つの項によって、探索と目的の達成という二つの動機を扱うことが可能となる。

そしてこの期待自由エネルギーを考慮する未来の長さだけ、ポリシーごとに加算を行う(式4)。

$$G(\pi) = \sum_{t=T}^{t+L} G(\pi, t) \quad (4)$$

最後に式5によってポリシーが推論されることにより、行動を生成してエージェントは意思決定を行うことが可能となる。ただし式5で初出の変数・記号とその説明は表5に示す。

$$P(\pi) = \sigma(-\gamma G(\pi)) \quad (5)$$

ソフトマックス関数を利用することで、ポリシーごとの期待自由エネルギーをポリシーのカテゴリ分布へと変換することができる。これによって行動の決定をポリシーの推論として扱うことができる。また、 $\gamma$ はソフトマックス関数における温度を表し、現在の信念における信頼性を反映する。そして推論したポリシーから行動を生成することで、エージェントは意思決定を行うことが可能となる。

### 2.3 能動推論を用いたマッピングシステム

本報告の関連研究として、能動推論を使用して、マッピングを行っている研究文献 [7] がある。探索範囲は倉庫の内部として、一台のロボットをエージェントとみなし、ロボットにカメラを一つ搭載して、カメラの画像を用いて倉庫の内部の地図と目標地点までの経路を把

握ることが目的である。エージェントが保有している生成モデルは階層モデルとして実装されており、上位モデルでナビゲーションを、下位モデルでマッピングを行っている。また、ナビゲーションにおいては、スタート地点とゴール地点、そしてナビゲーション開始時の自身の向きが初期状態として与えられる。下位モデルでは初期状態は特に与えられずに、探索をすることで倉庫内のマップを作成していく。ただし、 $\pi$  のもとでの状態の遷移確率や  $s_t$  のもとでの  $o_t$  の生成確率については、事前に倉庫を走行することにより、強化学習により与えられている。研究で使用された倉庫の内部には視覚的に類似している場所が多く、倉庫の内部の環境や自身の位置の状態があいまいになりやすいといった環境となっている。そのため能動推論を利用しない場合には、一回の探索でマップを作成することはできず、三回の探索の結果を合わせることによってマップを作成するに至った。しかし能動推論を利用することにより、自身の位置の状態を推論しながら探索することができるため、一度の探索でマップを構成することができた。また、作成したマップを基に能動推論を利用することによって、スタート地点からゴール地点までの最短経路の導出にも成功している。一方でこの研究において制御しているのは、単一のセンサとしてのカメラと単一のアクチュエータとしてのロボットである。複数台のセンサ・アクチュエータを連携して制御することができればより多くのユースケースへの適用が可能となることが考えられる。

### 3 能動推論を利用した人物探索システム

本報告では複数のセンサ・アクチュエータを用いた人物探索に対して能動推論を適用し、CPS における仮想空間の不確実性を考慮しながら、仮想空間で情報を分析して物理空間上の複数のセンサ・アクチュエータを効率よく制御することの有効性をシミュレーションによって評価する。以下では、人物探索システムの構成を説明する。

#### 3.1 システム概要

自由エネルギー原理では生成モデルによって、外界の状態を信念として表現することが可能である。また、得られた観測から変分自由エネルギーの最小化によって生成モデルと信念を更新することが可能である。そして表現された信念を利用することで、能動推論によってその信念の不確実性を減少させる行動をとることができる。そのため CPS において、仮想空間上の情報の欠損や集約によって生じる不確実性を信念として表現し、その不確実性を減少させるような行動と観測の実現が期待される。人物探索においては、ターゲットを常にセンシングし続けることは困難であり、仮想空間上でターゲットの位置に関して情報が欠損することで不確実性が生じる。その不確実性に対して自由エネルギー原理を利用することにより、過去のセンシング結果からターゲットの位置を信念として表現することが可能である。よって、能動推論によってターゲットの不確実性を減少させる行動計画を行い、物理空間に存在する可動式カメラを適切に制御して人物探索を進めることができる。

そこで能動推論を利用した、複数の可動式カメラを利用した人物探索を行うシステムの概要を図3に示す。環境としては、都市において複数の可動式カメラが配置されている状況を想定する。そして人物の探索依頼が出された場合に、可動式カメラの映像と CPS における仮想空間に集積された情報を利用して、都市を移動するターゲットの探索を行う。能動推論を行うにあたってエージェントは複数の可動式カメラと仮想空間が構築されたサーバ上の生成モデルを保有しているとみなす。まず、可動式カメラによって、可動式カメラの観測位置におけるターゲットの有無を観測として取得してサーバへと送信する。そしてサーバ上ではターゲットの位置に対する事前の信念および得た観測、ポリシーに従って、変分自由エネルギーを計算することでターゲット位置を推論する。この時の推論結果が次の時間ステップにおけるターゲット位置に対する事前確率となる。そしてターゲット位置の推論結果からさらに、期待自由エネルギーを計算することで各カメラの観測方向の変更としてのポリシーを推論する。最後に推論したポリシーに従って、各カメラの観測方向を決定してサーバから可動式カメラへ命令を送信、可動式カメラは観測方向を変更したのち次の時間ステップで観測を行う。以上を繰り返すことにより、能動推論がターゲットの位置に関する不確実性を減少させるような行動計画を実現することで、人物探索を進めていくことが可能となる。

本報告においては、シミュレーションを行うにあたって様々な仮定を置いている。ターゲットが移動するマップはサーバ上に予め用意されているとして、事前知識として保持しているとする。カメラによる探索は本来、観測結果から個人を特定できなければ人物探索を行うことは難しくなる。しかし今回は個人を判別できるという仮定のもとで進めていく。そして可動式カメラは交差点の道路の一つを選択して、観測するものとする。観測が可能な時間は、ターゲットが交差点に進入するときと離脱するときであり、交差点の移動中は、個人を特定するための画像処理に時間が当てられているとする。街中を移動する人物には、移動先の傾向があるとする。これは狭い道路よりも広い道路を選びやすかったり、暗い道路よりも明るい道路を好んだり、見通しの悪い道路よりも見晴らしの良い道路が使われやすかったりするなどを表現する。本報告では、このようなターゲットの傾向を交差点において直進する確率と簡略化して扱う。エージェントはこの確率を学習する必要があるとする。また、ターゲットの移動には制約を設ける。具体的には、ターゲットは交差点を除いて進行方向を変化させることはないとする。交差点においては、直進確率に従って、ランダムに1つ道路を選択するものとする。そのためターゲットは可動式カメラを避けることや近づくことはせず、可動式カメラの向きの変更に対しても独立に動くものとする。また、ターゲットはステップごとに交差点の侵入と離脱を繰り返し、道路の経路長は考えないものとする。

### 3.2 システム詳細

人物探索システムの詳細について説明する。まず、マップはグラフ  $GR(N, V)$  として表現され、交差点がノード、交差点間の経路をリンクとする。この時、交差点間の接続行列  $C$  を定義することができ、交差点  $i, j$  間の経路の数が  $C$  の要素  $P_{i,j}$  となる。また、可動式カメラは交差点の一部に設置されており、カメラが設置されている交差点集合は  $N_c (C N)$  と表す。よって可動式カメラ間の接続行列を  $C_c$  として、 $C$  および  $N_c$  を用いて定義することが可能である。また、時刻  $t$  におけるカメラ  $i$  の方向は  $a_t^{(i)} (1 \leq i \leq N_c)$  として表現できる。エージェントはカメラの方向ベクトルを  $a_t = (a_t^{(1)}, \dots, a_t^{(i)}, \dots, a_t^{(N_c)})$  として推論する。これらによって状態  $s_t$  はターゲットの位置および各可動式カメラが現在向いている方向のベクトルとして表現することが可能となる。可動式カメラ  $i$  が時刻  $t$  において取得する観測は  $o_t^{(i)}$  として、ターゲットの有無を取得する。よってエージェントが扱う観測ベクトルは  $o_t = (o_t^{(1)}, \dots, o_t^{(i)}, \dots, o_t^{(N_c)})$  と表される。

また、生成モデルは2.2項で述べたように、 $P(s_0)$ 、 $P(\pi)$ 、 $P(s_t | s_{t-1}, \pi)$ 、 $P(o_t | s_t)$  に分解される。 $P(s_0)$  は初期状態として、ターゲットの初期位置とおよび可動式カメラが向いている方向となる。ターゲットの位置に関しては、仮想空間上に集積された情報を解析することによって過去の位置を得られると仮定しているが、現在の位置は不確実性が存在するとし

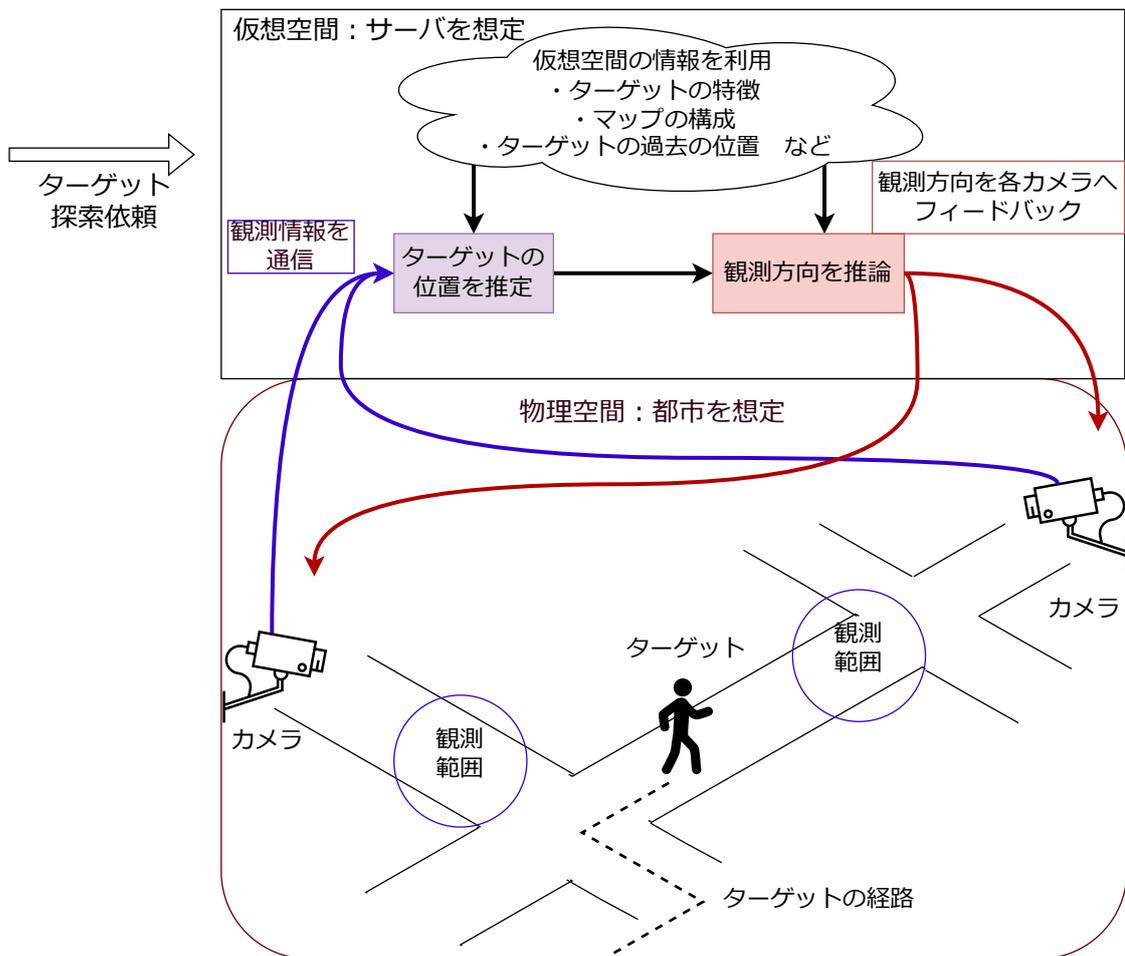


図 3: システムの概要図

て位置の候補をいくつか与える。また、可動式カメラの観測方向に関して不確実性はないとしている。 $P(\pi)$  はポリシー  $\pi$  の事前確率であり、可動式カメラの観測方向の変更の傾向である。今回は可動式カメラの観測方向の変更については制約をかけない。そのためすべての行動に等確率で割り当てている。 $P(s_t|s_{t-1}, \pi)$  はポリシー  $\pi$  のもとでの状態の遷移確率であり、マップ情報及びターゲットの移動規則、可動式カメラの観測方向の変更規則を表す。マップ情報は事前知識として保持していると仮定を置いているため、生成モデルに正確に与える。ターゲットの移動規則については、交差点において選択する道路については確率で選択されるため、生成モデルは実際に選ばれる道路を事前に知ることは出来ない。ただし、選択される道路の確率は、エージェントによってターゲットの移動規則がベイズ推定により学習されることにより更新される。 $P(o_t|s_t)$  は  $s_t$  のもとでの  $o_t$  の生成確率であり、可動式カメラの観測の精度を表す。今回は可動式カメラによる観測は正確であるという前提を置いているため、 $P(o_t|s_t)$  には不確実性はないものとする。そのため、各可動式カメラの方向と

ターゲットの位置が合致すれば、確実に観測を得られる。

ここから人物探索における能動推論について説明する。まず生成モデルは、現在時刻  $t$  において、事前信念  $P(s_{t-1})$  と遷移状態  $P(s_t | s_{t-1}, \pi)$ 、 $t-1$  における各カメラの行動  $a_{t-1}$  を利用して、式 6 によってターゲットの位置についての信念  $P(s_t)$  を生成する。

$$P(s_t) = P(s_{t-1}) \cdot P(s_t | s_{t-1}, a_{t-1}) \quad (6)$$

また、生成した  $P(s_t)$  と尤度  $P(o_t | s_t)$  を利用することで、式 7 によって可動式カメラが得られる観測とターゲットの位置状態の同時信念  $P(o_t, s_t)$  を生成する。

$$P(o_t, s_t) = P(s_t) \cdot P(o_t | s_t) \quad (7)$$

そして、実際に可動式カメラ  $i$  が得た観測  $or_t^{(i)}$  のベクトル  $or_t = (or_t^{(1)}, \dots, or_t^{(i)}, \dots, or_t^{(N_c)})$  を利用することで、式 8 に従って変分自由エネルギー  $F_t$  を計算することができる。

$$F_t = \mathbb{E}_Q[\log Q(s_t) - \log P(or_t, s_t)] \quad (8)$$

ただし  $Q(s_t)$  はターゲットの位置に対する近似の事後信念であり、すべての状態を等確率で初期化する。この  $F_t$  を最小化するように、変分サイズによってターゲット位置の事後信念  $Q(s_t)$  を推論する。また時刻  $t$  における事後信念  $Q(s_t)$  と時刻  $t-1$  における事後信念  $Q(s_{t-1})$  におけるターゲットの位置信念の変化から、生成モデルにおける遷移状態  $P(s_t | s_{t-1}, \pi)$  におけるターゲットの移動規則を表す直進確率をベイズ推定によって更新する。

続いて、将来の可動式カメラの観測方向の変更  $\pi$  の計画を行うために、事後信念  $Q(s_t)$  およびターゲットの遷移状態  $P(s_t | s_{t-1}, \pi)$ 、観測の尤度  $P(o_t | s_t)$ 、事前選好  $p(o_t)$  に従って式 9 によって期待自由エネルギー  $G(\pi, t)$  を計算する。

$$G(\pi, t) = -\mathbb{E}_{Q(o_t | \pi)}[D_{KL}[Q(s_t | o_t, \pi) || Q(s_t | \pi)]] - \mathbb{E}_{Q(o_t | \pi)}[\log p(o_t)] \quad (9)$$

そして可動式カメラの観測方向の変更  $\pi$  による  $G(\pi, t)$  を式 10 によってポリシーのカテゴリ分布とする。

$$P(\pi, t) = \sigma(G(\pi, t)) \quad (10)$$

このカテゴリ分布に従って可動式カメラの観測方向の変更  $a_t$  を選択することで、期待自由エネルギーを最小化する行動を選択することができ、ターゲットの位置状態の不確実性を下げる、すなわちターゲットを捕捉しやすい行動を選択することが可能となる。そして行動を選択することで、次の時間ステップ  $t+1$  においてもターゲットの位置の事後信念  $Q(s_t)$  を事前信念  $P(s_t)$  として扱うことで、生成モデルによる状態の生成から始めることができる。以上を繰り返すことにより、エージェントは人物探索を行うことが可能となる。

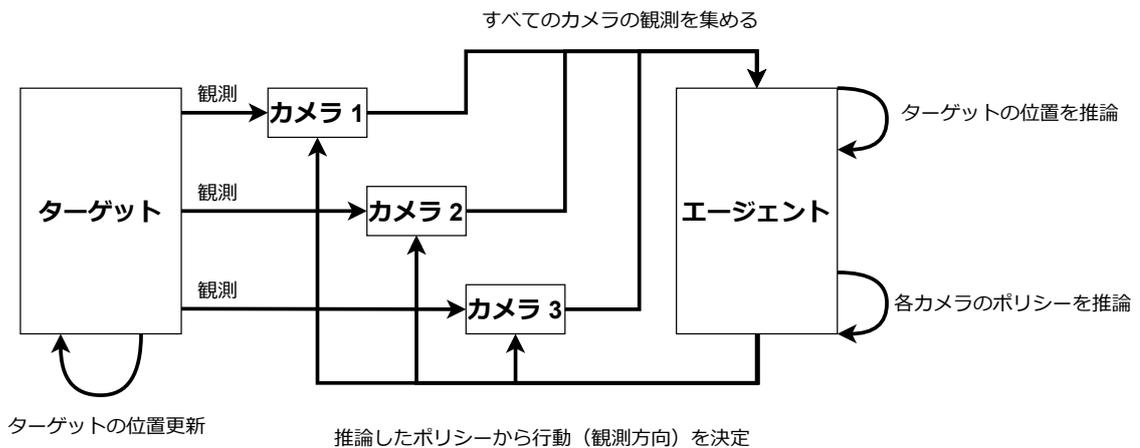


図 4: 複数の観測を利用した能動推論のシステム図

### 3.3 ポリシー推論の高速化

自由エネルギー原理における能動推論では、エージェントが複数の行動が可能である場合、ポリシーを推論する際には、すべての行動の組み合わせをポリシーとして推論する。これは行動の組み合わせによって、外界の状態に与える影響が変化する可能性があるからである。例えば本報告における人物探索システムの場合、図4のように複数の可動式カメラの観測方向の組み合わせをポリシーとして推論する。しかしすべての行動の組み合わせをポリシーとすることで、計算時間が大きくなるという問題も生じる。

一方本報告で扱うように、複数の可動式カメラの観測方向の変更を行動とする場合、複数の可動式カメラの行動の組み合わせによって外界の状態の変化が異なることはない。すなわちこの場合、各可動式カメラの行動は独立であるとみなすことができ、すべての行動の組み合わせをポリシーとして推論する必要はない。そのため、複数の行動が独立であるという条件の下では、ポリシーの推論の分散化が可能である。

能動推論を利用して、3台の可動式カメラを用いた人物位置推定を行う場合において、ポリシーの推論を分散化したシステムの概要図を図5に示す。

生成モデルにおいて、ターゲットの位置の推論を終えたのち、各カメラのポリシーを個別に推論する。そして推論したポリシーに従って、各カメラの次の観測方向を決定している。このモデルではカメラの数だけ期待自由エネルギーを計算してポリシーを推論することになり、各可動式カメラの行動の組み合わせを考えるよりも、計算時間が減少する。

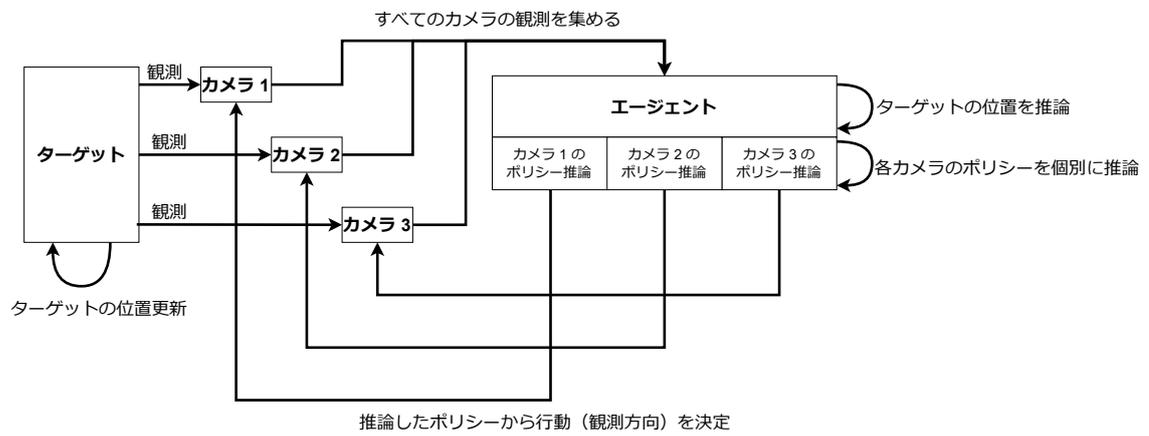


図 5: ポリシー推論を高速化したシステム図

## 4 シミュレーション評価

### 4.1 評価環境

本報告において使用したマップの一例を図6に示す。図6については、文献 [10] で利用されたマップと同じものを利用している。グレーで塗られた箇所が道路を表し、ターゲットが移動可能なことを表す。また、この時の交差点は6つであり、6つの交差点の遷移行列  $C$  は式11のように表される。

$$C = \begin{pmatrix} 0 & 3 & 0 & 0 & 0 & 1 \\ 3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 1 & 0 & 0 & 0 & 2 & 0 \end{pmatrix} \quad (11)$$

可動式カメラは交差点において、方向を一つ選択して観測することが可能である。図7では、カメラ1は青、カメラ2は緑、カメラ3はオレンジ、カメラ4は赤に塗られた場所のうちの一つを観測可能であることを表している。観測によってターゲットの有無を知ることができ、選択した方向にターゲットが存在する場合、観測値としてターゲットの位置が取得できる。また、選択した方向にターゲットが存在しない場合は、観測値はヌルとなる。

ターゲットは図6で示したマップに基づいて、ステップごとに交差点の侵入と離脱を繰り返す。交差点において方向を選択する方向は、直進確率  $prob$  に従って決定される。このとき、右折確率と左折確率は、 $(1 - prob)/2$  で表されるとする。ただし、分岐が直進と右左折のいずれかしかできない場合は、右左折の確率は  $1 - prob$  となる。また、直進ができず、右左折とどちらかしかできない場合は、それぞれ0.5の確率で推移する。

### 4.2 評価方法

探索の対象となるターゲットは移動し続けるため、常にセンシングし続けることは困難であり、ターゲットに関する情報が欠損する。そのためターゲットの状態を推定しながら、その信念において最も不確実性を減少させるように可動式カメラの観測方向を変えることで、効率よく人物探索を行うことが可能になると考えられる。よって、1つ目の指標として、ターゲットが交差点に存在するときに、可動式カメラがターゲットを観測できていた確率として、捕捉割合を利用する。捕捉割合は (捕捉割合) = (可動式カメラがターゲットをとらえた回数) / (ターゲットが可動式カメラで捉えられる範囲にいた回数) で表される。これが高いほどターゲットの位置や規則性をとらえていることを表す。そのため、ターゲットの規

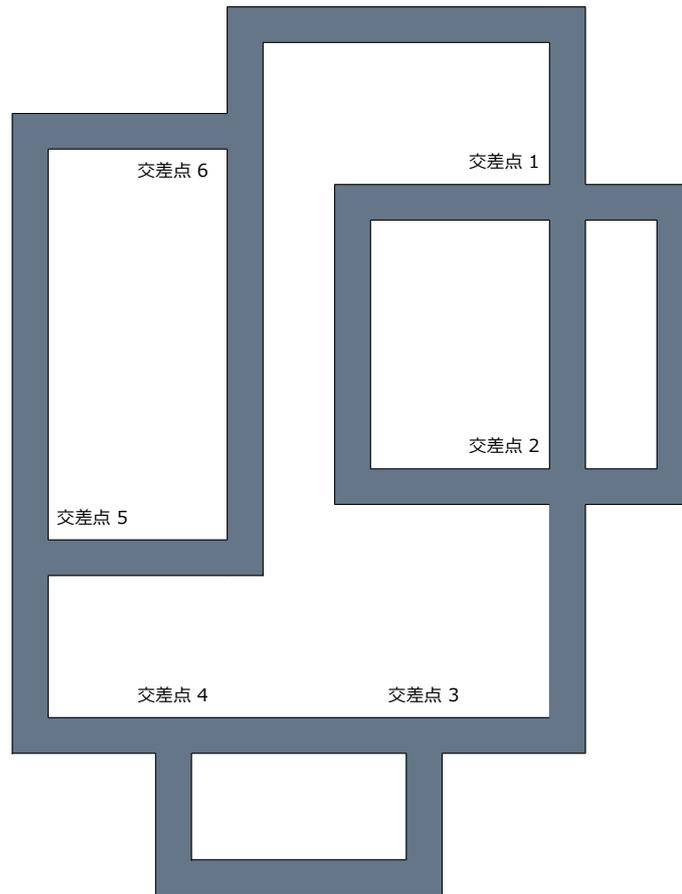


図 6: 使用したマップの一例

則性を表している直進確率を変化させたときの、捕捉割合の変化について評価する。直進確率を 0 から 1 まで変化させた場合において、評価を行う。ターゲットの進行方向に対して推定を行い、その不確実性を減少させるための行動の有効性を確かめるため、1 つ目の比較対象として、無作為に可動式カメラの方向を決定して観測する場合と比較を行う。これはターゲットの位置を信念として表現せず、行動に組み込まない手法である。比較手法の 2 つ目としてターゲットの直進確率に 0.33 を与えて、生成モデルのパラメータ更新を行わない能動推論である。これはターゲットの規則性としての直進確率を利用しない手法となる。

また、エージェントの生成モデルにおけるターゲットの規則性の学習を評価する方法としては、実際にエージェントが推定した直進確率を示す。エージェントは実際に人物探索をしながら、自身が保持している直進確率について修正していく必要がある。そのため、エージェントには実際のターゲットの直進確率と異なる確率を事前知識として与え、その事前知識をもとに人物探索を行う。そして時間ステップごとに、ターゲットの直進確率の信念がどのように推移していくかを確認する。具体的には、エージェントは、ターゲットの直進確率

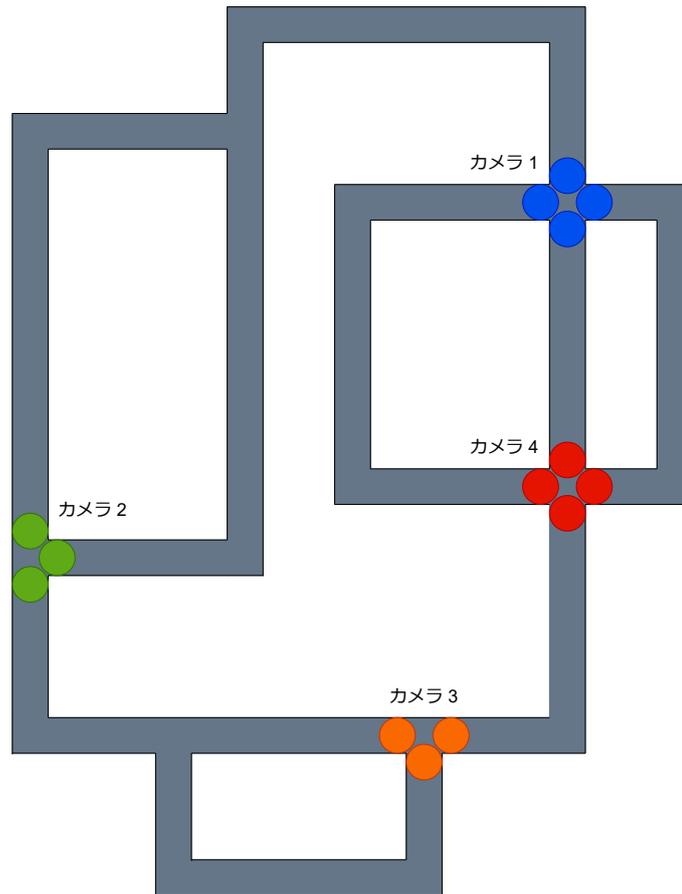


図 7: 可動式カメラの位置と観測範囲

が 0.33, 0.5, 0.1 であるという事前知識から学習を行う。一方、実際のターゲットの直進確率は 0.7 として移動させる。直進確率の信念は時間ステップごとに変化すると考えられるため、時間ステップごとの学習した直進確率の変化を示す。

### 4.3 評価結果

評価結果を図 8 に示す。図 8 では横軸がターゲットの直進確率、縦軸がターゲットの捕捉割合を表す。無作為に観測方向を変更した場合には、ターゲットの規則性は全く利用できず、ターゲットの直進確率に関わらず約 0.3 の捕捉割合となった。そして能動推論については、規則性が最も弱い直進確率 0.33 の時点で 0.65 となっている。無作為に観測方向を変更した場合よりも捕捉割合が向上した理由としては、ターゲットの位置に関する信念を保持していることが挙げられる。ターゲットの位置信念に存在する不確実性を減少させるように可動式カメラを制御することで、結果的にターゲットを捕捉しやすくなる。

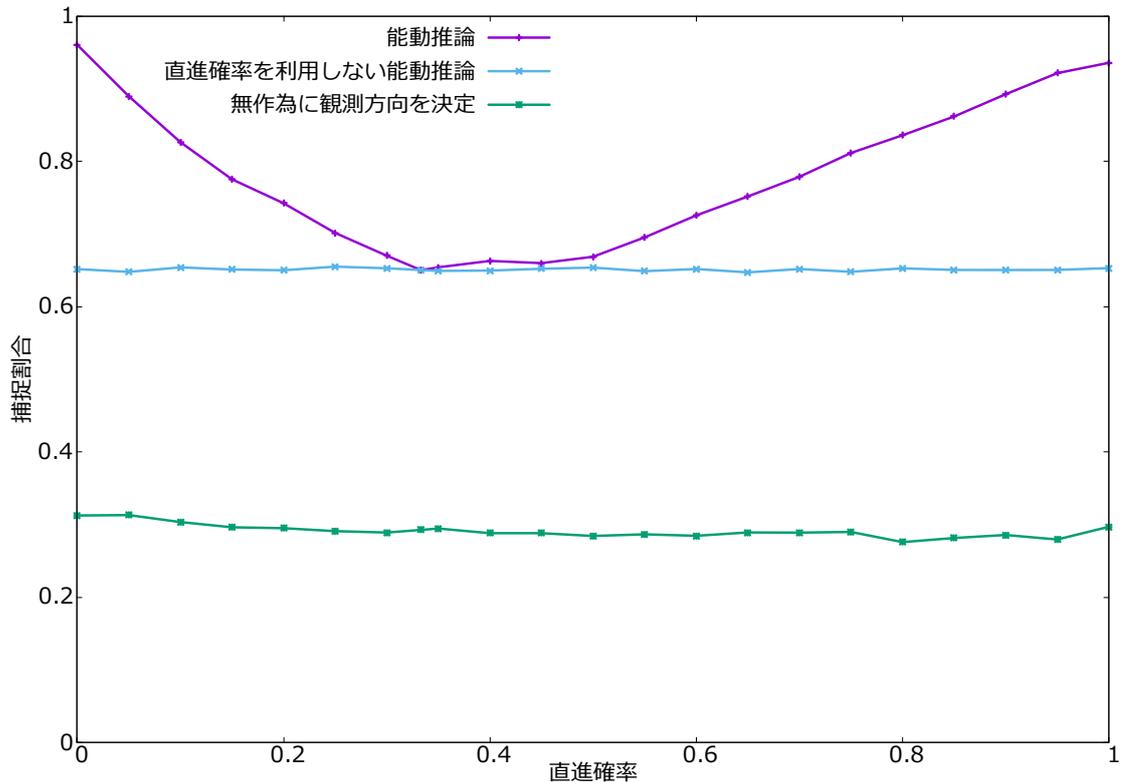


図 8: 直進確率による可動式カメラのターゲット捕捉割合の変化

また、能動推論では直進確率が変化するにしたがって捕捉割合が向上し、直進確率 1.0 のときに 0.94 の割合で、直進確率 0.0 の時に 0.96 の割合で捕捉できている。これは遷移状態  $P(s_t | s_{t-1}, \pi)$  におけるターゲットの移動規則を表す直進確率を更新することで、将来のターゲットの位置予測の精度が向上したためである。ただし能動推論においても直進確率が 0.33 から 0.5 までの範囲では、能動推論において、捕捉割合の向上がほとんど見られない。これは今回の環境において、直進確率が 0.33 のときには十字路でターゲットの移動に傾向が存在せず、直進確率が 0.5 のときには T 字路ではターゲットの移動の傾向が存在しないためであると考えられる。直進確率を利用しない能動推論においては、直進確率に関わらず捕捉割合が 0.65 となっている。すなわちターゲットの移動に規則が存在しても、生成モデルに組み込むことができなければ、ターゲットを多く捕捉することができないということであり、環境に適した行動計画を行うには生成モデルを更新する必要があることを表す。

また、時間ステップごとの生成モデルにおける直進確率の推定を図 9, 10, 11 に示す。横軸が時間ステップ、縦軸が推定した確率を表す。事前知識が 0.33 であるときには、2500 ステップほどで信念は 0.7 と推定することができている。そして 2500 ステップ以降にはその推定はほとんど変わらず、最終的に 5000 ステップで 0.72 である推定結果となった。事前

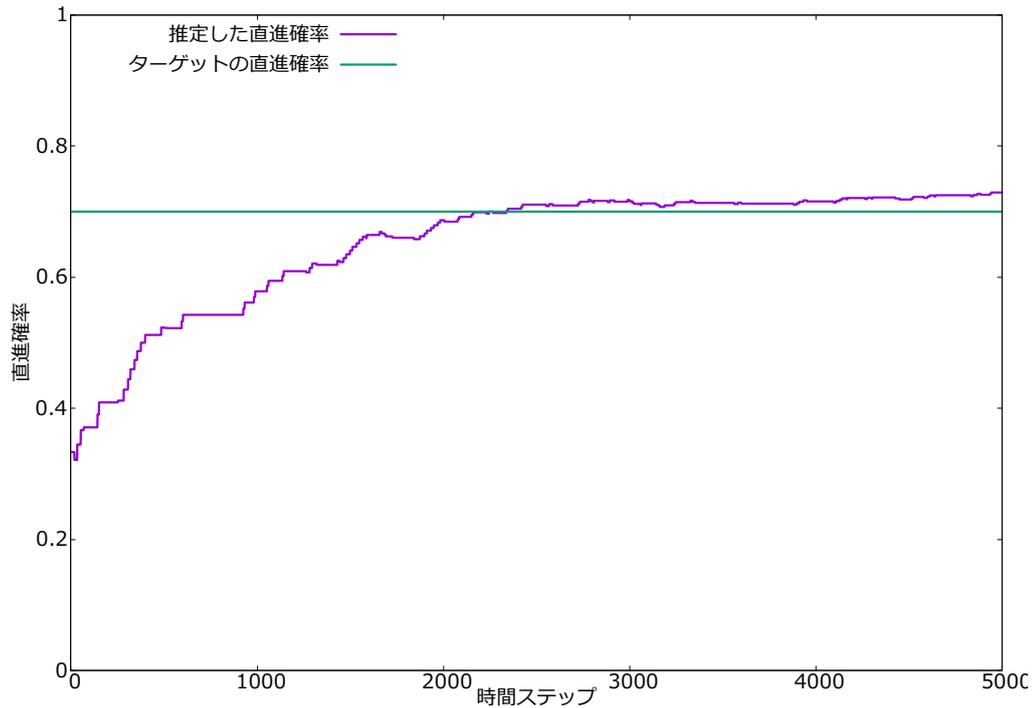


図 9: 事前知識が直進確率 0.33 であるときの生成モデルの直進確率の更新

知識が 0.5 であるときには生成モデルが推定する直進確率は、1700 ステップほどで真の直進確率に達したのちは、誤差が最大 0.04 で推移し、最終的には直進確率が 0.706 と得られている。事前知識が 0.33 である時よりも収束が早く 800 ステップほど早くなっており、この理由としては事前知識と真の確率の誤差が少なかったことが挙げられる。事前知識が 0.1 であるときには、5000 ステップのなかで直進確率が 0.7 に収束することはなく、最終的に 0.604 となった。これは事前知識と真の確率の差が大きいため、生成モデルを更新しきれなかったことが考えられる。

事前知識が真の値と乖離しているほど収束に時間がかかることから、事前知識の精度も重要であることが分かる。また、この更新は可動式カメラがターゲットを連続で捕らえることにより、その状態の変化からターゲットの移動規則として直進確率を推定している。そのためより効果的に更新を行うためには、ドローンやロボットなどのモビリティを有するセンサ・アクチュエータを利用するなどして、より能動的に推論を行う必要があると考えられる。

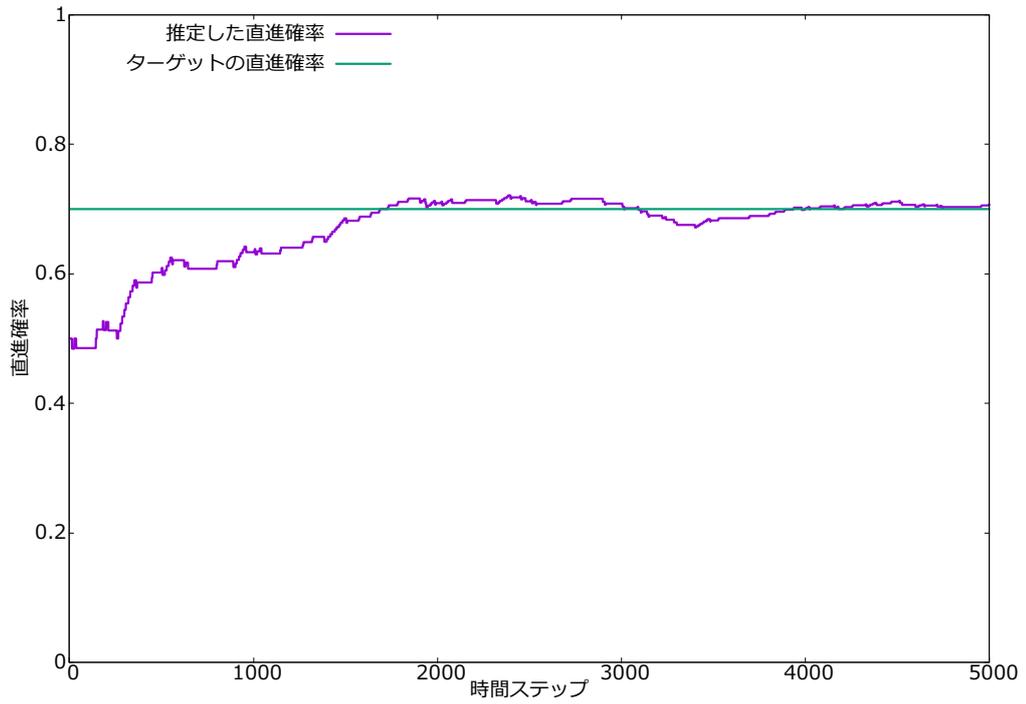


図 10: 事前知識が直進確率 0.5 であるときの生成モデルの直進確率の更新

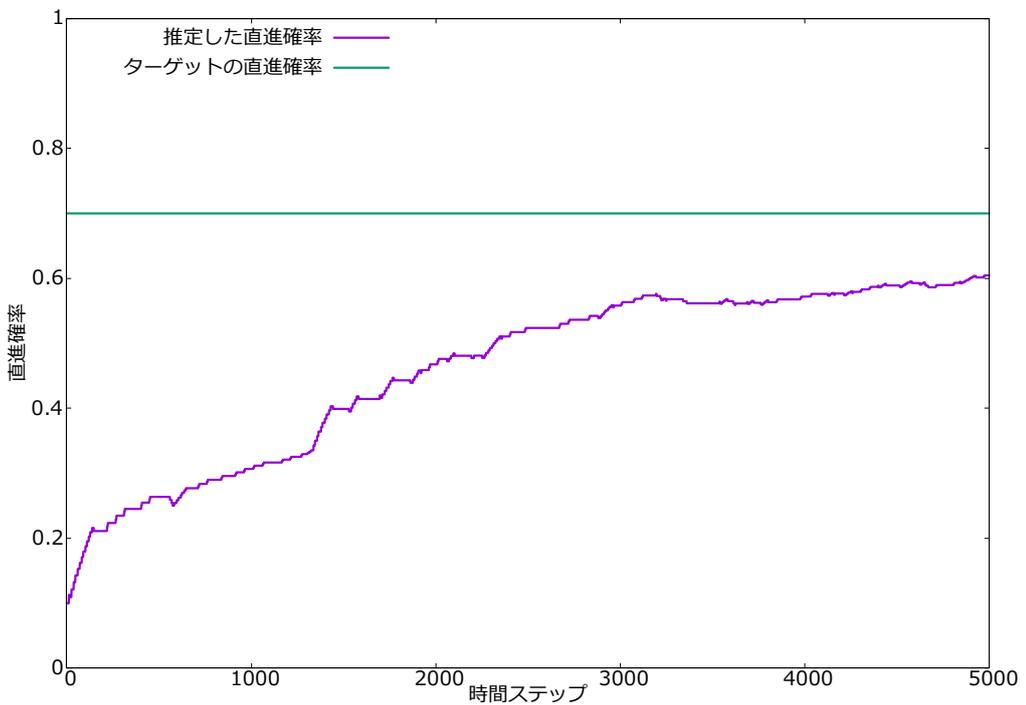


図 11: 事前知識が直進確率 0.1 であるときの生成モデルの直進確率の更新

## 5 おわりに

本報告では、仮想空間における情報の不足に対して、物理空間へと働きかけを行うことでセンサ・アクチュエータを効率よく利用して、情報を適切に収集することを目的としていた。その中でユースケースの 1 つとして、街中において複数の可動式カメラを利用しての人物探索を例にして能動推論を利用したシステムの評価を行った。その結果、能動推論を利用しない場合よりもターゲットを効率よく捕捉することが可能であったことから、不足している情報を適切に判断し、物理空間へ効率よく働きかけていたと考えられる。また、人物探索をしながら、その移動している人の傾向をとらえ、その傾向をもとにシステムの修正を行えることも確認した。

今後の課題として、より厳密な条件下での評価を検討する必要がある。例えば、ターゲットの移動規則が人の容姿、年齢、性別などで変化する場合は考慮する必要や、場所や時間帯など様々な要因で変化する場合は考慮する必要がある。あるいは実際の人物探索システムにおいては、複数人を同時に探索する場合は存在すると考えられるが、そのような状態がより複雑になる場合は検討する必要がある。また、ユースケースの精査が必要である。今回は人物探索を例に挙げたが、ほかにも交通情報推定や事件・事故の把握など、街中において情報の把握が必要な様々な事象が考えられる。さらに、災害時におけるユースケースとして、利用可能・不可能である道路の判別や救助を必要としている人の把握など、より高度な探索が求められるシナリオも考える必要がある。そしてセンサについても、今回は RGB カメラを想定しているが、深度カメラ、Lidar デバイスなどを考える必要がある。また、可動式カメラだけでなく、ドローンやロボットなど、モビリティを有するアクチュエータにセンサを搭載して、利用することも検討する必要がある。

## 謝辞

本報告を遂行するにあたり、大阪大学大学院情報科学研究科村田正幸教授にはご多忙の中、終始適切なお助言を賜りました。ここに深謝の意を表します。同研究科荒川伸一准教授には、平素より研究の方針や本報告の作成をはじめとした、様々な面で適切な御指導を頂きました。厚く御礼申し上げます。最後に、村田研究室の皆様には、本報告の執筆にあたり多くのお助言、激励を頂き、家族、友人には日々の学生生活を支えてくださりました。本当にありがとうございました。

## 参考文献

- [1] S. Zanero, “Cyber-physical systems,” *Computer*, vol. 50, pp. 14–16, Apr. 2017.
- [2] D. G. Pivoto, L. F. de Almeida, R. da Rosa Righi, J. J. Rodrigues, A. B. Lugli, and A. M. Alberti, “Cyber-physical systems architectures for industrial internet of things applications in industry 4.0: A literature review,” *Journal of Manufacturing Systems*, vol. 58, pp. 176–192, Jan. 2021.
- [3] 内閣府, “Society 5.0.” [https://www8.cao.go.jp/cstp/society5\\_0/index.html](https://www8.cao.go.jp/cstp/society5_0/index.html), 2022. (Accessed: 2023-12-10).
- [4] X. Tang, K. Yang, H. Wang, J. Wu, Y. Qin, W. Yu, and D. Cao, “Prediction-uncertainty-aware decision-making for autonomous vehicles,” *IEEE Transactions on Intelligent Vehicles*, vol. 7, pp. 849–862, July 2022.
- [5] Z. Pang, J. Li, P. Tokmakov, D. Chen, S. Zagoruyko, and Y.-X. Wang, “Standing between past and future: Spatio-temporal modeling for multi-camera 3d multi-object tracking,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 17928–17938, June 2023.
- [6] K. Friston, F. Rigoli, D. Ognibene, C. Mathys, T. Fitzgerald, and G. Pezzulo, “Active inference and epistemic value,” *Cognitive Neuroscience*, vol. 6, pp. 187–214, May 2015.
- [7] O. Çatal, T. Verbelen, T. Van de Maele, B. Dhoedt, and A. Safron, “Robot navigation as hierarchical active inference,” *Neural Networks*, vol. 142, pp. 192–204, May 2021.
- [8] K. Friston, “The free-energy principle: a unified brain theory?,” *Nature Reviews Neuroscience*, vol. 11, pp. 127–138, Jan. 2010.
- [9] R. Smith, K. J. Friston, and C. J. Whyte, “A step-by-step tutorial on active inference and its application to empirical data,” *Journal of Mathematical Psychology*, vol. 107, pp. 1–60, Apr. 2022.
- [10] R. He, A. Bachrach, and N. Roy, “Efficient planning under uncertainty for a target-tracking micro-aerial vehicle,” in *Proceedings of 2010 IEEE International Conference on Robotics and Automation*, pp. 1–8, May 2010.