

大阪大学 1

能動推論に基づく広域エリア人物探索手法の有効性の評価

大阪大学大学院 情報科学研究科 情報ネットワーク学専攻
小菅 瑛斗 荒川 伸一 村田 正幸

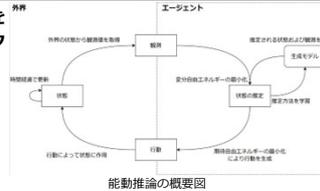
情報ネットワーク研究会 2024.10.11

研究背景 2

- **物理空間の情報の把握**
 - 例：医療現場や工場、農場の状況把握などの工業目的
 - 人手不足の解消、大規模システムの管理
 - 例：交通情報推定や迷子 / 徘徊老人 / ペット探索などの生活目的
 - Society 5.0、スマートシティ
 - 物理空間の情報を集積することが重要
- **集積された情報に不確実性が生じるのは不可避**
 - センサによるサンプリングは有限
 - 情報を保存するコンピューティングリソースの制約
- **情報の不確実性を表現し、減少させる枠組みが必要**
 - 得られた観測により欠損情報の不確実性を表現
 - 不確実性を減少させるための行動を推論
 - 観測と行動を繰り返して効果的に情報の不確実性を減少

能動推論 [1] 3

- **エージェントの環境推定と行動推論をモデル化した統一的なフレームワーク**
 - 変分自由エネルギー：環境を信念として不確実性を含め推定
 - 期待自由エネルギー：信念の不確実性を減少させる行動を推論
 - エージェントは観測と行動の繰り返しにより環境を理解
 - 環境状態の予測を行う生成モデルを更新
- **能動推論を利用して倉庫内でマッピングを行う先行研究 [2]**
 - 倉庫内のマップを環境状態として表現
 - マップ内で 1 台のロボットを制御して前方に取り付けられたカメラで環境を観測
 - 観測と制御の繰り返しによりトポロジカルマップを生成

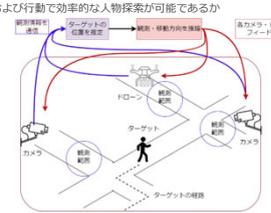


能動推論の概要図

[1] K. Friston, F. Rigoli, D. Ogburne, C. Mathys, T. Fitzergald, and G. Pezzulo, "Active inference and epistemic value," Cognitive Neuroscience, vol. 6, pp. 187-214, May 2015.
[2] O. C. and T. Verbeke, T. Van de Marel, B. Dhoosh, and A. Saffron, "Robot navigation as hierarchical active inference," Neural Networks, vol. 142, pp. 192-204, May 2021.

研究目的と研究手法 4

- **研究目的**
 - 人物探索における能動推論の有効性の評価
 - 可動式カメラやドローンの観測情報から物理空間の状態（ターゲットの位置）を信念として推定
 - 可動式カメラとドローンをアクチュエートする行動計画
 - 観測と行動の相互フィードバックにより、少ない観測および行動で効率的な人物探索が可能であるか
- **研究手順**
 - 人物探索システムの策定
 - 能動推論のフレームワークを適用
 - シミュレーション評価
 - ターゲットの捕捉までの時間
 - 能動推論により生成される行動の妥当性



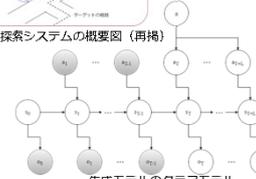
人物探索システムの概要図

本研究で想定する人物探索システム 5

- **複数の可動式カメラとドローン、マップ、ターゲットで構成**
 - 離散ステップで生成モデルを作成
 - 観測 o_t ：ターゲットの有無
 - 状態 s_t ：マップ上でのターゲットの位置
 - マップは離散的なグリッドで表現
 - 行動 a_t ：ドローンの移動・カメラの向き変更
- ターゲット規則を利用 [3]
 - 地点による差
 - 移動目的による差
 - 時間帯による差
- こうした移動規則から生成モデルを更新
 - ターゲットの位置予測に利用



人物探索システムの概要図 (再掲)



生成モデルのグラフモデル

[3] 富士交通省, "まちの活性化を測る歩行者数調査のガイドライン," <https://www.mlit.go.jp/commos/001239608.pdf> 2018

人物探索システムにおける能動推論 6

- **ターゲットの位置推定は変分自由エネルギー F_t の最小化により実行**
 - $F_t = E_q[\log Q(s_t)] - \log P(o_t, s_t)$
 - $Q(s_t)$ がターゲットの位置の信念を表現
 - ターゲットの位置をグリッドごとに確率的に表現
 - 最小化によりターゲットの位置の推定を推進
- **カメラやドローンの行動計画は期待自由エネルギー $G(\pi, t)$ の最小化により実行**
 - $G(\pi, t) = -E_q[D_{KL}[Q(s_t|o_t, \pi)]|Q(s_t|\pi)] - \frac{E_q[\alpha_t]}{\alpha_t} \log P(o_t)$
 - 情報利得 事前の選好
 - 第一項目 (情報利得) は減少させられるターゲットの位置の不確実性を表現
 - 第二項目 (事前の選好) はターゲットを捕捉できる可能性を表現
 - 最小化によりターゲットを効率的に捕捉する行動を計画

能動推論を適用した人物探索の流れ

7

- 観測と可動式カメラ、ドローンの制御を繰り返すことにより効率的に人物探索
 - 可動式カメラとドローンの観測結果によりターゲットの位置信念 $Q(s_t)$ および生成モデルを更新
 - 変分自由エネルギー $F_t = E_Q[\log Q(s_t) - \log P(a_t, s_t)]$ を最小化
 - 信念 $Q(s_t)$ に従って、次の可動式カメラの方向とドローンの移動計画 π を決定
 - 期待自由エネルギー $G(\pi, t) = -E_Q[\log P(a_t | Q(s_t, \pi))] - E_Q[\log P(a_t)]$ 進行方向 ターゲット 観測範囲
 - 観測と行動を繰り返すことにより可動式カメラ及びドローンにより捕捉が可能

ターゲット位置の信念更新例

直進: 33.3%	直進: 0%	直進: 32.8%	直進: 0%
右折: 33.3%	右折: 50%	右折: 34.3%	右折: 100%
左折: 33.3%	左折: 50%	左折: 32.8%	左折: 0%

カメラの観測誤差

8

- カメラでの探索において、人物を認識できない可能性 (観測誤差)
 - 観測誤差は変分自由エネルギー F_t に影響
 - $F_t = E_Q[\log Q(s_t) - \log P(a_t, s_t)]$
 - $= E_Q[\log Q(s_t) - \log \hat{P}(a_t | s_t) P(s_t | s_{t-1}, \pi) Q(s_{t-1})]$
 - $= E_Q[\log Q(s_t) - \log a P(a_t | s_t) P(s_t | s_{t-1}, \pi) Q(s_{t-1})]$
 - $= E_Q[\log Q(s_t) - \log P(a_t | s_t) Q(s_{t-1}) - \log a P(s_t | s_{t-1}, \pi)]$
 - $= E_Q[\log Q(s_t) - \log P(a_t | s_t)]$ は観測誤差を含む尤度 ($0 \leq a \leq 1$)
 - $P(s_t | s_{t-1}, \pi)$ がターゲットの遷移確率であることから $\log a P(s_t | s_{t-1}, \pi)$ は負となり F_t の値は増大
 - 信念の不確実性の増大
 - 一方でターゲット規則を表す $P(s_t | s_{t-1}, \pi)$ の不確実性が減少すれば F_t の値は減少
 - ターゲット規則を正確に推定することにより観測誤差の影響を緩和

評価環境

9

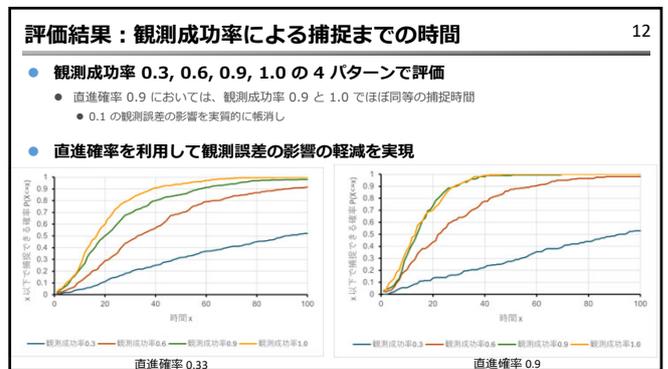
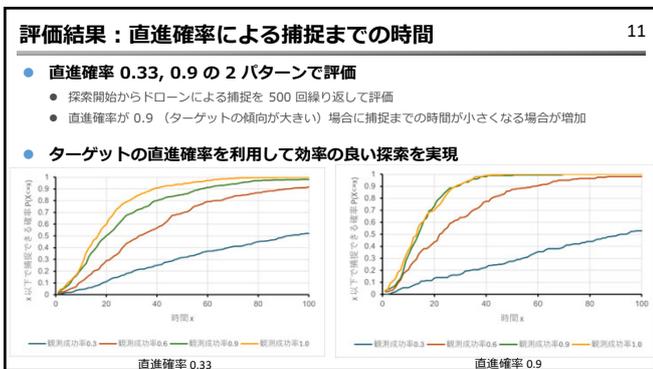
- マップ
 - 右回のように複数のグリッドにより構成
- 可動式カメラ
 - 配置された交差点の1つのグリッド(色付き)を一つ決定して観測
 - ターゲットを観測できない可能性(認識誤差)が存在
- ドローン
 - ドローン待機所が初期位置
 - 最大グリッド2マス移動可能
 - 認識誤差が存在
- ターゲット
 - グリッド1マスを移動
 - 交差点においては、いずれか一つの道を選択
 - 移動の規則性: 交差点において直進確率(直進を選択する確率)を特つ

マップと可動式カメラの観測可能位置

評価方法

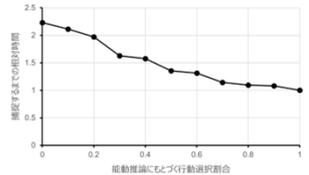
10

- ターゲットの直進確率およびカメラの観測成功率別に捕捉時間を比較
 - ターゲットが交差点において直進する確率を変化
 - 直進確率 0.33, 0.9 の2パターンで評価
 - ターゲットが存在した際に、観測に成功する確率(観測成功率)を変化
 - 観測成功率 0.3, 0.6, 0.9, 1.0 の4パターンで評価
 - ドローンによる捕捉までの時間により評価
 - 一度の観測および行動が単位時間ステップ
 - ドローンによる捕捉でターゲットの追跡が可能
 - 捕捉までの時間は累積度数分布として表現
- 能動推論によりターゲットを捕捉する行動を計画できているかを評価
 - 能動推論に基づく行動選択割合を変化
 - ドローンによる捕捉までの時間により評価



評価結果：能動推論にもとづく行動選択割合による捕捉時間 13

- 能動推論にもとづく行動選択割合に対する相対捕捉時間
 - 能動推論にもとづく選択割合が 1.0 の場合の捕捉までの時間を 1 とみなして相対時間で評価
 - 行動選択割合が 1.0 の場合に捕捉までの時間が最も短い
- 能動推論によりターゲットを捕捉するための効率の良い行動を計画



能動推論にもとづく行動選択割合に対する相対捕捉時間

まとめと今後の課題 14

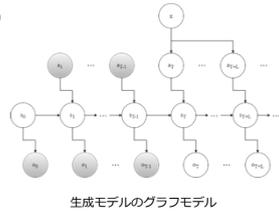
- 能動推論による複数の可動式カメラとドローンによる人物探索を実施
 - 人物探索においてターゲットの位置の不確実性を減少させるように可動式カメラとドローンを制御
 - ターゲットの規則性を可動式カメラやドローンの制御に反映
 - 単純化した環境で有効性が表れていることを確認
- 今後の課題
 - 様々なユースケース
 - 交通情報推定
 - 事件・事故の把握 など
 - 様々なセンサの特性に応じた観測情報の利用
 - 現在はターゲットの存在の有無の情報のみ
 - 深度カメラ、Lidar デバイスなどの様々な情報を提供するセンサ
 - より大規模なエリアでの検証（スケール性）
 - 特に期待自由エネルギーの計算時間がアクチュエータの数に比例
 - 一定のエリアごとに分割するなどの分散・階層化が必要

能動推論における生成モデル 15

- 予測される未来の状態および観測を生成
 - $P(\delta, s|\pi) = P(s_0)P(\pi) \prod_{t=1}^T P(s_t|s_{t-1}, \pi) \prod_{t=0}^T P(o_t|s_t)$

生成モデルの構成要素

記号	説明
$P(s_0)$	初期のターゲット位置
$P(\pi)$	カメラの方向・ドローンの移動
$P(s_t s_{t-1}, \pi)$	ターゲットの移動規則
$P(o_t s_t)$	観測におけるターゲットの有無の判別



生成モデルのグラフモデル

不確実性を表現するための仕組み 16

自由エネルギー原理 [4]

- 生物の知覚や学習を変分自由エネルギーの最小化により説明
 - 生物（エージェント）は環境を表現する生成モデルを保持
 - 観測情報（五感からの情報）から将来の観測情報や環境の状態を予測
 - 得られた観測情報から生成モデルを更新（学習）
- 変分自由エネルギー F_t の最小化 環境の状態予測
 - $F_t = E_Q[\log Q(s_t) - \log P(o_t, s_t)]$
 - エージェントは真の状態を知ることが出来ないため近似事後信念 $Q(s_t)$ を導入
 - 変分ベイズによって生成モデルから得られる $P(o_t, s_t)$ に近似 \Rightarrow 変分自由エネルギーの最小化

[4] K. Friston "The free-energy principle: a unified brain theory?" Nature Reviews Neuroscience, vol. 11, pp. 127–138, Jan 2010.

先行研究 17

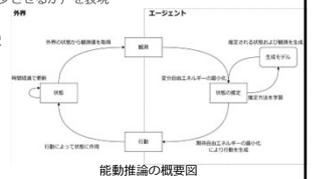
- 能動推論を利用して倉庫内でマッピングを行う先行研究 [2]
 - 倉庫内のマップを状態として表現
 - マップ内で 1 台のロボットを制御して前方に取り付けられたカメラで環境を観測
 - 観測と制御の繰り返しによりトポジカルマップを生成
 - さらにトポジカルマップを利用することでスタートとゴールの最短経路を導出
- 先行研究と本研究の違い
 - 先行研究では単一のセンサとアクチュエータを利用
 - 本研究では複数のセンサとアクチュエータを利用
 - より効率的に情報を集め、効率的に不確実性を減少
 - より広範な応用への期待

[2] O. C. et al, T. Verbeke, T. Van de Maelde, B. Dhoedt, and A. Saffon, "Robot navigation as hierarchical active inference," Neural Networks, vol. 142, pp. 192–204, May 2021.

能動推論 18

自由エネルギー原理の環境推定と行動推論をモデル化したフレームワーク

- 変分自由エネルギー：環境を信念として推定
- 期待自由エネルギー：信念の不確実性を減少させる行動を推論
 - $G(\pi, t) = -E_Q[D_{KL}[Q(s_t|o_t, \pi) || Q(s_t|\pi)]] - E_{Q(o_t|s_t)}[\log P(o_t)]$
- 期待自由エネルギーは将来予想される変分自由エネルギーの期待値
- 第一項目は情報利得（どの程度信念の不確実性を減少させるか）を表現
- 第二項目は事前選好（報酬）を表現
- $G(\pi, t)$ を最小化するポリシー π （行動計画）を選択
- 観測と行動の繰り返しにより環境を理解
 - エージェントが主体となって能動推論を実行
 - 状態の予測を行う生成モデルを更新



能動推論の概要図

現在実用化されている人物探索サービス

19

- **伊丹市：まちなかミマモルメ [5]**
 - 高齢者および小学生向けのサービス
 - 街中に設置されたビーコン受信機が小型タグを検知
- **セコム：セコムみまもりホン [6]**
 - 小型デバイスに搭載された GPS により居場所を把握
 - 本人がブザー用ストラップを引っ張ることで緊急通報
- **本人がタグ・デバイスを保持していることが前提**
 - 忘れる、落とす、故障などのリスク
- **誰もが安心して生活できるためにはデバイスやタグによらない探索が必要**
 - 可動式カメラやドローンによる探索

[5] 伊丹市, "位置情報通知サービス「まちなかミマモルメ」," https://www.city.itami.lg.jp/SOSIKU/TOSHIKOTU/TOSI_ANZEN/camera/mimamorme.html, 2022

[6] SECOM, "高齢者向けケイタイとして最適な「セコムみまもりホン」のご紹介," <https://www.secom.co.jp/mimamori/column/03.html>, 2019

複数の可動式カメラによるターゲット位置推定

20

- **先行研究 [2] では単一のセンサ・アクチュエータを利用**
 - 複数のセンサ・アクチュエータの利用によってより多くの応用への適用が期待
- **複数のアクチュエータの行動推論**
 - 全ての行動の組み合わせを考えるため計算量大
- **行動推論は各アクチュエータごとに個別に推論**
 - 組み合わせにより環境が変化しない場合
 - アクチュエータごとに期待自由エネルギーを計算し最小化する行動を推論



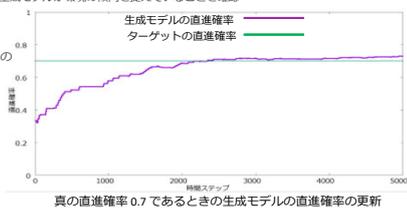
ポリシー推論を高速化したシステム

[2] O. C. amk, T. Verbelen, T. Van de Meele, B. Dhoedt, and A. Saffon, "Robot navigation as hierarchical active inference," *Neural Networks*, vol. 142, pp. 192–204, May 2021.

評価結果：生成モデルの更新

21

- **ターゲットの真の直進確率が 0.7 である場合**
 - シミュレーション開始時の初期値としてターゲットの直進確率 0.33 を与える
 - エージェントによる生成モデルの更新：0.33 ⇒ 0.72
 - 推定した状態から生成モデルを更新
 - 観測と行動を繰り返すことで生成モデルが環境の傾向を捉えていることを確認
- **課題**
 - 真の直進確率を得るまで多くの観測を要する



真の直進確率 0.7 であるときの生成モデルの直進確率の更新