Maximum entropy searching

Rui Jiang1, Hui Zhou2, Han Wang2, Shuzhi Sam Ge1,2

1Institute for Future (IFF), Qingdao University, Qingdao 266071, People’s Republic of China
2School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore

Abstract: This study presents a new perspective for autonomous mobile robots path searching by proposing a biasing direction towards causal entropy maximisation during random tree generation. Maximum entropy-biased rapidly-exploring random tree (ME-RRT) is proposed where the searching direction is computed from random path sampling and path integral approximation, and the direction is incorporated into the existing rapidly-exploring random tree (RRT) algorithm. Properties of ME-RRT including degenerating conditions and additional time complexity are also discussed. The performance of the proposed approach is studied, and the results are compared with conventional RRT/RRT* and goal-biased approach in 2D/3D scenarios. Simulations show that trees are generated efficiently with fewer iteration numbers, and the success rate within limited iterations has been greatly improved in complex environments.

1 Introduction

Navigating autonomous robots from the origin to the target with satisfactory performance is a crucial problem which has not been fully solved. One of the challenges in robot navigation lies in the area of motion planning, which can be naturally separated into two main tasks: path searching and path following. Path searching aims to find a collision-free (either topological or geometric) path given two end-points, and then path following ensures the robot is tracking the pre-set route [1, 2]. Apart from path searching, trajectory planning or motion planning has been proposed as a direct approach to combine path searching and path following together. A trajectory contains speed information along the path, thus kinematics and dynamics of the robotic systems are usually considered in planning algorithms [3]. To find a feasible path or trajectory, objectives are designed from different perspectives to guide path searching. Those objectives include optimality with regard to energy, execution time and trajectory length [4–8].

Unfortunately, high-dimensional, unstructured or large-scale environments usually result in unacceptable execution time if the optimal path is required. Sampling-based approaches, which do not try to represent environment explicitly but only plan the path according to collision detection results, are more frequently used in practice, with asymptotic optimality [9]. The ‘environmental-independent’ feature of sampling-based approaches simplifies the algorithm design but may cause performance degradation as environment information (which could be regarded as a map) is extremely helpful in searching. In this work, we aim to strengthen sampling-based searching by considering environment information such that the algorithm performance could be improved especially in complex environments with a narrow passage or U-shape obstacles.

For years, people are keen to link robot behaviours with intelligence. Some researchers believe that there is a possible deep connection between intelligence and entropy maximisation: in [10], by proposing the physical relation between intelligence and causal entropy maximisation and applying it to physical systems, two intelligent behaviours have been produced spontaneously: tool use and social cooperation. Inspired by Wissner-Gross and Freer [10], this paper further presents a path searching framework driven by entropic intelligence. The main contributions are highlighted as follows:

1. Based on the concept of causal entropy, the entropic force is introduced in robot navigation as a new perspective which aims at maximising future feasible paths and minimising external constraints from the environment. The entropic force, which is computed according to the results of random path sampling in a finite time horizon, enables perception of the surrounding environment without formulating environment explicitly.
2. Based on the entropic force, the maximum entropy-biased rapidly-exploring random tree (ME-RRT) is proposed to grow random tree towards not only the goal or unsearched areas but also entropy maximisation. Compared with RRT and goal-biased RRT, the proposed approach avoids fully random sampling or naive goal-biasing behaviours. Environment information is incorporated in the searching algorithm such that macrostates with more feasible paths in a finite time horizon is expected to be explored.
3. Comparative experiments in simulated environments have been conducted for a disc shape holonomic mobile robot in 2D and 3D spaces among ME-RRT, RRT, goal-biased RRT and their optimal version in three scenarios such that the efficiency and robustness of the proposed approach have been demonstrated.

The paper proceeds as follows. We first discuss related work with an emphasis on existing solutions to sampling-based path searching. Subsequently, Section 3 presents an overview of the concept of causal entropy and the force towards entropy maximisation. Section 4 provides the main flow of the proposed entropic searching approach, where random path sampling, path integral approximation and extension function design are presented. Properties of the newly proposed ME-RRT is discussed in Section 5. Experimental results are demonstrated in Section 6. Finally, Section 7 concludes the paper.
RoadMap planner (PRM) is one of the initial attempts which demonstrated its potentials [11–14]. While PRM tends to explore the entire free space, single-query planners have been proposed such that searching efficiency is improved. The expansive-spaces tree (left), a single-query, bidirectional, lazy collision-checking (SBL) planner has been designed [16]. The well-known rapidly-exploring random tree (RRT) [9] and its variants [17–19] have been implemented in robotic manipulators and mobile robots [20–22].

Performance of sampling-based searching algorithms is related to the complexity of configuration space. Though the ‘Goals NonReachable with Obstacles Nearby (GNRON)’ problem [33] does not theoretically exist for sampling-based searching, the searching efficiency is of concern when a narrow passage becomes the only feasible channel between initial and desired configurations. A primitive way is to focus on configuration space geometry. For instance, Baginski [24] proposed to widen narrow passages by slightly dilating obstacles. In [25, 26], the generalised voronoi diagrams are extracted and samples with small interspace around the branches of the diagrams are preferred. Another perspective is to evaluate on samples and rule out those unpromising ones. In [27], the bridge test is used in sampling process so that samples in wide-open free space are more likely to be discarded. Inspired by the bridge test, Du et al. [28] further improve RRT to identify and estimate the width and depth of the narrow passage by ‘X-test’. In addition, researchers are also making efforts in proposing new sampling strategies to enhance the probability of ‘good’ samples and reject unpromising ones. Quasirandom alternatives are proposed to provide better discrepancy than conventional random samples [29]. To overcome the deficiency of aimless exploration, biased sampling has been introduced in [30], where weighted hybrid sampling is adopted to balance exploration and goal reaching behaviours. In [31], a path-biased sampling strategy is proposed such that system’s non-holonomic constraints can be incorporated after a feasible any-angle path is generated. Kobbelt [32] proposed to bias sampling in the direction of optimal trajectories based on an adaptive model. Trying to obtain the optimal path, the RRT* [33] and its variants [34–37] have been proposed with guaranteed asymptotic optimality.

Recently, benefiting from the development of hardware performance and learning models, learning-based approaches, which aim to imitate human path searching behaviours, have become popular. An end-to-end motion planner has been presented in [38], where a deep convolutional neural network is designed to process data from the onboard laser scanner. Finn and Levine [39] have proposed a deep convolutional long short term memory (LSTM) model to enable vision-based robot planning and manipulation without detailed human supervision. In [40], the role playing learning scheme has been proposed to enable mobile robot planning with a companion. Although deep learning models have strong power in approximating complex models, the data fed into the model would significantly influence the resulting planner performance. To eliminate the uncertainty of model quality caused by training data, researchers are also aiming at uncovering the intrinsical mechanism of intelligent searching behind the data.

How to select searching direction is a problem worthy of study. In literature, the goal is the commonly used direction but may be insufficient especially in environments with narrow passages and U-shape obstacles, as illustrated in Fig. 1. To overcome the mismatchedness of goal-biased tree expansion, curiosity [41–43] has been proposed as one possible direction. However, the definition and computation of ‘curiosity’ is still a problem under study. In this work, we propose to bias searching direction towards entropy maximisation, which is computable, thanks to the clear concept.

### 3 Preliminaries

This section is prepared for a comprehensive understanding of maximum entropy searching. It first introduces microstates and macrostates. With basic assumptions, causal entropy is then defined on each macrostate within a finite time horizon. Lastly, after expressing the causal entropic force, we discuss the implication of causal entropy maximisation.

#### 3.1 Microstates and macrostates

We consider a general state equation \( \dot{x} = f(x(t), u(t), t) \), where \( x(t) \) and \( u(t) \) denote the state vector and the control vector, respectively; \( f(\cdot) \) represents a C∞ function with respect to \( x(t), u(t) \) and \( t \). Prior to introducing the causal entropy, the following concepts are presented, as shown in Fig. 2a:

**Definition 1: (Microstate):** Given a finite time horizon \( \tau \), a microstate \( x(t) \) is defined as a feasible path in state space where \( 0 \leq t \leq \tau \).

**Definition 2: (Macrostate):** Given a set \( S \) as all possible microstates of the dynamic system, an equivalence relation is defined as \( x_i(t) \equiv x_j(t) \) if \( x_i(0) = x_j(0) \) in the finite time horizon. A macrostate \( X \) is defined as an equivalence class of elements of \( S \) under the equivalence relation.

#### 3.2 Entropic force and causal entropy

In a physical system, entropy with respect to macrostate is defined as \( S(X) = -\sum_i P(x^t) \ln P(x^t) \), for a random macrostate \( x^t \) with probability \( P(x^t) \). For any current macrostate \( X_0 \), entropic force \( F(X_0) \) is defined as

\[
F(X_0) = TVX S(X)|_{X_0} \tag{1}
\]

where \( T \) denotes the equivalent temperature; \( S(X) \) is the entropy associated with macrostate \( X \).

Causal entropy has been proposed in [10] using the concept of path integral, which was originally introduced to explain the quantum mechanics based on the principle of least action in classical mechanics [44]. As shown in Fig. 2b, to formulate the path integral in a finite time horizon \( \tau \), we divide time interval

![Fig. 1](image1.png)  
**Fig. 1** Goal-biased RRT performs well in ‘regular obstacle’ environment (left) but grows badly in complex ‘maze’ scenario (right) with U-shape obstacles. Obstacles are represented in blue circles

![Fig. 2](image2.png)  
**Fig. 2** Illustration of macrostates, microstates and path integral formulation
0 \leq t \leq \tau \) into \( N \) segments with equal width \( \varepsilon \). For every time instant \( t_i \), one specific point \( x_i(t_i) \) will be selected as the representative element of the set that contains all possible points at \( t_i \). We then linearly connect all points \( x_i(t) \) to construct one possible path. As \( \varepsilon \to 0 \), the constructed paths gradually approach to the whole set of all possible paths.

Given the initial state \( x_i(0) \), the path integral formulation for all possible paths \( x_i(t) \) within a finite time horizon \( \tau \) can be interpreted as

\[
\sum_{x_i(t)} \Phi[x_i(t)] = \int_{x_i(t)(0)}^{x_i(t)(\tau)} \Phi[x_i(t)] d\tau
\]

Equation 2

\[
\sum_{x_i(t)} \Phi[x_i(t)] = \int_{x_i(t)(0)}^{x_i(Ne)} \Phi[x_i(t)] d\tau
\]

Equation 3

where \( \Phi[x_i(t)] \) denotes a functional of \( x_i(t) \). As \( x_i(t) \) share the same initial state, we only integrate from \( x_i(t) \) to \( x_i(Ne) \). It is noted that (2) is not well defined as \( \lim_{\varepsilon \to 0} \sum_{x_i(t)} \Phi[x_i(t)] \to \infty \). Although mathematical difficulties exist, above interpretation provides us the physical prospect of path integral formulation. Particularly, if \( \Phi[x_i(t)] \) is defined as the probability of path \( x_i(t) \), the path integral \( \int_{x_i(t)(0)}^{x_i(t)(\tau)} \Phi[x_i(t)] d\tau \) is also the probability considering contribution from all possible paths, given initial state \( x_i(0) \). With the concept of path integral, the causal entropy at macrostate \( X \) in a finite time horizon \( \tau \) is defined as

\[
S_c(X, \tau) = -k_B \int_{x_i(t)} P(x_i(t)|x_i(0)) \ln P(x_i(t)|x_i(0)) d\tau
\]

Equation 4

with computable form

\[
S_c(X, \tau) = -k_B \sum_{x_i(t)} \Omega_i P(x_i(t)|x_i(0)) \ln P(x_i(t)|x_i(0))
\]

Equation 5

where \( k_B \) is the Boltzmann constant connecting macrostates and microstates, \( x_i(0) \) and \( P(\cdot | \cdot) \) denote current state and the conditional probability density, respectively. In (5), the \( i \)th path \( x_i(t) \) is assigned a volume \( \Omega_i \) to represent its proportion in the entire state space [10].

Massive possible paths in state space make it compulsory to sample enough paths such that causal entropy in (5) can be obtained. Suppose we have \( M \) sampled path, by assuming \( \Omega_i \propto 1/P(x_i(t)|x_i(0)) \) and to satisfy the unit measure axiom

\[
\sum_{i=1}^{M} \Omega_i P(x_i(t)|x_i(0)) = 1
\]

Equation 6

we obtain

\[
\Omega_i = \frac{1}{M \cdot P(x_i(t)|x_i(0))}
\]

Equation 7

Then the following proposition on causal entropy can be derived.

**Proposition 1:** The computable causal entropy satisfies

\[
S_c(X_1, \tau) \leq S_c(X_2, \tau)
\]

Equation 8

if either condition (a) or (b) holds.

(a) \( \sum_{i=1}^{M} \Omega_i^1 = \sum_{i=1}^{M} \Omega_i^2 \) and all microstates in \( X_2 \) are equiprobable;

(b) \( \sum_{i=1}^{M} \Omega_i^1 \leq \sum_{i=1}^{M} \Omega_i^2 \) and all microstates in their corresponding macrostate are equiprobable.

where \( \Omega_i^1, \Omega_i^2 \) denotes volume of sample \( x_i(t) \) in \( X_1 \) and \( X_2 \), respectively.

**Proof:** By eliminating \( P(x_i(t)|x_i(0)) \) using (7), we have

\[
S_c(X, \tau) \propto -\frac{1}{M} \sum_{i=1}^{M} \ln \frac{1}{\Omega_i} = -\ln M + \frac{1}{M} \sum_{i=1}^{M} \ln \Omega_i
\]

Equation 9

\[
= \frac{1}{M} \ln \prod_{i=1}^{M} \Omega_i
\]

Equation 10

Since \( \ln (\cdot) \) is monotonically increasing in domain \((0, + \infty)\), \( S_c \) can be compared by analysing \( \prod_{i=1}^{M} \Omega_i \).

(a) Consider the multivariate function \( f(\Omega_1, \ldots, \Omega_M) = \prod_{i=1}^{M} \Omega_i \) with constant constraint \( \sum_{i=1}^{M} \Omega_i = c \). By introducing Lagrangian \( \mathcal{L}(\Omega_1, \ldots, \Omega_M, \lambda) = f(\Omega_1, \ldots, \Omega_M) - \lambda (\sum_{i=1}^{M} \Omega_i - c) \) and solving \((\partial / \partial \Omega_i) \mathcal{L}(\Omega_1, \ldots, \Omega_M, \lambda) = (\partial / \partial \Omega_i) \mathcal{L}(\Omega_1, \ldots, \Omega_M, \lambda) \propto \lambda \) we have the maximum of \( f \) at \( \Omega_i = \cdots = \Omega_M = (1/c) \). In other words, while all microstates in \( X_i \) are equiprobable, \( S_c(X_i, \tau) \leq S_c(X_j, \tau) \). (b) As all microstates in their corresponding macrostate are equiprobable, we know that \( \Omega_i^1 = \cdots = \Omega_i^M \) and \( \Omega_i^2 = \cdots = \Omega_i^M \), which lead to \( \Omega_i^1 \leq \Omega_i^2 \) \( \forall i = 1, \ldots, M \). Thus, we get \( \prod_{i=1}^{M} \Omega_i^1 \leq \prod_{i=1}^{M} \Omega_i^2 \) and further \( S_c(X_i, \tau) \leq S_c(X_i, \tau) \).

Remark 1: Proposition 1 of causal entropy corresponds to two possible sources of entropy reduction: (a) Internal constraints, such as non-holonomic constraints, reduce entropy by breaking uniform distribution of possible paths. (b) External or environmental constraints, which mainly include obstacles, limit state-space volume such that the entropy is also reduced.

### 3.3 Implication of causal entropy maximisation

Causal entropic force has been defined analogously to conventional entropic force to drive states towards entropy maximisation as

\[
F(X, \tau) \propto \nabla_X S_c(X, \tau)_{|X=0}
\]

Equation 11

Distinguished from instantaneous entropy maximisation, causal entropy maximisation focuses on system evolution in a future time horizon. The force towards maximum has been derived and expressed in [10] as

\[
F(X, \tau) = \frac{2T_c}{T_r} \int_{x_i(t)} f(0) P(x_i(t)|x_i(0)) \frac{d}{d\tau} \ln P(x_i(t)|x_i(0)) d\tau
\]

Equation 12

where \( T_r \) and \( T_c \) are temperature parameters regulating force’s magnitude, \( f(0) \) is the zero-mean Gaussian initial force. The ‘causal entropic force’ drives the system to states with more feasible paths in a finite time horizon. This implication inspires us to bias tree searching towards causal entropy maximisation as an explorative direction.

### 4 Maximum entropy-biased rapidly-exploring random tree

The ME-RRT algorithm is outlined in Algorithm 1 (see Fig. 3), where **sampleState** first randomly selects a sample in feasible configuration space. Then the nearest tree node to the sample is selected as the tree growing point. Entropic force is computed with respect to the growing point after path sampling and path integral calculation. Next, **extendTree** tries to extend the tree, given the entropic force, random sample and goal direction. While sampling conditions are satisfied, the algorithm continues extending the tree until the initial state and the goal state are both merged in a single
After obtaining $M$ evolutions of $\mathcal{X}(t)$ as in (13), we now discuss calculation of the conditional probability $\mathbb{P}(\mathcal{X}(t)|x(0))$ for a particular microstate $\mathcal{X}(t)$. By using (7) and noting $f(0)$ is zero-mean, (12) can be approximated as

$$ F(X_0, \tau) \approx -\frac{2T}{I} \sum_{i=1}^{M} \Omega q_i f_i(0) \ln \mathbb{P}(\mathcal{X}(t)|x(0)) $$

$$ = \frac{2T}{I M} \left( \ln M \sum_{i=1}^{M} f_i(0) + \sum_{i=1}^{M} f_i(0) \ln \Omega_i \right) $$

(16)

where $\Omega_i$ can be estimated by first using kernel density estimator (KDE) and then utilising (7). In KDE, a sample path is deemed as a point $p = [x_0(e)^T, \ldots, x_N(e)^T] \in \mathbb{R}^{N}$. Given a set containing $M$ paths $\{p_1, \ldots, p_M\}$, its KDE gives probability density

$$ \mathbb{P}(\mathcal{X}(t)|x(0)) = \frac{1}{M h} \sum_{i=1}^{M} K\left(\frac{p_i - p}{h}\right) $$

(17)

where $h$ is the bandwidth parameter; $K$ is selected as the Epanechnikov (parabolic) kernel function: $K(p - p/h) = \frac{3}{4}(1 - (p - p/h)^2)$ for $|p_i - p/h| \leq 1$, otherwise $K = 0$.

4.3 Hybrid weighted tree extension

The extending function tries to grow the tree towards a particular bias direction. Unlike conventional RRT which extends the tree towards the sample configuration $q_{smp}$, a simple maximum entropy tree is biased towards the entropic force. Given $q_{smp}$, the nearest node on $T$ is selected as $q_0$. Then a naive extending function at $q_0$ can be written as $q_{new} = q_0 + F(X_0, \tau)$, where $X_0$ denotes the macrostate begins at $q_0$, and $F(X_0, \tau)$ denotes the entropic force at $X_0$ with a finite time horizon $\tau$. To overcome the poor adaptability of single sampling strategy in complex environments, a hybrid weighted sampling algorithm is proposed such that the extending function becomes

$$ q_{new} = q_0 + \text{norm} \left( \frac{\rho_1 F(X_0, \tau)}{\|F\|_{\text{max}}} + \frac{\rho_2 (q_{\text{goal}} - q_0)}{\|q_{\text{goal}} - q_0\|} \right. $$

$$ + \left. \frac{\rho_3 (q_{smp} - q_0)}{\|q_{smp} - q_0\|} \right) $$

(18)

where norm normalises the vector to a unit vector; $\|F\|_{\text{max}}$ denotes maximum force obtained in current environment; $\rho_1, \rho_2$, and $\rho_3$ are weighting coefficients. Since $\|F(X_0, \tau)\|$ is varying with respect to $X_0$, another round of normalisation is required to ensure a constant step size of the extension function.

5 Properties of ME-RRT

5.1 Maximum entropy sampling strategy

Proposition 2: Without considering collision force $h$, the discrete form of sampling strategy (13) leads to maximum causal entropy with a specified sampling covariance matrix.
Proof: Without environment influence, the discrete path sampling strategy can be written explicitly as

\[
\begin{bmatrix}
    x \\
    y \\
    \dot{x}_{j+1} \\
    \dot{y}_{j+1}
\end{bmatrix}
= \begin{bmatrix}
    1 & 0 & \epsilon & 0 \\
    0 & 1 & \epsilon & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    \dot{x} \\
    \dot{y}
\end{bmatrix}
+ \begin{bmatrix}
    \frac{\epsilon^2}{2m} & 0 \\
    0 & \frac{\epsilon^2}{2m} \\
    \frac{\epsilon}{m} & 0 \\
    0 & \frac{\epsilon}{m}
\end{bmatrix}
\begin{bmatrix}
    f_i - \lambda i \\
    f_j - \lambda j
\end{bmatrix}
\]

(19)

As \( f = (f_i, f_j) \) is zero-mean Gaussian, we suppose

\[ f_j \sim N(0, \Sigma), \quad \forall j = 0, \ldots, N - 1 \]  

Based on properties of Gaussian distribution, we know that

\[
\begin{bmatrix}
    \dot{x} \\
    \dot{y} \\
    \dot{x}_{j+1} \\
    \dot{y}_{j+1}
\end{bmatrix}
\sim N\left( M_{\dot{x},j}^{(1)} \dot{x}_0, \sum_{k=0}^{j} M_1 M_1^\top \sum_{k=0}^{j} M_2 M_2^\top (M_1^\top)^k \right) 
\]

(21)

\[
\begin{bmatrix}
    x \\
    y \\
    \dot{x}_{j+1} \\
    \dot{y}_{j+1}
\end{bmatrix}
\sim N\left( \sum_{k=0}^{j} M_1 M_1^\top \sum_{k=0}^{j} M_2 M_2^\top (M_1^\top)^k \right) 
\]

(22)

hold for all \( j = 1, \ldots, N - 1 \), where \( M_1 = (1 - (\epsilon \lambda / m)) I \), \( M_2 = (\epsilon / m) I \), \( M_3 = (\epsilon - (\epsilon^2 / 2m)) I \), \( M_4 = (\epsilon^2 / 2m) I \). From (21) and (22), we could further obtain that \( p = [x(\epsilon), \ldots, x(N\epsilon)] \in \mathbb{R}^M \) is also Gaussian.

It is noted that the definition of causal entropy (4) can be deemed as identical to conventional entropy definition for multivariate dimensional random variable \( p \). Given constraints on the covariance matrix, the multivariate Gaussian distribution leads to maximum entropy [46]. Thus, the proposed sampling strategy leads to maximum causal entropy with any pre-defined covariance matrix.

Above proof can be generalised to any dimensional state space without difficulty.

5.2 Degeneracy of ME-RRT

Under some circumstances, ME-RRT could degenerate to goal-biased RRT. According to (18), degeneracy is caused by vanished entropy. In this work, from (16) we know that the degenerate condition is satisfied at local maximum when \( \Omega_j = \Omega_0 \), \( \forall i, j = 1, \ldots, M \). In other words, ME-RRT degenerates when the probability density function is constant for any microstate \( x(t) \). From Proposition 2 we know that entropy force vanishes if there is no interaction (i.e. collision) between sampling paths and obstacles. Thus, the time horizon \( \tau \) needs to be tuned depending on application scenarios such that the environment can be sufficiently explored.

5.3 Time complexity

We discuss extra time complexity of the proposed approach compared with goal-biased RRT. The time consumed by entropy computation would be pathIntegral in Algorithm 1 (Fig. 3). Suppose \( M \) paths with \( N \) discrete steps are required, time complexities for path sampling, KDE and averaging are \( O(MN) \), \( O(NM^2) \) and \( O(M) \), respectively. Thus, the combined extra complexity is \( O(NM^3) \). In practice, a trade-off between algorithm efficiency and approximation accuracy is necessary. From (16), it is obvious that a large \( M \) is desired to obtain an acceptable estimation. Furthermore, a small \( N \) may lead to insufficient exploration of nearby environment and cause degenerated entropic force.

In this work, since the entropic force can be computed given the environment and path sampling strategy, we propose to generate the vector field \( F(X, \tau) \) before searching such that online computation is avoided. With the pre-computed \( F(X, \tau) \), the execution efficiency of the proposed approach is greatly improved.

6 Experiments

In this section, we first explore parameters’ influence on the resulting entropic force field. Then, experiments are conducted in simulated environments to assess the quantitative and qualitative performance of the proposed searching approach. A disc shape holonomic mobile robot with radius 0.5 m is simulated for collision detection. The sampling process follows system dynamics while the planner considers translational states only. The proposed tree extension strategy can be also applied to RRT*, thus results of ME-RRT* is listed in experiments.

6.1 Experimental configurations

Results from the proposed approach have been compared with well-known planners including [Code for environment design and conventional 2-D RRT is obtained from http://coecsl.ece.illinois.edu/ge423/spring13/RickRekoskeAvoid/rrt.html.] RRT/RRT* and goal-biased RRT/RRT*. All experiments in 2D scenarios are carried out with MATLAB R2017b on a mobile workstation with Intel Core i5-8250U CPU and 8 GB RAM. All experiments in 3D scenarios are carried out with MATLAB R2017b on a workstation with Intel Core i7-4770 CPU and 15.6 GB RAM.

To avoid online computation of entropic force, entropic force is pre-computed at uniformly sampled point set \( S \) with resolution 2.5 m × 2.5 m. Given any configuration \( q_0, F_{X_0, \tau} \) is approximated as \( F_{X_0, \tau} \) which satisfies \( \|q_0 - q_0^*\| \leq \|q_0^* - q_0\| \), \( \forall q_0^* \in S \).

Parameters in experiments are \( \kappa = 1.38 \times 10^{-23} \), \( T_r = 0.5 \times 10^{-22} \), \( T_r = 500T_r, M = 5000, \epsilon = 0.1, m = 0.1, \lambda = 0.1 \) and others in Table 1. The step size is set to 2. The radius of searching percentage for optimal planner is 5. Experiments are repeated for 1000 times and 500 times for 2D and 3D scenarios, respectively, under each configuration and only successful paths are taken during evaluation to obtain statistically reasonable results.

6.2 Influence of parameters on entropic force field

Since the time horizon \( \tau \) is an essential parameter that embeds in path integral calculation without explicit expression, we investigate the influence of \( \tau \), as shown in Fig. 4. It is noted that \( \tau \) is positively related to ‘detection range’ to obstacles. For a small \( \tau \), only local environment can be explored such that the force vanishes in the

| Table 1 Experimental parameters
| Method | \( \rho_1 \) | \( \rho_2 \) | \( \rho_3 \) | \( \tau \) |
|----------|----------------|----------------|----------------|---------------|
| ME-RRT/ME-RRT* | 2 | 1 | 3 | 6 |
| RRT/RRT* | 0 | 0 | 3 | — |
| goal-biased RRT/goal-biased RRT* | 0 | 1 | 3 | — |
| Scenario 1: regular obstacles |
| ME-RRT/ME-RRT* | 2 | 2 | 3 | 100 |
| RRT/RRT* | 0 | 0 | 3 | — |
| goal-biased RRT/goal-biased RRT* | 0 | 2 | 3 | — |
| Scenario 2: narrow passage |
| ME-RRT/ME-RRT* | 2 | 0 | 3 | 10 |
| RRT/RRT* | 0 | 0 | 3 | — |
| goal-biased RRT/goal-biased RRT* | 0 | 1 | 3 | — |
Fig. 4 Influence of varying time horizon $\tau$ in scenario 2: narrow passage. Entropic force (vector) fields are illustrated using black line segments dotted at each sampled point, while the corresponding volume (scalar) fields are demonstrated with triangular surface plots. Numbers in colour bars denote total volume $\sum_{t=1}^{T} \Omega_{t}$ for a particular sample point. Sample size is set to $M = 5000$ for Figs. 4a-c and $M = 6000$ for Fig. 4d.

Table 2 Quantitative results for 2D scenarios

| Planner      | Tree size | Iteration no. | Execution time | Success rate |
|--------------|-----------|---------------|----------------|--------------|

Table 3 Quantitative results for 3D scenarios

| Planner      | Tree size | Iteration no. | Execution time | Success rate |
|--------------|-----------|---------------|----------------|--------------|

central zone which is far from obstacles. Readers may also notice that points are pushed away from the corridor because the time horizon is too small such that paths passing through the corridor are unlikely to be generated. Setting $\tau$ to a large value will increase the global vision by producing paths that fully explore the environment, but may also lead to little difference in volumes for two near points.

Thus, a favourable entropic force field can be obtained by selecting an appropriate $\tau$ according to environment size and desired exploration range.
6.3 Comparative results

We have designed several testing scenarios in 2D and 3D spaces including (i) regular obstacles, (ii) narrow passage and (iii) maze. The experimental results for the proposed approach and benchmark planners in 2D and 3D spaces are listed in Tables 2 and 3, respectively, where the best results are highlighted in bold. Evaluation criteria include the number of nodes in generated trees, the total iteration number during searching, the execution time, the success rate within the maximum iteration numbers and the node-iteration ratio (NIR) which measures extended node number in a unit iteration. Note that 2D and 3D experiments are conducted on different platforms, thus execution time between Tables 2 and 3 are not comparable.

As for quantitative results in 2D scenarios, the proposed approach results in smaller tree size in half of the testing scenarios and always leads to the smallest iteration number. The larger NIR shows that ME-RRT is more efficient in tree extension. Although the extension function is more complex than conventional approaches, the proposed searching direction saves more time such that ME-RRT performs best in execution time, except for experiments (i) and (v), where the success rates of ME-RRT are higher than its competitors. The superiority of ME-RRT is particularly evident for complex scenarios such as in experiment (v) – where goal-biased RRT performs even worse than RRT because of its naive tree extension strategy. The 3D results are similar to 2D results. As we simply elongate 2D obstacles in z-axis to build 3D obstacles, the theoretical entropy force in z-axis is zero, which makes tree extension only be guided in xy-plane. That may explain why sometimes RRT* outperforms ME-RRT* in NIR.

The qualitative results in 3D scenarios are illustrated in Fig. 5. The proposed extension function helps to explore the environment efficiently with easier obstacle bypassing and avoidance, which can be seen in scenario 2 with less trees generated before entering the narrow passage compared with RRT* and goal-biased RRT*.

Fig. 5 Three 3D scenarios and example paths planned with ME-RRT* (upper), RRT* (middle) and goal-biased RRT* (lower). Each figure shows successfully planned paths (in red) in ten trials, while trees generated are coloured green. Start states and goal states are represented by green and red circles, respectively.

a Scenario 1: regular obstacles, ME-RRT*
b Scenario 2: narrow passage, ME-RRT*
c Scenario 3: maze, ME-RRT*
d Scenario 1: regular obstacles, RRT*
e Scenario 2: narrow passage, RRT*
f Scenario 3: maze, RRT*
g Scenario 1: regular obstacles, goal-biased RRT*
h Scenario 2: narrow passage, goal-biased RRT*
i Scenario 3: maze, goal-biased RRT*
7 Conclusion

This paper presents a path searching approach where the ME-RRT is proposed to guide the tree expansion. Environment information has been considered in ME-RRT without explicit mapping but by random path sampling. We have evaluated the performance of ME-RRT by comparing with RRT, goal-biased RRT and their optimal descendants in three different 2D/3D scenarios. The results show that ME-RRT is more efficient in tree generation and has a significantly higher success rate in complex scenarios.

As a new heuristics, maximum entropy has shown its potential by providing a search direction to maximise the number of optimal descendants in three different 2D/3D scenarios. Although simulated experiments have been conducted to verify the effectiveness of the maximum entropy-biased sampling under certain conditions, conventional exploration and searching approaches cannot be fully replaced but need to be modified carefully to balance among optimality, efficiency and adaptability in different scenarios. Moreover, how to select parameters in entropy maximisation is another issue worthy of study. In this work, the vector field representing entropic force has been generated in advance before searching. Online path sampling using parallel acceleration is expected in the future. The online computation will expand the scope of the proposed approach to search in dynamic, large-scale or higher-dimensional environments.

8 References

[1] Wang, M., Shan, H., Lu, R., et al.: ‘Real-time path planning based on hybrid-vanet-enhanced transportation system’, IEEE Trans. Veh. Technol., 2015, 64, (5), pp. 1664–1678
[2] Tanner, H.G., Loizou, S.G., Kyriakopoulos, K.J.: ‘Nonholonomic navigation and control of cooperating mobile manipulators’, IEEE Trans. Robot. Autom., 2003, 19, (1), pp. 16–28
[3] Choset, H.M.: ‘Principles of robot motion: theory, algorithms, and implementation’ (MIT Press, USA, 2005)
[4] Lit, X.-C., Ge, S.S., Al Mamun, A.: ‘Hierarchical incremental path planning and situation-dependent optimized dynamic motion planning considering accelerations’, IEEE Trans Syst Man Cybern. B (Cybern.), 2007, 37, (6), pp. 1541–1554
[5] Jan, G.E., Chang, K.Y., Parberry, I.: ‘Optimal path planning for mobile robot navigation’, IEEE/ASME Trans. Mechatronics, 2008, 13, (4), pp. 451–460
[6] Soulignac, M.: ‘Feasible and optimal path planning in strong current fields’, IEEE Trans. Robot., 2011, 27, (1), pp. 89–98
[7] Bowen, C., Ye, G., Alterovitz, R.: ‘Asymptotically optimal motion planning for learned tasks using time-dependent cost maps’, IEEE Trans. Autom. Sci. Eng., 2015, 12, (1), pp. 171–186
[8] Shum, A., Morris, K., Khajepour, A.: ‘Robots and Systems (IROS), Seattle, WA, USA, 2015, pp. 3227–3234
[9] Lai, X.-C., Ge, S.S., Al Mamun, A.: ‘Robots and Systems (IROS), San Francisco, CA, USA, 2000, vol. 2, pp. 995–1001
[10] Perez, A., Karaman, S., Shkolnik, A., et al.: ‘Asymptotically-optimal path planning for manipulation using incremental sampled skeletal tree’, IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), San Francisco, CA, USA, 2011, pp. 4307–4313
[11] Salzman, O., Halperin, D.: ‘Asymptotically near-optimal RRT for fast, high-quality motion planning’, IEEE Trans. Robot., 2016, 32, (3), pp. 473–483
[12] Palomar, L., Arras, K.O.: ‘A novel RRT extend function for efficient and smooth mobile robot motion planning’. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), Korea, 2014, pp. 434–441
[13] Moon, C.-B., Chung, W.: ‘Kinodynamic planner dual-tree RRT (DT-RRT) for two-wheeled mobile robots using the rapidly exploring random tree’, IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), Daejeon, Korea, 2016, pp. 3714–3721
[14] Ge, S.S., Cui, Y.J.: ‘New potential functions for mobile robot path planning’, IEEE Trans. Robot. Autom., 2000, 16, (5), pp. 615–620
[15] Bajpai, B.: ‘Local motion planning based on shrinking and growing geometry models’. Proc. 1996 IEEE Int. Conf. Robotics and Automation, Minneapolis, Minnesota, USA, 1996, vol. 4, pp. 3303–3308
[16] Wang, M., Shan, H., Lu, R.: ‘MD-RRT: A new random path sampling using hardware-accelerated computation of generalized voronoi diagrams’. Proc. 2000 IEEE Int. Conf. Robotics and Automation, San Francisco, CA, USA, 2000, vol. 3, pp. 2931–2937
[17] Liu, G., Lien, J.-M.: ‘Fast global-to-local path planning using cross-entropy’. Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), Hamburg, Germany, 2015, pp. 3262–3267
[18] Sun, Z., Hsu, D., Jiang, T., et al.: ‘Narrow passage sampling for probabilistic roadmap planning’, IEEE Trans. Robot., 2005, 21, (6), pp. 1105–1115
[19] Du, M., Chen, J., Zhao, P., et al.: ‘An improved RRT-based motion planner for autonomous vehicle in cluttered environments’. IEEE Int. Conf. Robotics and Automation (ICRA), Hong Kong, China, 2014, pp. 4674–4679
[20] Branicky, M.S., LaValle, S.M., Olson, K., et al.: ‘Quasi-randomized path planning’. Proc. 2001 IEEE Int. Conf. Robotics and Automation, Seoul, Korea, 2001, vol. 2, pp. 1481–1487
[21] Urmon, C., Simmons, R.: ‘Approaches for heuristically biasing RRT growth’. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 2013, vol. 2, pp. 1176–1182
[22] Palmieri, L., Koening, S., Arras, K.O.: ‘RRT*-based nonholonomic motion planning using any-angle path biasing’. IEEE Int. Conf. Robotics and Automation (ICRA), Stockholm, Sweden, 2016, pp. 2775–2781
[23] Kolarov, M.: ‘Cross-entropy motion planning’. Int. J. Rob. Res., 2012, 31, (7), pp. 855–871
[24] Karaman, S., Frazzoli, E.: ‘Sampling-based algorithms for optimal motion planning’. Int. J. Rob. Res., 2011, 30, (7), pp. 846–868
[25] Webb, D.J., van den Berg, J.: ‘Kinodynamic RRT*: asymptotically optimal motion planning for robots with linear dynamics’. IEEE Int. Conf. Robotics and Automation (ICRA), Hong Kong, China, 2014, pp. 4674–4679
[26] Gammell, J.D., Srinivasa, S.S., Barfoot, T.D.: ‘Informed RRT*: optimal sampling-based path planning focused via direct sampling of an admissible ellipsoidal potential field’. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), Chicago, Illinois, USA, 2014, pp. 2970–2977
[27] Devaurs, D., Siméon, T., Cortés, J.: ‘Role of the learning function in extremum seeking control’. Proc. 2016 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), Singapore, 2017, pp. 4240–4247
[28] Noren, I., Khan, A., Habib, Z.: ‘Optimal path planning using RRT* based approaches: a survey and future directions’. Int. J. Adv. Comput. Sci. Appl., 2017, 8, pp. 247–255
[29] Pfeiffer, M., Schaubeil, M., Nieto, J., et al.: ‘From perception to decision: A data-driven approach to end-to-end motion planning for autonomous ground robots’. IEEE/RSJ Int. Conf. Robotics and Automation (ICRA), Singapore, 2017, pp. 5127–5133
[30] Finn, C., Levine, S.: ‘Deep visual foresight for planning robot motion’. IEEE Int. Conf. Robotics and Automation (ICRA), Singapore, 2017, pp. 2786–2793
[31] Li, M., Jiang, R., Ge, S.S., et al.: ‘Role of the learning function in extremum seeking control of a heterogeneous mobile robot navigation’, CAAI Trans. Intell. Technol., 2018, 3, (1), pp. 49–58
[32] Wu, Q., Miao, C.: ‘Curiosity: from psychology to computation’, ACM Comput. Surv. (CSUR). 2013, 46, (2), pp. 1–18
[33] Hester, T., Stone, P.: ‘Intrinsically motivated model learning for developing curious robots’. Artif. Intell., 2017, 247, pp. 170–186
[34] Laversanne-Finot, A., Pérè, A., Oudeyer, P.-Y.: ‘Curiosity driven exploration of learned disentangled goal spaces’, arXiv preprint arXiv:1807.01521, 2018
[35] Feynman, R.P., Hibbs, A.R., Styer, D.F.: ‘Quantum mechanics and path integrals’ (Courier Corporation, USA, 2005)
[36] Barraggard, J., Latombe, J.C.: ‘A monte-carlo algorithm for path planning with many degrees of freedom’. Proc. 1990 IEEE Int. Conf. Robotics and Automation, Cincinnati, Ohio, USA, 1990, pp. 1712–1717
[37] Chirikjian, G.S.: ‘Gaussian distributions and the heat equation’, in ‘Stochastic models, information theory, and lie groups’ (Birkhäuser, Switzerland, 2009), vol. 1, pp. 31–61