J Intell Robot Syst (2014) 75:291–311
DOI 10.1007/s10846-013-9981-9

Active SLAM and Exploration with Particle Filters Using Kullback-Leibler Divergence

Luca Carlone · Jingjing Du · Miguel Kaouk Ng · Basilio Bona · Marina Indri

Received: 26 March 2013 / Accepted: 26 September 2013 / Published online: 13 October 2013

Abstract Autonomous exploration under uncertain robot location requires the robot to use active strategies to trade-off between the contrasting tasks of exploring the unknown scenario and satisfying given constraints on the admissible uncertainty in map estimation. The corresponding problem, namely *active SLAM (Simultaneous Localization and Mapping) and exploration*, has received a large attention from the robotic community for its relevance in mobile robotics applications. In this work we tackle the problem of active SLAM and exploration with Rao-Blackwellized Particle Filters. We propose an application of Kullback-Leibler divergence for the purpose of evaluating the particle-based SLAM posterior approximation. This metric is then applied in the definition of the *expected information from a policy*, which allows the robot to autonomously decide between exploration and place revisiting actions (i.e., *loop closing*). Extensive tests are performed in typical indoor and office environments and on well-known benchmarking scenarios belonging to SLAM literature, with the purpose of comparing the proposed approach with the state-of-the-art techniques and to evaluate the maturity of truly autonomous navigation systems based on particle filtering.

Keywords Mobile robot · Active SLAM · Autonomous exploration · Rao-Blackwellized Particle Filters

L. Carlone (✉) · J. Du · M. Kaouk Ng · B. Bona · M. Indri
Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy
e-mail: luca.carlone@polito.it

J. Du
e-mail: hmily1119@gmail.com

M. Kaouk Ng
e-mail: meng400@hotmail.com

B. Bona
e-mail: basilio.bona@polito.it

M. Indri
e-mail: marina.indri@polito.it

1 Introduction

In the last decades, the increasing use of mobile robotic systems in service and domestic applications has highlighted the importance of autonomous navigation. In particular, for a mobile robot, performing rescue or coverage tasks in emergency or disaster situations, *autonomous exploration* capability is a crucial prerequisite for mission accomplishment. When no absolute localization service is available to the robot (e.g., GPS), the exploration task, which can be considered a decisional process, needs to be accompanied by

the concurrent estimation of robot pose and a map of the environment. This estimation process is usually referred to as *Simultaneous Localization and Mapping* (SLAM). While the maturity of SLAM in both single robot and multi robot scenarios has been recognized by the robotic community, see [1, 2], the problem of active planning under uncertainty with application to robotic exploration still remains an open issue. The corresponding problem, also referred to as *active SLAM and exploration*, requires the robot to plan its motion in order to maximize the explored areas and at the same time minimize the uncertainty in SLAM. The former objective is clearly accomplished by visiting unknown places, while the latter requires the robot to perform *loop closing* actions, i.e. to come back to already visited areas.

While the importance of the exploration task is quite intuitive, the need of imposing suitable constraints on the admissible uncertainty stems from two different reasons: on one hand, such constraints can be due to mission requirements (e.g., high uncertainty in the map model can be undesirable in a rescue task, if human intervention is planned on the basis of such world model); on the other hand, *the need of imposing bounds on the admissible uncertainty stems from reasons that are intrinsic to the estimation process*. Probabilistic approaches to the SLAM problem are usually based on some assumptions on the underlying probability densities, which allow to solve the problem with a sustainable computational effort. In *Extended Kalman Filter* (EKF) SLAM and variants the main hypothesis consists in assuming the posterior to be normally distributed. In order to preserve the Gaussianity of the posterior, the nonlinear models describing robot motion and the available measurements need to be linearized. The linear model properly describes the behavior of the nonlinear system in a neighborhood of the linearization point; for large amount of uncertainty, the linear approximation may not correctly model the problem and the effect of such phenomenon is that the actual probability density of the SLAM posterior can no longer be approximated with a Gaussian density. This is not a mere approximation issue: if the posterior description is not consistent with the actual estimation errors, the filter is likely to diverge leading to failure of the exploration (and mapping) task [3]. A second widespread category of approaches to SLAM is based on Rao-Blackwellized Particle Filters (RBPF). In this paradigm, the Rao-Blackwell factorization is applied to decouple the localization and mapping problems and the pose estimation is entrusted to a particle filter, see [4, 5]. The assumption underneath the RBPF algorithm is that the posterior density can be properly approximated by a point-mass (particle) representation. Since the complexity of the approach grows with the particle set size, common processors can only handle few hundreds particles. Therefore, a limited number of particles can be not adequate for assuring a proper approximation of the trajectory posterior, and this issue becomes critical when the uncertainty in the posterior increases [6]. Because of the factorization, a bad approximation of the trajectory posterior affects the estimated map quality, hence, a constraint on the uncertainty of the trajectory posterior is desirable to accomplish consistent mapping.

At this point, it should be clear that the active SLAM task requires suitable metrics for evaluating the uncertainty, so that the corresponding uncertainty bounds can be assessed. When dealing with EKF-based approaches, this metric can be naturally found in the entropy of the multivariate belief describing both map and robot pose. Although the evaluation of the *expected gain* of an action can be computationally demanding, effective and well founded approaches exist, see for instance [7–9], and the references therein. In [10], the exploration problem is solved by applying a Bayesian optimization method that allows to reduce the number of simulations needed for planning. More recently, a very lucid treatment of EKF-based active SLAM is provided in [11], which also discusses the use of different uncertainty metrics. The works [12, 13] instead discuss the active SLAM problem when using pose graph optimization to estimate robot trajectory. The problem of active SLAM and exploration with particle filters, instead, is not fully understood and recent literature on the topic remarks several drawbacks of naive entropy-based metrics [14].

In this work, which extends and completes the results obtained in [15, 16], and [17], we investigate the problem of active SLAM and exploration

with Rao-Blackwellized Particle Filters. We first derive a measure of RBPF uncertainty using the Kullback-Leibler divergence [18]. The divergence allows to obtain an upper bound on the error committed when the true posterior is approximated with a particle-based representation. This metric is then applied in the definition of the *expected information from a policy*; the proposed information measure quantifies the advantage in reaching a target point, taking into account both the possibility of exploring new areas and the necessity of bounding the admissible uncertainty. It is worth noticing that the term "expected" is used in the probabilistic sense and not only to resemble the predictive nature of the metric. Finally we propose an extensive evaluation of state-of-the-art approaches to active SLAM and exploration with particle filters, applied to the estimation of occupancy grid map representations of an indoor environment. The experimental analysis provides several insights on related approaches and highlights the advantage in using the proposed technique. For instance, we show that, depending on the design parameters, joint entropy-based techniques [19] can only trade-off between two suboptimal strategies: maximizing the number of visited cells, or moving towards the nearest exploration target. In both cases, filter uncertainty plays a minor role in planning, hence producing strategies that are not robust in presence of large estimation uncertainty. Other metrics (e.g., expected map information [14]) can better capture the estimation uncertainty, but at the price of extra-computation; also these techniques, however, cannot guarantee the robustness required in real world operation and are prone to failure. The proposed technique combines computational efficiency with robustness. In particular, the complexity is linear in the number of particles in the filter. Moreover, we explicitly model the probability of estimating a correct map during exploration using Kullback-Leibler divergence (KLD). The KLD-based metric essentially quantifies the probability/risk of failure when reaching a given target, and enables to design uncertainty-aware exploration strategies.

The paper is organized as follows. Related works on active SLAM and exploration are reviewed in Section 2. After some preliminary notions, Kullback-Leibler divergence is applied for the evaluation of SLAM posterior uncertainty in Section 3. Section 4 discusses how to use this metric in the context of active SLAM and exploration and introduces the *expected information from a policy*. Experiments are reported in Section 5, while conclusions are drawn in Section 6.

2 Related Work

This section discusses the state-of-the-art of robotic exploration using occupancy grid maps [20]. Literature on robotic exploration can be classified in two research lines: exploration with perfect position knowledge and exploration under uncertain localization. In the former scenario, accurate prior knowledge of robot position is available, and the robot decision making reduces to the choice of the exploration target that maximizes the opportunity of visiting unknown areas. Early contributions to this setup can be found in [21], in which *frontiers*, that is, cells on the boundary between known areas and unexplored space, are selected and used as potential motion targets for exploration. In these approaches the robot simply chooses the nearest target, regardless complex decision making strategies. Similar approaches can be found in more recent contributions, see for instance [22]. Also in the case of perfect knowledge of robot pose, the onboard sensors have limited range and are affected by uncertainty, hence suitable metrics are required to evaluate the so called *information gain*, i.e., the expected amount of information that can be collected when reaching a motion target. In order to model the uncertainty of the occupancy grid map representation of the environment, Bourgault et al. [7] and Moorehead et al. [23] proposed the use of the entropy of the cells in the grid. In such a case the gain, computed from the entropy, quantifies the expected amount of information that can be acquired by taking a sensor reading in a certain area of the map.

The approaches mentioned so far do not take into account pose uncertainty when selecting the next target point. When dealing with the general problem of exploration under uncertain robot position, several challenges arise. Stachniss et al. addressed the problem of active SLAM and exploration using Rao-Blackwellized Particle

Filters, improving exploration performances via active loop closing, through a heuristic method for re-traversing loops [24]. A decision-theoretic approach was proposed in [19] in which active loop closing is achieved by monitoring the uncertainty in RBPF. Many drawbacks of such method are analyzed in [14], where the authors also proposed a measure of RBPF-SLAM uncertainty. Other approaches to the problem of measuring uncertainty in SLAM exist in literature, involving the computation of volume covered by particles or non-entropy-based utility functions applied to exploration. For further discussion and models the reader is referred to [14] and [25]. Formal definitions of the information gain in related work are also discussed in Section 5.1.

3 Rao-Blackwellized Particle Filters SLAM

3.1 Preliminaries

Although the high dimensionality of state space in grid-based SLAM makes challenging the application of sample-based representations of the posterior of robot pose and occupancy grid map, an elegant solution to reduce dimensionality of the sampling space can be obtained through Rao-Blackwellization [4]. Since the map probability can be computed analytically given robot path, it is convenient to factorize the joint probability as follows:

$$p(x_{1:t}, m \mid z_{1:t}, u_{0:t-1}) = p(m \mid x_{1:t}, z_{1:t}) \cdot p(x_{1:t} \mid z_{1:t}, u_{0:t-1}) \quad (1)$$

In Eq. 1 the state includes robot trajectory $x_{1:t} = \{x_1, x_2, \ldots, x_t\}$ and the map m, both estimated from measurements $z_{1:t} = \{z_1, z_2, \ldots, z_t\}$ and odometry $u_{0:t-1} = \{u_0, u_1, \ldots, u_{t-1}\}$. In the following, we alternatively use the notation $d_{1:t} = \{z_{1:t}, u_{0:t-1}\}$ to denote all the available data until time t. Equation 1 provides the basis for RBPF-SLAM: the particle filter is applied to the problem of estimating potential trajectories of the robot and a map is associated to each trajectory hypothesis (also called *sample* or *particle*). According to particle filter framework, the posterior of robot trajectory is approximated by a set of weighted random samples:

$$p(x_{1:t} \mid z_{1:t}, u_{0:t-1}) \approx \sum_{i=1}^{n} \omega_t^{[i]} \delta\left(x_{1:t} - x_{1:t}^{[i]}\right) \quad (2)$$

where n is the particle set size, $x_{1:t}^{[i]}$ is the trajectory (i.e., the collection of the poses assumed from time 1 to time t) of the i-th particle, $\omega_t^{[i]}$ is the corresponding weight $\left(\sum_{i=1}^{n} \omega_t^{[i]} = 1\right)$, and $\delta(\cdot)$ is the Dirac delta function. Filter prediction is obtained by drawing particles from the *proposal distribution* $\rho(x_{t+1} \mid x_t, u_t)$, which is often approximated with a Gaussian density, whose mean and covariance depend on the odometric information u_t, whereas the weights are updated according to [26]:

$$\omega_t^{[i]} = \omega_{t-1}^{[i]} p\left(z_t \mid x_{t-1}^{[i]}, u_{t-1}, m_{t-1}^{[i]}\right), \quad i = 1, \ldots, n \quad (3)$$

in which the likelihood of current observation is included in the importance weight of a sample. Particles degeneracy (i.e., the situation in which most part of the sample set has negligible weight) is then prevented by a resampling phase that randomly chooses the samples that best fit current and past observations, according to particles weights. A common condition for resampling is based on the *effective sample size* [27], which is an approximated measure of particle diversity:

$$\tilde{N}_{\text{eff}} = \frac{1}{\sum_{i=1}^{n} \left(\omega_t^{[i]}\right)^2} \quad (4)$$

Particles are re-sampled if the previous quantity drops below a given threshold, usually fixed to $n/2$ [19].

3.2 Reverse KLD-Sampling for Measuring the Uncertainty in RBPF-SLAM

This section introduces the first contribution of the article. When the uncertainty in robot localization increases (e.g., the robot is traversing a long trajectory in open loop, without revisiting known places), the particle-based belief approximation degrades and loop-closing occurrence becomes

crucial to reduce the uncertainty. For the Rao-Blackwell factorization, a poor approximation of the trajectory posterior causes a degradation of the quality of the estimated map, hence performing simple exploration (i.e., choosing the motion targets so to maximize the explored areas) may lead to inconsistent map models, making the exploration process inefficient.

The problem of quantifying the quality of the particle-based approximation was firstly addressed by Fox [28] with application to Monte Carlo Localization (MCL); Fox applied the Kullback-Leibler divergence (KLD) $\xi(p_1, p_2)$, which is a measure of fit between two probability distributions p_1, p_2, to derive the number of particles needed to achieve a desired approximation error with given probability (Eq. 5 in [28]):

$$n \doteq \frac{k_t - 1}{2\bar{\xi}_t}\left[1 - \frac{2}{9(k_t-1)} + \sqrt{\frac{2}{9(k_t-1)}}\zeta_{1-\gamma}\right]^3 \tag{5}$$

The previous relation describes the number of samples n that are needed to assure that, with probability $1-\gamma$, the error between the particle-based belief and the true probability density (assumed to be piecewise constant, with support on k_t *bins*) is smaller than $\bar{\xi}_t$; the term $\zeta_{1-\gamma}$ is the upper $1-\gamma$ quantile of the standard normal distribution. Roughly speaking, the state space is discretized in bins and k_t (number of bins being support of the piecewise constant distribution) is the number of bins in which at least one particle falls.

Conversely, when considering a fixed particle set size n, it is possible to calculate the bound $\bar{\xi}_t$ on the divergence between the true posterior $p(x_t \mid d_{1:t})$, describing robot pose at time t, and its point mass approximation $\hat{p}(x_t \mid d_{1:t})$, i.e.,

$$\xi(\hat{p}(x_t \mid d_{1:t}), p(x_t \mid d_{1:t})) < \bar{\xi}_t, \quad \text{with}$$
$$\bar{\xi}_t = \frac{k_t - 1}{2n}\left[1 - \frac{2}{9(k_t-1)} + \sqrt{\frac{2}{9(k_t-1)}}\zeta_{1-\gamma}\right]^3 \tag{6}$$

where $\bar{\xi}_t$ is computed from Eq. 5. For the sake of simplicity, the divergence will be lately written as $\xi(p(x_t \mid d_{1:t}))$, omitting the dependence on the two compared distributions. According to Eq. 6, for each pose $x_\tau, \tau = 1, 2, \ldots, t$, an upper bound $\bar{\xi}_\tau$ on the approximation error can be defined with a given probability $1-\gamma$. Let the error of approximation of the whole trajectory posterior be defined as:

$$\xi(p(x_{1:t}, d_{1:t})) = \xi(p(x_t^*, d_{1:t})),$$
$$x_t^* = \text{argmax}_{\tau\in[1:t]}(\bar{\xi}_\tau) \tag{7}$$

i.e., the error of the pose that is expected to provide the worst approximation is assigned to $\xi(p(x_{1:t}, d_{1:t}))$. Equation 7 formalizes the observation that the trajectory estimation, thus the map quality, can be compromised if a single pose is affected by large errors. Combining Eqs. 6 and 7, and since $\bar{\xi}_\tau$ is monotonically increasing in the number of non-empty bins, the probability of having a trajectory error less than $\bar{\xi}$ can be derived as:

$$p(\xi(p(x_{1:t}, d_{1:t})) < \bar{\xi}) = p(\xi(p(x_t^*, d_{1:t})) < \bar{\xi})$$
$$= \mathcal{F}\left\{\sqrt{\frac{9(k_t^*-1)}{2}}\left(\sqrt[3]{\frac{2n\bar{\xi}}{k_t^*-1}} + \frac{2}{9(k_t^*-1)} - 1\right)\right\} \tag{8}$$

where $k_t^* = \max\{k_\tau, \tau \in [1:t]\}$ and $\mathcal{F}(\cdot)$ is the cumulative distribution function of a standard Gaussian density. The evaluation of Eq. 8, at each time step t, only requires the computation of k_t, since k_t^* can be computed iteratively as $k_t^* = \max\{k_t, k_{t-1}^*\}$. Furthermore, the computation of k_t has complexity $O(n)$, since for each sample only the corresponding support bin has to be evaluated.

Remark 1 Equation 8 returns the probability that the Kullback-Leibler divergence $\xi(p(x_{1:t}, d_{1:t}))$ between the true posterior and our particle-based approximation is smaller than a given threshold $\bar{\xi}$. Roughly speaking, $p(\xi(p(x_{1:t}, d_{1:t})) < \bar{\xi})$ is the probability that the quality of our particle-based belief is sufficiently good. Conversely, $1 - p(\xi(p(x_{1:t}, d_{1:t})) < \bar{\xi})$ represents the probability (or the *risk*) of failure, i.e., the probability of having a bad particle-based belief representation. The probability $p(\xi(p(x_{1:t}, d_{1:t})) < \bar{\xi})$ is function of the number of samples n and the number of support bins k_t^*. In particular, it is increasing in

n, meaning that the quality of our particle-based representation improves when adding particles in the filter. More interestingly, the probability in Eq. 8 is decreasing in k_t^* (the red curve in Fig. 3 shows the trend of the probability in Eq. 8 as a function of k_t^*). Intuitively, the number of bins k_t^* occupied by the particles measures the spread of the particles in the state space: if the particles are very close to each other, they are likely to occupy few bins (hence k_t^* is small, and the probability (8) is close to 1); in this case the particle-based belief is a very good approximation or the true posterior. Conversely, when the particles are spread across the state space, they occupy several bins, leading to a large k_t^* and to a small probability (8). In a typical exploration scenario, when the robot starts performing SLAM, all particles are initialized at the same pose (zero uncertainty) and the probability (8) is 1. When the robot travels in open loop, visiting unknown places, the particle spread across the state space, to model possible trajectory hypotheses. Therefore, k_t^* increases and the probability (8) decreases. After a loop closing, few particles are re-sampled by the particle filter, re-establishing a small value for k_t^* (the re-sampled particles usually cover few bins), and a probability (8) close to 1. For this reason, the probability $p(\xi(p(x_{1:t}, d_{1:t})) < \bar{\xi})$ given in Eq. 8 is also a good metric to detect loop closings, as remarked in Section 5.4 and in our previous work [16].

3.3 Parameters Setting

Equation 8 describes the probability of having a divergence smaller than $\bar{\xi}$. The goal is to devise a suitable selection of model parameters, such that the threshold $\bar{\xi}$ has a meaningful link with the quality of the estimated map. Roughly speaking, the choice of the model's parameters (including the threshold $\bar{\xi}$) must be done so that a divergence smaller than $\bar{\xi}$ corresponds to a consistent map estimate, whereas a divergence larger than $\bar{\xi}$ corresponds to "bad" map estimates. In this way, $p(\xi(\hat{p}(x_{1:t}, d_{1:t})) < \bar{\xi})$ can be used as the probability of estimating a consistent map.

For a given particle set size, Eq. 8 depends on the number of non-empty bins, k_t, and the error bound $\bar{\xi}$. The former is connected with the bin size selected for state space discretization (recall the assumption that the true distribution is piecewise constant).

The variable k_t ranges from 1 to n and the bin size has to be designed to obtain an outcome of the model as informative as possible: extremely small grid sizes lead to fast saturation of k_t (particles cover all possible bins in few iterations of the filter), whereas large grid sizes do not give a fine granularity in uncertainty evaluation (particles always cover few bins, hence only small values of k_t occur). These considerations can be easily understood from Fig. 1, which reports the trend of k_t for the case in which the robot travels a long loop, revisiting the starting position at approximately time step 1250. In order to formalize the bin size selection, the concept of *bin efficiency* (η) is introduced as:

$$\eta = \left(\frac{t_r}{T}\right)\left(\frac{\max_{t\in[1,\ldots,T]} k_t}{n}\right) \tag{9}$$

where t_r is the rise time, defined as the number of time steps required for k_t to go from 10 % to 90 % of the dynamics, and T is a normalization factor corresponding to the total duration of the experiment. The first term in Eq. 9 penalizes small bin sizes, for which k_t quickly saturates to its upper bound, instead the second one penalizes large bins, which prevent the use of the whole dynamics. Figure 2 reports the efficiency value η for different selection of the bin size. The bin size is described

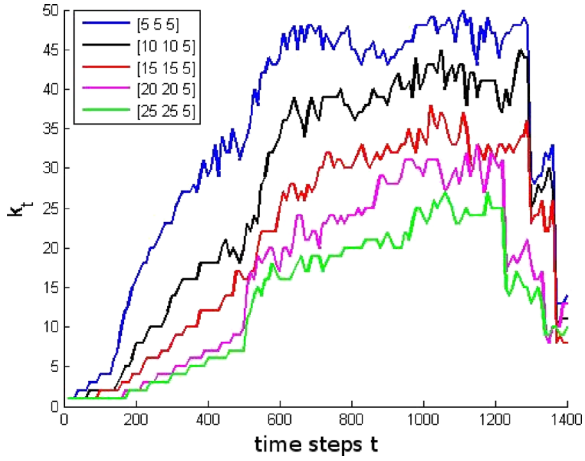


Fig. 1 Number k_t of non-empty bins for different bin sizes $[\Delta x\ \Delta y\ \Delta\theta]$ using 50 particles. Δx and Δy are expressed in centimeters, whereas $\Delta\theta$ is expressed in degrees

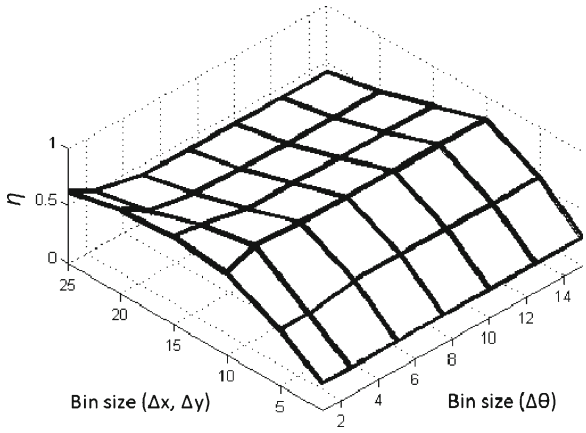


Fig. 2 Efficiency η for different bin sizes, computed according to Eq. 9. The bin size for the Cartesian variables ($\Delta x = \Delta y$) is expressed in centimeters, whereas the bin size for the angular variable $\Delta\theta$ is measured in degrees

by three variables: Δx and Δy that correspond to the resolution used to discretize the Cartesian space in which the robot operates, and $\Delta\theta$ that is the resolution used to discretize possible robot orientations. We selected a bin size with $\Delta x = \Delta y = 10$ cm and $\Delta\theta = 5°$, which corresponds to a large value of the bin efficiency metric of Eq. 9.

The second parameter of the model is the bound $\bar{\xi}$. Inconsistencies in trajectory estimation are unavoidably reflected in incorrect map estimation; therefore, as anticipated at the beginning of this section, it is desirable to select $\bar{\xi}$, so that the probability (8) corresponds to the probability of estimating a consistent map. Although in grid-based SLAM visual inspection is a common approach to evaluate map quality, a quantitative evidence of the estimated map quality can be given by comparing it with the corresponding ground truth (available in simulation). For this purpose the following metric, based on the definition of *acceptance index*, introduced by Carpin [29], is used.

Definition 1 Let M_1 and M_2 be two grid maps. The agreement between M_1 and M_2 (indicated as $\mathrm{agr}(M_1, M_2)$) is the number of cells in M_1 and M_2 that are both free or both occupied. The disagreement between M_1 and M_2 (indicated as $\mathrm{dis}(M_1, M_2)$) is the number of cells such that M_1 is free and M_2 is occupied or vice-versa, moreover the disagreement also includes the number of cells that are unvisited in M_1 and turn out to be observed in M_2. The *acceptance index* $\omega(M_1, M_2)$ between the maps is defined as:

$$\omega(M_1, M_2) \doteq \frac{\mathrm{agr}(M_1, M_2)}{\mathrm{agr}(M_1, M_2) + \mathrm{dis}(M_1, M_2)} \tag{10}$$

The acceptance index, ranging between 0 and 1, measures map similarity, once a suitable roto-translation is applied.

Contrarily to Carpin's definition in [29], unvisited cells are considered in the evaluation of the disagreement: assuming M_1 to be the true map, it is evident that the estimated map cannot contain more information than the actual one, hence it has been decided to include in the disagreement the fact that cells that are unknown in the true map, cannot have a probability different from the prior 0.5 in the estimated map.

Extensive tests on RBPF-SLAM were performed in a simple scenario with a single loop, recording for each run the maximum k_t and the corresponding acceptance index. If the latter is higher than a given threshold, the map was classified as consistent (successful experiment), otherwise the map building process failed (in our tests the threshold was fixed to 0.85, although the classification showed low sensitivity on this parameter). The histogram in Fig. 3 describes the outcome of the tests discussed in [15]: the

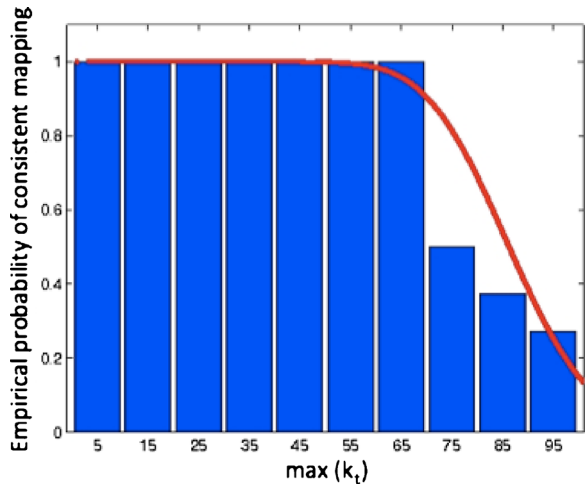


Fig. 3 Histogram represents the frequency of successful experiments for each maximum observed value of k_t. Tests were performed with a particle set size $n = 100$. These results are also discussed in our previous work [15]

frequency of successful experiments is associated to each maximum observed k_t, that is the ratio between the number of experiments in which a consistent map is estimated and the total number of experiments in which the same maximum k_t was observed. Therefore $\bar{\xi}$ can be simply computed so that the probability obtained from Eq. 8 fits the empirical probability in the histogram. In our implementation a least-squares fitting was used, obtaining $\bar{\xi} = 0.42$ (fitting curve is shown in red in Fig. 3). If a more conservative metric is needed, $\bar{\xi}$ can be selected so that the red curve is always below the histogram of experimental data.

4 Active SLAM and Exploration with RBPF

It is possible to interpret active SLAM and exploration as an optimization problem: given the information acquired until the current time (included in the SLAM posterior), the robot has to choose an exploration target (i.e., a goal point to be reached) so that the corresponding *information gain* is maximized. All the techniques for active SLAM and exploration with particle filters cannot compute a global optimum, i.e., they cannot estimate the best target over the whole map, because of the lack of closed-form expressions for the information gain as a function of the possible robot trajectories. Therefore, state-of-the-art approaches make use of simulation to evaluate the information gain in few possible target points (lately referred to as *candidate targets*). A widespread strategy for selecting such candidate targets will be reviewed in Section 4.1, with application to occupancy grid map representations. Then, in Section 4.2 the second contribution of the article is introduced, proposing a formal definition of the *expected information from a policy*, which is used to quantify the information gain at each target.

4.1 Target Selection

This section briefly recalls the *frontier-based exploration* approach for target selection, proposed in [21]. A *frontier* cell is on the boundary between visited and unexplored areas. Using occupancy grid map representation [20], unvisited areas are modeled with cells having occupation probability of 0.5, whereas visited cells assumes values close to 1 (respectively 0) if they contain obstacles (respectively free space). It is clear that the boundary between an area containing cells with values smaller than 0.5 and an area with cells having value equal to the prior probability 0.5 offers the robot a chance of visiting new places, hence they are desirable candidate targets. Therefore, from the occupancy grid map we can easily build a set of *frontier targets*, which are on the boundary of the explored region. In addition to frontier targets, in this paper we consider further candidate targets, named *trajectory targets*; trajectory targets are added along the trajectory traveled by the robot and allow the robot to come back to known places when filter uncertainty is high. Therefore, the set of candidate targets includes both *frontier targets* (corresponding to frontiers of the explored region), and *trajectory targets* (added along the trajectory of the robot). We remark that, when no frontier does exist in the map, the exploration process can be considered concluded and no target needs to be reached.

4.2 The Expected Information from a Policy

Naive criteria for selecting the exploration goal among the candidate targets do exist; for instance, goal selection can be based on the number of cells that the robot is expected to visit when reaching a given target, i.e., the goal point is the target for which the number of visited cells (cells with probability of occupation different from 0.5) is maximized. It is common, in these approaches, to "simulate" the motion of the robot approaching the target: taking assumptions on the unknown areas (e.g., obstacle-free) and considering the sensor range, one can predict the number of cells that can be visited when reaching a given target. These naive approaches only address the exploration problem, since the probabilistic nature of the estimation process is completely neglected. Definition 2 will use the results of Section 3.2 to introduce a new metric for information gain, that considers both the advantage of exploring unvisited areas, and the risk of incorrect trajectory approximation. The concept of *map information*, mentioned in the following definition, will be formally specified in Definition 3.

Definition 2 (Expected information from a policy) Let $I(m_t) = I(p(m|x_{1:t}, d_{1:t}))$ be the current *map information*, available to the robot at time t. Assume that a reasonable prediction of the amount of information, which can be acquired when applying the motion policy π to reach a given exploration target, can be computed (e.g., using simulation). If we call $I(m_{t+T(\pi)}) = I\big(p(m|x_{1:t+T(\pi)}, d_{1:t+T(\pi)})\big)$ the predicted map information after the target is reached, the *expected information from policy* π is defined as:

$$E[I(\pi)] = p(\pi)[I(m_{t+T(\pi)}) - I(m_t)] + (1 - p(\pi))[-I(m_t)] \quad (11)$$

where $p(\pi) = p(\xi(p(x_{t:t+T(\pi)}, d_{t:t+T(\pi)}) < \bar{\xi})$ is the probability of having an error smaller than $\bar{\xi}$ in the posterior approximation, computed as in Eq. 8.

The expected information from a policy is simply the expected value over the possible results of policy application: with probability $p(\pi)$ the estimated map is consistent, hence the robot has the possibility to gain the information $[I(m_{t+T(\pi)}) - I(m_t)]$, whereas with probability $1 - p(\pi)$ the filter becomes inconsistent and the robot can no longer model SLAM posterior in a proper way, loosing the information acquired before time t. When the robot is confident on its SLAM approximation ($p(\pi) \approx 1$), Eq. 11 is dominated by the first summand, hence the robot will prefer to maximize the gain $I(m_{t+T(\pi)})$ selecting exploration actions (moving towards unknown areas). When the probability $p(\pi)$ drops, that is the robot has traveled for long time in unknown areas, its policy will trade-off between loop closing actions (which contribute to $E[I(\pi)]$ with an increase of $p(\pi)$) and exploration actions (which contribute to $E[I(\pi)]$ with an increase of $I(m_{t+T(\pi)})$). The expected information at some targets may result in a negative number, truly mirroring the risk of information loss. It must be noted that the trajectory evaluation is restricted to the time interval $[t, t + T(\pi)]$ since no decision at time t can improve $p(\xi(p(x_{1:t}, d_{1:t}) < \bar{\xi})$ (the particle filter does not allow to add hypothesis on past poses).

Common entropy-based information metrics (e.g., the metrics discussed in Section 5.1) can be used to compute the map information $I(m_t)$ in Definition 2; in the following definition, an even simpler measure is proposed.

Definition 3 (N-based map information) The N-based map information N_t is the number of visited cells in the map, i.e., cells with occupancy probability different from 0.5.

According to the previous definition of map information, the expected information from a policy (Definition 2) is simplified as follows.

Definition 4 (N-based expected information from a policy) Considering N_t as map information metric, i.e., $I(m_t) = N_t$, Eq. 11 simplifies to:

$$E[I(\pi)] = p(\pi)N_{t+T(\pi)} - N_t \quad (12)$$

where current and predicted observed cells N_t and $N_{t+T(\pi)}$ are computed at time t and $t + T(\pi)$, and $T(\pi)$ is the time required to execute the motion policy π. The term N_t is common to all the targets, then can be neglected when comparing the gain at different targets. Therefore the *N-based expected information from policy* π is defined as:

$$G_{\mathrm{EI}} \doteq p(\pi)N_{t+T(\pi)} \quad (13)$$

The following definition finally allows to link the expected information to the length of the trajectory corresponding to the motion policy π.

Definition 5 (N-based normalized expected information from a policy) The *N-based normalized expected information from a policy* π is defined as:

$$F_{EI} = \frac{G_{\mathrm{EI}}}{l} \quad (14)$$

in which the expected information is normalized by the distance l that has to be traveled to execute policy π.

The N-based normalized expected information assumes the meaning of a gain for each traveled meter, and intrinsically takes into account the disadvantage of reaching targets which are far away.

In the rest of the article, F_{EI} will referred to as *expected information (from a policy)*, implying that an N-based metric is used and that the information is normalized by the distance to the goal.

5 Experiments and Discussion

During autonomous exploration, the robot first selects the candidate targets, according to a given procedure (e.g., the frontier-based selection presented in Section 4.1). Then the robot uses a metric to evaluate the information gain in reaching each candidate target: the candidate target for which the expected gain is maximum will be selected as a goal for robot motion. After reaching this *exploration goal* (or *goal target*), the decisional process is iterated until the target selection phase will produce no candidate target. The first part of this section describes the information gain metrics used in state-of-the-art techniques. In particular, the *entropy-based* gain, proposed in [19], the *expected map information* [14], and a simpler metric (*naive* gain) are recalled. Then, the results of the implementation of the proposed method are presented and compared with the performances of these related techniques. The tests were designed in MobileSim and the robotic platform used was an *ActivMedia Pioneer P3-DX* equipped with a laser range sensor and odometric pose estimation. In our RBPF-SLAM implementation, two important aspects are incorporated such as the *adaptive resampling technique* [19] to reduce the problem of particle depletion, and the *laser-stabilized odometry* [30] to improve the prediction phase of the filter. As in [14] and [19] the predicted laser measurements are obtained via ray-casting on the map of the best particle, assuming that the unknown areas are obstacle free. The map of the best particle is also used to plan a path to reach a given target, using the A* algorithm.

5.1 State-of-the-Art Approaches

5.1.1 Joint Entropy

This technique is based on the joint entropy of map and trajectory posterior. The *joint entropy* of the SLAM posterior is approximated as [19]:

$$H\left(p\left(x_{1:t}, m \mid d_{1:t}\right)\right) \approx H\left(p\left(x_{1:t} \mid d_{1:t}\right)\right) + \sum_{i=1}^{n} w_t^{[i]} \times H\left(p\left(m^{[i]} \mid x_{1:t}^{[i]},\ d_{1:t}\right)\right) \tag{15}$$

where n is the number of particles in the filter and $w_t^{[i]}$ is the weight of the i-th particle at time step t. The first summand in Eq. 15 quantifies trajectory entropy, whereas the second one corresponds to map entropy. The entropy of the i-th map in the filter can be computed as:

$$H\left(p\left(m^{[i]} \mid x_{1:t}^{[i]},\ d_{1:t}\right)\right) = \sum_{\forall u,v} H\left(p\left(m_{uv}^{[i]} \mid x_{1:t}^{[i]},\ d_{1:t}\right)\right) \tag{16}$$

where $p(m_{uv}^{[i]} \mid x_{1:t}^{[i]},\ d_{1:t}))$ describes the probability of a single cell in the occupancy grid map (corresponding to the indices u and v) and the binary entropy function $H(\cdot)$ is defined as $H(p) \doteq -p\log(p) - (1-p)\log(1-p)$.

For the computation of trajectory posterior entropy, Roy et al. [31] and subsequently Stachniss et al. [19] proposed to approximate the entropy of the point mass distribution with the entropy of the Gaussian distribution fitting particle poses:

$$H\left(p(x_{1:t} \mid d_{1:t})\right) \approx \frac{1}{t} \sum_{\tau=1}^{t} H\left(p(x_\tau \mid d_{1:\tau})\right) \tag{17}$$

where $p(x_\tau \mid d_{1:\tau})$ is a Gaussian approximation of pose posterior at time τ. Therefore, the highest information gain is attained by the target that maximizes the joint entropy reduction:

$$G_{\mathrm{H}} = H\left(\, p(x_{1:t},\ m \mid d_{1:t})\right) - H\left(\, p(x_{1:t+T(\pi)},\ m \mid d_{1:t+T(\pi)})\right) \tag{18}$$

where current and predicted joint entropies are computed at time t and $t + T(\pi)$.

Although joint entropy is a widespread approach for particle-filter based exploration, some drawbacks are commonly recognized:

1. it depends on the overall grid area, instead of the actual observed cells;
2. it is function of the grid resolution, which is the reason why the entropy of any map increases unbounded when the resolution increases;
3. if the robot entropy is computed using Gaussian fitting as in [19, 31], the method leads to the entropy divergence when only one particle survives from resampling (common after loop closing) unless ad-hoc heuristics are applied;
4. trajectory entropy contribution is negligible with respect to map entropy.

The computational complexity of joint entropy is $O(n(M+l))$, where M is the number of cells in the map and l is the length of trajectory to reach the target. The objective function F_{H}, which the target has to maximize, is usually augmented with some cost that penalizes far targets [7, 19]:

$$F_{\mathrm{H}} = G_{\mathrm{H}} - \alpha l \tag{19}$$

where α is a suitable weight for the path length l.

5.1.2 Expected Map Information

In order to avoid some of the undesirable properties of joint entropy, the *expected map information* has been proposed in [14]. Before introducing this information metric, the concept of *Expected Map (EM)* for RBPF needs be briefly reviewed. Taking into account all the maps associated to the samples in the filter, EM is defined as the mathematical expectation $E[\cdot]$ of the map hypotheses over the particle set:

$$p(EM \mid d_{1:t}) = E_{x_{1:t}}\left[p(m \mid x_{1:t}, d_{1:t})\right] \tag{20}$$

Each cell u, v of the expected map can be obtained as:

$$p(EM_{uv} \mid d_{1:t}) \approx \sum_{i=1}^{n} w_t^{[i]} p\left(m_{uv}^{[i]} \mid x_{1:t}^{[j]},\ d_{1:t}\right) \tag{21}$$

The information of the cell (u, v) is hence defined as:

$$I\left(p(EM_{uv} \mid d_{1:t})\right) = 1 - H(p(EM_{uv} \mid d_{1:t})) \tag{22}$$

Accordingly, the *expected map information* can be computed as:

$$I\left(p(EM \mid d_{1:t})\right) = \sum_{\forall u,v} I(p(EM_{uv} \mid d_{1:t})) \tag{23}$$

Therefore, the information gain after the execution of the motion action π is:

$$G_{\mathrm{EMI}} = I\left(p(EM \mid d_{1:t+T(\pi)})\right) - I\left(p(EM \mid d_{1:t})\right) \tag{24}$$

The computation complexity of the expected map information (EMI) is $O(nM)$. As for the joint entropy, the objective function for exploration is augmented with a term penalizing long paths:

$$F_{\mathrm{EMI}} = G_{\mathrm{EMI}} - \alpha l \tag{25}$$

where α is a suitable weight for the path length l.

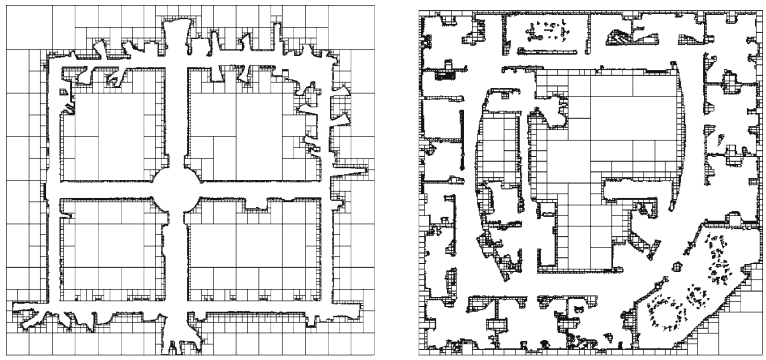
5.1.3 Naive Metric

The third metric considered for the experimental comparison is based on the number of cells in the occupancy grid map of the best sample (the particle with the highest weight) that can be visited when performing a policy. It is called *naive* since it neglects both the uncertain description of robot poses and the probabilistic interpretation of sensor measurements. A cell is considered as observed when its occupancy probability is different from 0.5. Starting from this consideration, the *naive* gain G_{N} is defined as the difference between the predicted number of cells observed after policy π is executed and the number of cells observed until the current time t (see also Definition 3):

$$G_{\mathrm{N}} = N_{t+T(\pi)} - N_t \tag{26}$$

5.2 Case Studies

For the purpose of evaluating the techniques presented so far in realistic scenarios, extensive tests were carried out in typical indoor environments, simulating the case in which a robot, equipped with a laser range finder is deployed in the scenario and is required to explore it, producing



(a) ACES (b) Intel research lab

(c) Freiburg 079 building

Fig. 4 **a** MobileSim map of the ACES building; **b** MobileSim map of the Intel Research Lab; **c** MobileSim map of the Freiburg 079 building

an occupancy grid map. The range of the laser-scanner is 16 m, whereas the number of particles in the filter was set to 50. Instead of considering manually designed simulation scenarios that can be an oversimplification of real exploration procedures and that may not catch the complexity of a real environment, real maps were considered and converted in the corresponding representation to be used in the MobileSim simulator [32]. The use of maps that are well-known to the robotic community makes the experiments clearer and repeatable: for this reason an online version of all the considered maps has been released [33]. In particular three test scenarios were considered: (1) the ACES building at the University of Texas, which covers an area of about 45 m by 40 m, in Fig. 4a; (2) the Intel Research Lab in Seattle with size of 28 m by 28 m, in Fig. 4b; and (3) University of Freiburg 079 building, in which the main corridor length is about 36 m, in Fig. 4c. The aforementioned maps were available online as occupancy grid maps [34] and they needed to be converted in a proper format to be used in MobileSim; Fig. 4a–c contain the already processed maps.

Two case studies were first considered to remark the peculiarity of each exploration approach. These case studies were selected during the tests on ACES building; this scenario is particularly relevant for the analysis because the presence of crossroads allows to clearly discern the differences in the decisional processes that lie behind each exploration approach.

Consider the first case study, depicted in Fig. 5a: the robot starts to explore the environment and reaches the point labeled with a red dot. At this point it performs target selection, according to the frontier-based technique presented in Section 4.1, and obtains the candidate targets that are indicated in the figure as blue dots. The candidate targets include both frontier targets, i.e., $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, and trajectory targets, i.e., $\{500, 501, 502\}$. This case study is important because the environment offers the robot the opportunity to perform a loop closing, reducing the SLAM posterior uncertainty. Moreover, because of the size of the environment, if the robot misses this loop closing and continues on the corridor traveling towards target 3, the uncertainty in the posterior unavoidably leads to a failure of

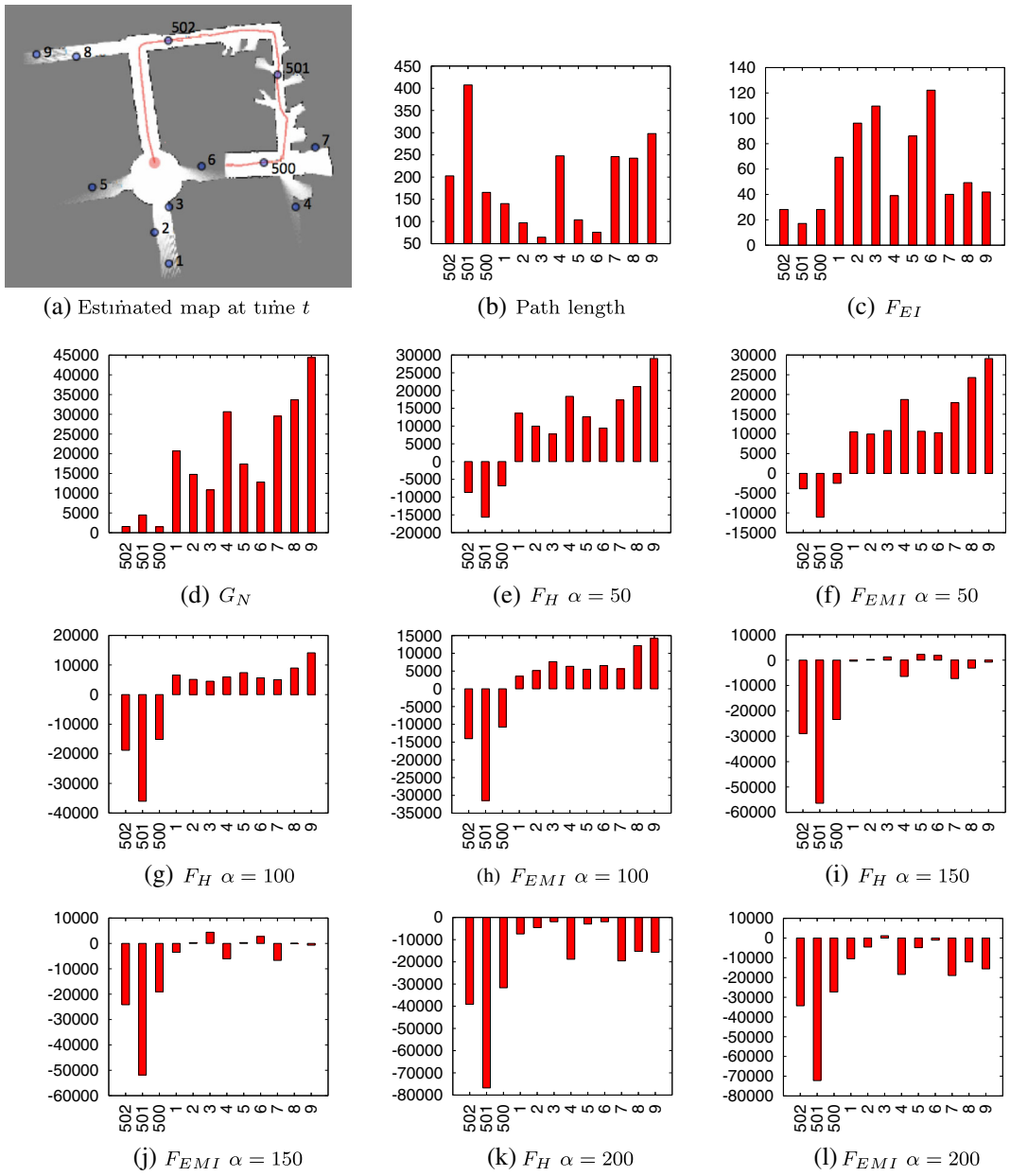


Fig. 5 Case study 1 (ACES building): **a** map of the scenario including robot position (*red dot*) and the position of the candidate targets (*blue dots*). The candidate targets are labelled as follows: targets $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ are the *frontier targets*, while targets $\{500, 501, 502\}$ are *trajectory targets* (see Section 4.1); **b** path length for each candidate target; **c–l** bar plots of the target scores: to each candidate target (*x-axis in the plots*) we associate a score (*y-axis in the plots*), computed according to the corresponding metric

the mapping process. The score assigned by the different approaches to each target are evaluated, comparing the following objective functions:

1. expected information gain from a policy (F_{EI});
2. joint entropy (F_H) with $\alpha = \{50, 100, 150, 200\}$;
3. expected map information (F_{EMI}), with $\alpha = \{50, 100, 150, 200\}$;
4. the naive metric G_N.

The results of the analysis in this case study are summarized in Fig. 5. Figure 5a shows the

estimated map at time t, and the candidate (blue dots), among which the robot (red dot) has to choose the next motion goal. Figure 5b shows the length of the path to reach each candidate target (targets are listed on the x-axis). Path length is computed as the number of steps required by the robot to reach a given candidate target; in our simulation a step corresponds to 0.1m. Figure 5c–l report the scores (objective function values) assigned by each technique to the targets: in all the cases the best target is the one that maximizes the objective function. Consider first the *joint entropy* approach. In Fig. 5e the cost of traversing long distances has smaller weight ($\alpha = 50$) and the information content of the target is dominant in the definition of the goal target: in this case F_H simply behaves like G_N, and this is evident comparing Fig. 5e and d. This comes at no surprise, since (i) trajectory entropy contribution is negligible with respect to map entropy; (ii) map entropies, with respect to Eq. 16, are almost equal in practice, hence the following simplifications are suggested by experimental evidence:

$$\begin{aligned}
&\sum_{i=1}^{n} w_t^{[i]} H\left(p(m^{[i]} \mid x_{1:t}^{[i]},\ d_{1:t})\right) \\
&\quad \approx H\left(p(m^{[*]} \mid x_{1:t}^{[*]},\ d_{1:t})\right) \sum_{i=1}^{n} w_t^{[j]} \\
&\quad = H\left(p(m^{[*]} \mid x_{1:t}^{[*]},\ d_{1:t})\right) = f(N_t)
\end{aligned}$$

where * denotes the best sample and $f(\cdot)$ is a monotonically decreasing function. Then the maximum of F_H corresponds to the target with the smallest entropy which, in turn, is the one having the maximum value of G_N. Figure 5i and k show the results obtained when $\alpha = 150$ and $\alpha = 200$ are used in F_H, respectively. In these cases the outcomes suggest that the behavior of the joint entropy can be directly deduced by the path length, see Fig. 5b: the scores are high for short paths and low for long paths, hence reducing the exploration approach to the choice of the nearest candidate target. When selecting $\alpha = 100$, the target selection is still unsatisfactory: the target of choice is the number 9, which is far from the robot and does not correspond to a loop closing.

Consider now the *expected map information*. When using F_{EMI} with $\alpha = 200$, see Fig. 5l, the scores of the targets resemble the corresponding scores in F_H. Few differences appear in general between F_{EMI} and F_H; in particular for $\alpha = 50$ and $\alpha = 100$ the best target is the same as the one suggested by F_H (target number 9), and also in this case the choice leads the robot to a target that is far and does not consider the uncertainty in robot posterior. Comparing Fig. 5j and i it can be noticed that the target labeled with number 3 acquires a higher score; however such difference does not improve the exploration effectiveness: target 3 does not correspond to a loop closing and both F_H and F_{EMI} underestimate the risk of filter inconsistency.

On the other hand, Fig. 5c remarks that the *expected information from a policy* (F_{EI}) returns a high score for target 6, which corresponds to a loop closing. Moreover, also the other high informative targets in F_{EI} (see targets $1, 2, 3, 5$ in Fig. 5c) are reasonable targets for exploration since they are close to the robot and allow to explore new areas. The behavior of F_{EI} can be easily understood from Eqs. 13 and 14. Equation 13 states that when the uncertainty in the filter is small ($p(\pi)$ close to 1 for all targets) the proposed metric F_{EI} will prefer targets that maximize the number of explored cells. However, when the uncertainty in the filter is large (as in our case study), the proposed approach will trade-off between the advantage of observing new cells (quantified by $N_{t+T(\pi)}$) and the probability of correct map estimation (which is $p(\pi)$). For instance, the proposed approach can completely disregard exploration targets leading to $p(\pi) \approx 0$, i.e., policies leading to low probability of consistent mapping. The probability ($1\text{-}p(\pi)$) is essentially the probability of failure in the mapping process, therefore the proposed approach, contrarily to state-of-the-art techniques, explicitly considers the risk of failure, leading to more robust and uncertainty-aware strategies. Finally, the normalization by the path length in Eq. 14 naturally penalizes far targets.

A second case study is reported in Fig. 6: in this scenario the robot, after completing a first loop closing, reaches a second crossroad, in which it has to decide between exploring new areas (i.e., going straight on towards target 1) or closing the loop (turning towards target 5). Also this case study remarks the previous observation on the state-of-the-art approaches: F_H ($\alpha = 200$) and F_{EMI}

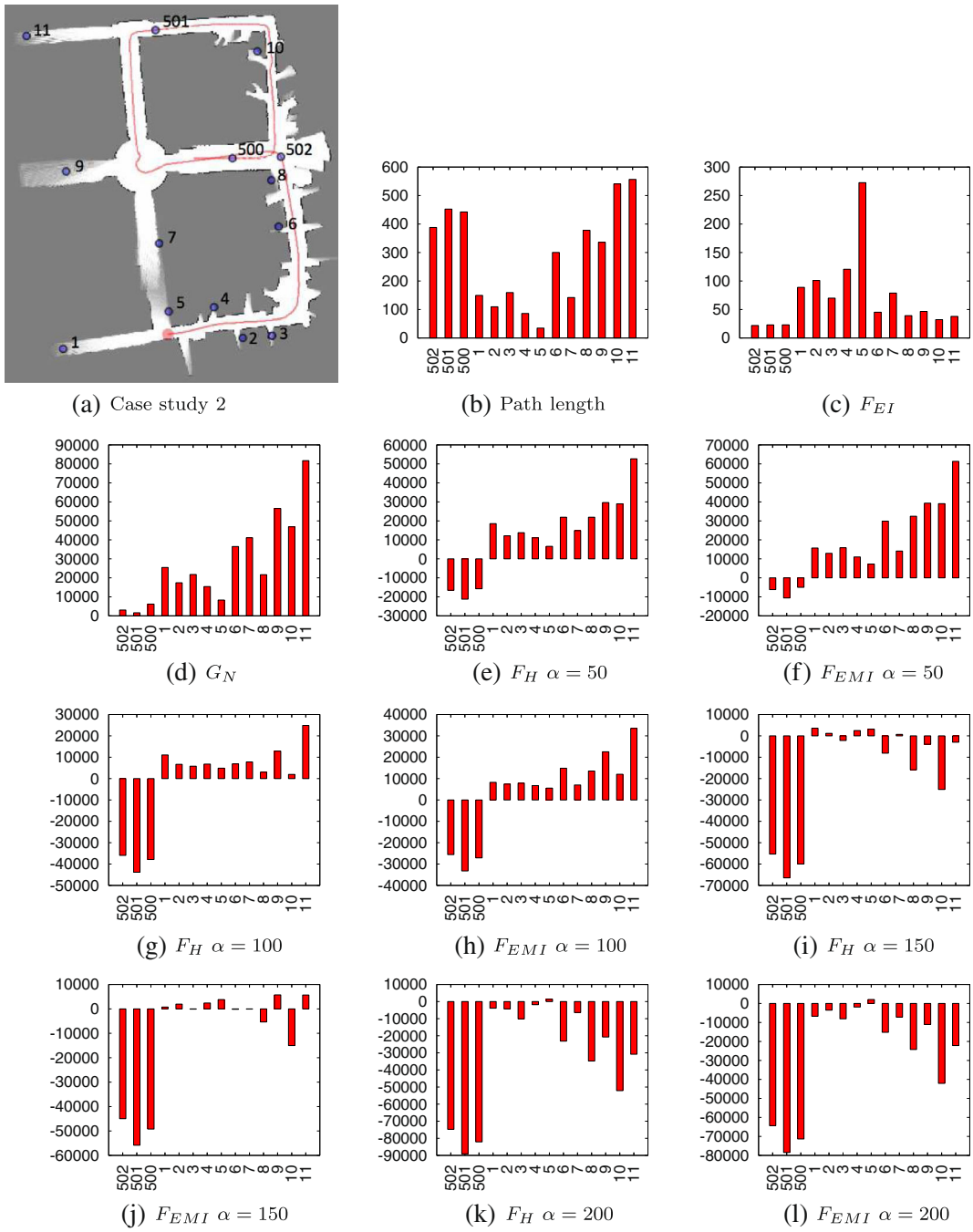


Fig. 6 Case study 2 (ACES building): **a** map of the scenario including robot position (*red dot*) and the position of the candidate targets (*blue dots*). The candidate targets are labelled as follows: targets {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11} are the *frontier targets*, while targets {500, 501, 502} are *trajectory targets* (see Section 4.1); **b** path length for each candidate target; **c–l** bar plots of the target scores: to each candidate target (*x-axis in the plots*) we associate a score (*y-axis in the plots*), computed according to the corresponding metric

($\alpha = 200$) resembles the inverse behavior of the path length, i.e., the distance from the path is crucial in the determination of the goal target; F_H ($\alpha = \{50, 100\}$) and F_{EMI} ($\alpha = \{50, 100\}$) present few differences with respect to G_N, hence questioning the need of using complex metrics in the

Table 1 Goal targets selected by each of the compared approaches for the case studies under analysis

Case	F_{EI}	F_H				F_{EMI}				G_N	Shortest
studies		50	100	150	200	$\alpha = 50$	100	150	200		path
1	6	9	9	6	6, 3	9	9	3	3	9	3
2	5	11	11	1	5	11	11	9	5	11	5

For F_H and F_{EMI} we report results for the following choices of α: {50, 100, 150, 200}

exploration process. The expected information from a policy F_{EI}, however, is able to detect the *chance* of loop closing.

Table 1 summarizes the results of the two case studies by reporting only the best target (i.e., the target that will be selected as goal for robot motion) for each exploration approach. When the goal target is not well defined, because two targets have comparable values, both targets are reported.

It should be clear at this point that, for F_H and F_{EMI}, the cost associated with the path length is hard to balance and the parameter α is only able to trade-off between two suboptimal strategies, i.e., (i) maximizing the number of visited cells and (ii) moving towards the nearest candidate target. On the other hand, the parameter $\bar{\xi}$ in the definition of *expected information from a policy* has a clear effect on the robot behavior: if the robot has to be more conservative in exploration we can simply set a lower value of $\bar{\xi}$ in Eq. 8. This corresponds to impose a stricter constraint on SLAM uncertainty, hence forcing the robot to prefer loop closing actions instead of visiting unknown areas. If the risk of map inconsistency is acceptable, the upper bound $\bar{\xi}$ is relaxed and the trade-off will lead towards exploration actions. When $\bar{\xi}$ is further increased, the probability $p(\pi)$ remains equal to 1, hence reducing the proposed approach to the naive gain (26).

5.3 Autonomous Exploration

The compared techniques are now applied in autonomous exploration experiments, so to evaluate the potential of the strategies in solving real navigation problems. Since a *ground truth* map is available, i.e., the map on which the simulation scenario is built, the acceptance index of Definition 1 can be used to measure the similarity between the true map and the map estimated with the different exploration approaches, so to evaluate the goodness of the mapping process. Examples of acceptance indices for maps estimated in the tests are reported in Fig. 7. For each map (Freiburg 079 building, Intel Lab, ACES building), we generated ten random initial conditions

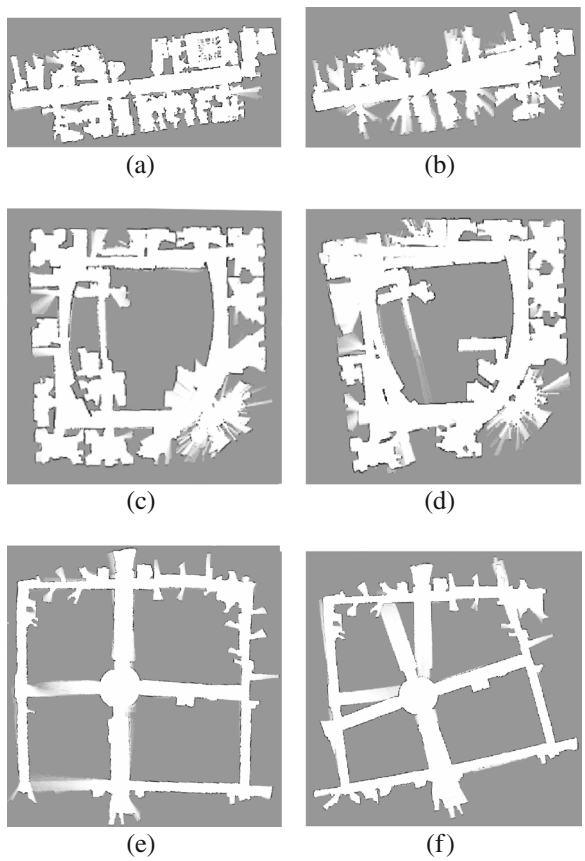


Fig. 7 Freiburg 079 building: **a** example of consistent map, acceptance index $\omega = 0.84$; **b** example of inconsistent map, acceptance index $\omega = 0.54$. Intel Research Lab: **c** example of consistent map, acceptance index $\omega = 0.84$; **d** example of inconsistent map, acceptance index $\omega = 0.74$. Aces building: **e** example of consistent map, acceptance index $\omega = 0.79$; **f** example of inconsistent map, acceptance index $\omega = 0.33$. All the acceptance indices are computed with respect to the ground truth map

(initial poses of the robot in the unknown map). For each initial condition, we simulated the behavior of the robot performing active SLAM and exploration using each of the following metrics:

1. expected information from a policy, F_{EI}
2. joint entropy, F_H with $\alpha = \{50, 100, 150, 200\}$
3. expected map information F_{EMI} with $\alpha = \{50, 100, 150, 200\}$
4. naive metric, G_N
5. a distance based metric, G_l, in which the robot always travels to the closest target

Table 2 reports the average acceptance index ω indices for the different approaches in the tested scenarios; the corresponding standard deviation is also shown in parenthesis. The table also contains the average time required to complete the exploration process and the number of experiments in which the mapping process produced a visually inconsistent occupancy grid map. The acceptance index indicates the average quality of the map estimated with a given approach, the exploration time is connected with the computational complexity of the approach, whereas the number of failures can be considered as an indicator of the robustness of the approach.

Consider first the results of autonomous exploration at the Freiburg 079 building (FR079). The scenario comprises a main corridor and several small offices. This structure suggests that the robot is forced to traverse several times the corridor, hence, imposing several loop closings. For this reason all the techniques are expected to have good performance, also because of the small size of the environment. This expectation is partially confirmed by the acceptance index of the different approaches in Table 2. However, it is possible to notice that in some cases (e.g., F_H with $\alpha = 50$, F_{EMI} with $\alpha = 50$ or F_{EMI} with $\alpha = 200$) the index drops, because of the failure of some test. These experiments show that the suboptimal strategies G_N and G_l provide poor results, and hence the probabilistic aspect of active SLAM and exploration has a remarkable importance also in small scenarios. According to the considerations reported in the previous section, also F_H with $\alpha = \{50, 200\}$ and F_{EMI} with $\alpha = \{50, 200\}$ are not effective solutions for the problem, since they resemble the behavior of the naive metrics G_N and G_l. The remaining metrics allow to have acceptable performance in terms of map quality. However, the expected information from a policy still outperforms the other approaches in terms of robustness (no failure was registered in the Freiburg scenario) and in terms of time required for exploration. The standard deviation of ω remains pretty small across the techniques, and is

Table 2 Autonomous exploration performance with state-of-the-art approaches

		F_{EI}	F_H				F_{EMI}				G_N	G_l
			50	100	150	200	50	100	150	200		
FR079	ω	0.77	0.71	0.79	0.78	0.78	0.71	0.77	0.79	0.71	0.71	0.68
		(0.06)	(0.08)	(0.03)	(0.07)	(0.06)	(0.08)	(0.04)	(0.02)	(0.09)	(0.11)	(0.07)
	Time	153	212	187	178	143	200	179	219	164	196	131
	#Fail	0	5	0	1	3	3	2	1	4	4	3
Intel	ω	0.80	0.75	0.80	0.78	0.76	0.75	0.76	0.77	0.73	0.73	0.75
		(0.03)	(0.14)	(0.05)	(0.04)	(0.04)	(0.08)	(0.09)	(0.04)	(0.08)	(0.07)	(0.07)
	Time	204	261	245	240	224	260	270	250	247	232	142
	#Fail	2	6	4	3	5	6	5	5	6	8	6
ACES	ω	0.63	0.50	0.57	0.49	0.58	0.42	0.53	0.54	0.59	0.47	0.58
		(0.13)	(0.16)	(0.06)	(0.09)	(0.12)	(0.10)	(0.18)	(0.14)	(0.09)	(0.13)	(0.10)
	Time	240	178	265	310	252	181	299	266	287	164	123
	#Fail	6	10	10	10	10	10	9	9	9	9	10

Average acceptance index ω, average time (in minutes) required to complete exploration, and number of failures for each of the compared approaches. The standard deviation of the observed ω is also reported in parenthesis. We compare the following metrics: expected information from a policy (F_{EI}), joint entropy (F_H), expected map information (F_{EMI}), the naive metric of Section 5.1.3 (G_N), and a distance-based metric (G_l). For F_H and F_{EMI} we report results for the following choices of α: {50, 100, 150, 200}

larger for techniques having several failures: a failure, in fact, usually corresponds to a small value of ω, increasing the spread (hence the standard deviation) of the observed ω values.

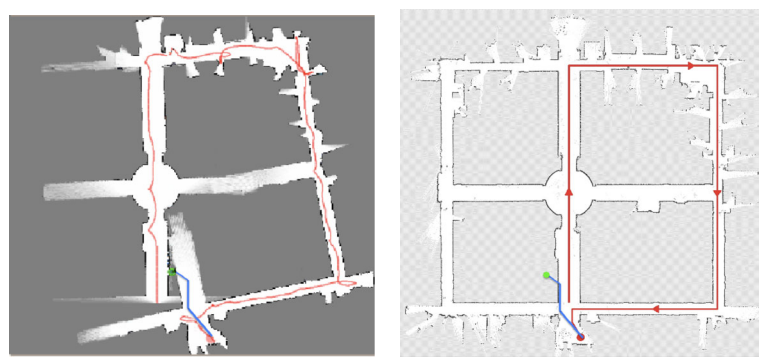
Similar observations can be done with regard to the Intel Research Lab. In the corresponding rows of Table 2 there is only a small difference between the average map quality. However in this case, the number of failures starts to be relevant for the presence of several areas to be explored and several obstacles. The proposed approach still shows the lowest number of failures and its exploration time is smaller than the state-of-the-art techniques.

The most challenging test bench is the ACES building, since in this case it is clear that robot decision making is crucial in accomplishing a consistent map estimation. In the previous scenarios, the size and the structure of the environment forces the robot to revisit known places, hence providing a natural reduction of the SLAM uncertainty. In this last case, instead, the presence of crossroads requires an effective exploration strategy: missing a loop closing means being unable to estimate a correct map, hence a wrong decision is clearly visible from the experimental results. The results reported in Table 2 show that the state-of-the-art approaches and the naive metrics failed in practically all the tests. This comes at no surprise from the considerations in Section 5.2. The proposed approach, based on the objective function F_{EI}, has a large number of failures, too, but it shows higher robustness for its capability of identifying loop closings. Also in this case, the $O(n)$ complexity of the approach allows to explore the whole environment in a reasonable amount of time.

It must be underlined that the experiment can be made easier considering, for instance, a higher range for the laser, or a larger number of particles. However, a more challenging situation was preferred in the tests to have a clear picture of the approach limitations: a longer sensor range or a larger particle set size can only hide the issues underlying the approaches in small environments, but they will still present the same drawbacks in large-scale scenarios.

Remark 2 A poor exploration strategy may affect map quality; however, also the reverse implication holds: a poor map estimate may easily lead to exploration failure. An example of this situation is reported in Fig. 8: the robot is not able to correctly close the loop and in the estimated map it appears a new corridor that needs to be explored; in the real map, however, the robot is in a known area.

We conclude this section with a last analysis of the results in Table 2. We want to exclude the case in which the variations of ω in Table 2 are simply due to statistical fluctuations and are not significant in assessing the difference among the analyzed techniques. For this purpose we apply a standard statistical tool, known as ANOVA (ANalysis Of VAriance) to compare the results obtained from simulations. We recall that for each technique we repeated 10 experiments and for each one we recorded the corresponding acceptance index ω; then in Table 2 we reported mean and standard deviation of the 10 realizations. ANOVA tests the *null hypothesis* that the 10 realizations of ω produced by different techniques are drawn from the same distribution. Roughly



(a) (b)

Fig. 8 ACES building—**a** Estimated map with: estimated robot position (*red dot*), estimated robot trajectory (*red line*) and planned robot trajectory (*blue line*); **b** Ground truth map with: actual robot position (*red dot*), real trajectory (*red line*) and planned trajectory (*blue line*)

speaking, ANOVA provides a statistical test of whether or not the means of several groups are equal, and in our particular application allows to assess that the differences in the mean values in Table 2 are statistically significant. We applied the ANOVA analysis for each scenario across all techniques, obtaining the following probabilities for the null hypothesis: $P_{\text{FR079}} = 0.001$, $P_{\text{Intel}} = 0.333$, and $P_{\text{ACES}} = 0.019$. The ANOVA analysis, for the FR079 and the ACES scenarios, indicates that the null hypothesis can be safely rejected. The probability is remarkably higher for the Intel scenario, where we already observed that all techniques perform similarly in terms of average ω.

5.4 Loop Closing Detection

For the sake of completeness a last experiment is reported to evaluate the capability of the metrics considered in this paper in detecting loop closing occurrence. Loop closing detection is important when dealing with the problem of nested loops, where the robot has to decide when to stop following an already traversed area, in order not to lose significant hypotheses of the particle filter [19]. The experiment was carried out in a real scenario with the robot performing SLAM in the corridors of Politecnico di Torino in an area of about 50 m × 45 m. The estimated map is shown in Fig. 9. The robot starts from point 1, first closing the loop denoted with number 2. Then it travels a longer path closing the loop corresponding to number 3. It explores a large area and closes the loop labeled with 4. The robot finally reaches position 5, traveling back to the starting point.

Fig. 9 Map estimated in real world experiment performed at Politecnico di Torino (100 particles)

In this case the following uncertainty metrics are compared:

1. the *joint entropy*, H, computed from Eq. 15;
2. the *expected map mean information* (EMMI), which is a normalized version of the expected map information [14]. It assumes values between 0 and 1, and it can be considered a measurement of robot certainty, in the sense that values close to 1 denote low uncertainty, whereas values close to 0 denote high uncertainty;

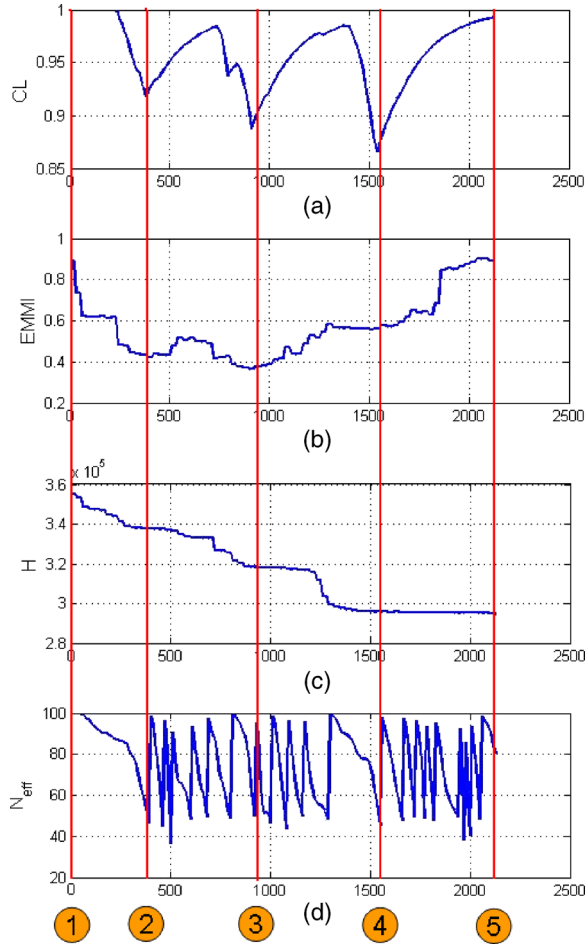


Fig. 10 Time-history of the different uncertainty metrics obtained in real experiment performed at Politecnico di Torino (100 particles): **a** proposed KLD-based metric (CL), **b** expected map mean information (EMMI), **c** joint entropy (H), and **d** effective sample size ($\tilde{N}_{\text{eff}}$)

3. the proposed metric based on Kullback-Leibler divergence computed from Eq. 8.

An exponential average on the values of the proposed metric was applied to smooth fast changes of the metric due to the effect of discretization of the state space; further discussion on the topic are reported in [16]. Such a smoothed metric is called *confidence level* (CL), since it is a reliable indicator of the *certainty* of the mapping process.

Comparing the behavior of the metrics in Fig. 10, it is evident that EMMI [14] is influenced by loop closing occurrence, although it is often a poor indicator of place revisiting. Joint entropy remains almost constant in correspondence of loop closing situations, but its dependency on grid extends and other previously mentioned drawbacks prevent its use for loop closing detection. The confidence level, instead, clearly shows place revisiting actions, which correspond to a quick rise of the metric (increasing certainty level). Further results on the behavior of these uncertainty metrics in real and simulated tests are reported in our previous work [16].

6 Conclusions

This work investigates the problem of active SLAM and exploration with Rao-Blackwellized Particle Filters. An application of Kullback-Leibler divergence is proposed to evaluate the particle-based SLAM posterior approximation. This metric is then employed in the definition of an *expected information from a policy*, which allows the robot to autonomously decide the best motion strategy to concurrently explore the environment and reduce the uncertainty in SLAM posterior. The probabilistic interpretation of the proposed information gain is then validated, by comparing it with state-of-the-art metrics of information gain. Several case studies highlight that the proposed technique enhances robot awareness in detecting loop closing occasions, which are often missed when using related approaches.

As a further contribution the results of an extensive experimental analysis on autonomous exploration are presented. Exploration procedures are compared in terms of time required for exploration (i.e., computational effort), average quality of the estimated map (i.e., fidelity to the ground truth map) and robustness (i.e., number of failures in the mapping process). The test scenarios are well-known SLAM benchmarking environments and the map used in simulations are available online [33] to stimulate future comparisons and discussions. The results provide an original insight on the exploration approaches, demonstrating that most of the mainstream approaches are not able to properly catch the probabilistic aspects of the problem, and several state-of-the-art solutions present a cumbersome computational cost that is not justified by any performance improvement with respect to naive strategies (e.g., moving to the nearest exploration target).

Acknowledgements This work was funded by Ministero dell'Istruzione, dell'Università e della Ricerca under MEMONET National Research Project.

References

1. Durrant-Whyte, H., Bailey, T.: Simultaneous localisation and mapping (SLAM): part I. The essential algorithms. Robot. Autom. Mag. **13**, 99–110 (2006)
2. Durrant-Whyte, H., Bailey, T.: Simultaneous localisation and mapping (SLAM): part II. State of the art. Robot. Autom. Mag. **13**, 108–117 (2006)
3. Castellanos, J.A., Martinez-Cantin, R., Tardós, J.D., Neira, J.: Robocentric map joining: improving the consistency of EKF-SLAM. Robot. Auton. Syst. **55**(1), 21–29 (2007)
4. Doucet, A., de Freitas, J., Murphy, K., Russel, S.: Rao-Blackwellized particle filtering for dynamic bayesian networks. In: Proc. of the Conference on Uncertainty in Artificial Intelligence, pp. 176–183 (2000)
5. Thrun, S., Burgard, W., Fox, D.: Probabilistic Robotics. MIT Press (2005)
6. Martinez-Cantin, R., De Freitas, N., Castellanos, J.: Analysis of particle methods for simultaneous robot localization and mapping and a new algorithm: marginal-SLAM. In: Proc. of the IEEE lnternational Conf. on Robotics and Automation (2007)
7. Bourgault, F., Makarenko, A.A., Williams, S.B., Grocholsky, B., Durrant-Whyte, H.F.: Information based adaptive robotic exploration. In: Proc. of the IEEE-RSJ Int. Conf. on Intelligent Robots and Systems, pp. 540–545 (2002)
8. Martinez-Cantin, R., De Freitas, N., Doucet, A., Castellanos, J.A.: Active policy learning for robot planning and exploration under uncertainty. In: Proc. of Robotics: Science and Systems (2007)

9. Sim, R., Roy, N.: Global A-optimal robot exploration in SLAM. In: Proc. of the IEEE Int. Conf. on Robotics and Automation, pp. 661–666 (2005)
10. Martinez-Cantin, R., De Freitas, N., Brochu, E., Castellanos, J., Doucet, A.: A Bayesian exploration-exploitation approach for optimal online sensing and planning with a visually guided mobile robot. Auton. Robot. **27**(2), 93–103 (2009)
11. Carrillo, H., Reid, I., Castellanos, J.A.: On the comparison of uncertainty criteria for active SLAM. In: Proc. of IEEE Int. Conf. on Robotics and Automation, pp. 2080–2087 (2012)
12. Valencia, R., Andrade-Cetto, J., Porta, J.M.: Path planning in belief space with pose SLAM. In: Proc. of the IEEE Int. Conf. on Robotics and Automation, pp. 78–83 (2011)
13. Valencia, R., Valls Miró, J., Dissanayake, G., Andrade-Cetto, J.: Active pose SLAM. In: Proc. of the IEEE Int. Conf. on Intelligent Robots and Systems, pp. 1885–1891 (2012)
14. Blanco, J.L., Fernandez-Madrigal, J.A., Gonzalez, J.: A novel measure of uncertainty for mobile robot SLAM with Rao-Blackwellized Particle Filters. Int. J. Robot. Res. **27**(1), 73–89 (2008)
15. Carlone, L., Du, J., Kaouk Ng, M., Bona, B., Indri, M.: An application of Kullback-Leibler divergence to active SLAM and exploration with particle filters. In: Proc. of the IEEE Int. Conf. on Intelligent Robots and Systems, pp. 287–293 (2010)
16. Carlone, L., Kaouk Ng, M., Du, J., Bona, B., Indri, M.: Reverse KLD-sampling for measuring uncertainty in Rao-Blackwellized Particle Filters SLAM. In: Proc. of the Workshop on Performance Evaluation and Benchmarking for Next Intelligent Robots and Systems. IEEE-RSJ Int. Conf. on Intelligent Robots and Systems (2009)
17. Du, J., Carlone, L., Kaouk Ng, M., Bona, B., Indri, M.: A comparative study on active SLAM and autonomous exploration with particle filters. In: Proc. of the IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics, pp. 916–923 (2011)
18. Kullback, S., Leibler, A.: On information and sufficiency. Ann. Math. Stat. **22**, 79–86 (1951)
19. Stachniss, C., Grisetti, G., Burgard, W.: Information gain-based exploration using Rao-Blackwellized Particle Filters. In: Proc. of Robotics: Science and Systems (2005)
20. Moravec, H.P.: Sensor fusion in certainty grids for mobile robots. AI Mag. **9**(2), 61–74 (1988)
21. Yamauchi, B.: A frontier-based approach for autonomous exploration. In: Proc. of CIRA vol. 97, pp. 146–151 (1997)
22. Fox, D., Ko, J., Konolige, K., Limketkai, B., Schulz, D., Stewart, B.: Distributed multirobot exploration and mapping. Proc. IEEE **94**(7), 1325–1339 (2006)
23. Moorehead, S.J., Simmons, R., Whitaker, W.L.: Autonomous exploration using multiple sources of information. In: Proc. of the IEEE Int. Conf. on Robotics and Automation, pp. 3098–3103 (2001)
24. Stachniss, C., Hahnel, D., Burgard, W.: Exploration with active loop-closing for FastSLAM. In: Proc. of the IEEE-RSJ Int. Conf. on Intelligent Robots and Systems, pp. 1505–1510 (2004)
25. Ko, J., Stewart, B., Fox, D., Konolige, K., Limketkai, B.: A practical decision-theoretic approach to multirobot mapping and exploration. In: Proc. of the IEEE-RSJ Int. Conf. on Intelligent Robotics and Systems, pp. 3232–3238 (2003)
26. Stachniss, C., Grisetti, G., Burgard, W.: Analyzing Gaussian proposal distributions for mapping with Rao-Blackwellized Particle Filters. In: Proc. of Int. Conf. on Intelligent Robots and Systems, pp. 3485–3490 (2007)
27. Arulampalam, S., Maskell, S., Gordon, N., Clapp, T.: A tutorial on particle filter for on-line nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. Signal Process. **2**(50), 174–188 (2002)
28. Fox, D.: Adapting the sample size in particle filters through KLD-sampling. Int. J. Robot. Res. **22**(12), 985–1003 (2003)
29. Carpin, S.: Fast and accurate map merging for multi-robot systems. Auton. Robot. **25**(3), 305–316 (2008)
30. Howard, A.: Multi-robot simultaneous localization and mapping using particle filters. In: Proc. of Robotics: Science and Systems, pp. 201–208 (2006)
31. Roy, N., Burgard, W., Fox, D., Thrun, S.: Coastal navigation: mobile robot navigation with uncertainty in dynamic environments. In: Proc. of Int. Conf. on Robotics and Automation, pp. 35–40 (1999)
32. MobileRobots Inc.: MobileSim-the mobile robots simulator. http://robots.mobilerobots.com/MobileSim (2011). Accessed 10 Apr 2013
33. Robotics Research Group (RRG): Politecnico di Torino. http://www.polito.it/LabRob (2011). Accessed 10 Apr 2013
34. Kümmerle, R., Steder, B., Dornhege, C., Ruhnke, M., Grisetti, G., Stachniss, C., Kleiner, A.: Slam benchmarking webpage. http://ais.informatik.uni-freiburg.de/slamevaluation (2009). Accessed 10 Apr 2013