Science-Aware Exploration Using Entropy-Based Planning

Shivam Gautam, Bishwamoy Sinha Roy, Alberto Candela, David Wettergreen

Abstract— **Efficient exploration of unknown terrains by extraterrestrial rovers requires the development of strategies that reduce the entropy in the geological classification of a given terrain. Without such intelligent strategies, teleoperation of the rover is reliant either on human intuition or on the exhaustive exploration of the entire terrain. This paper highlights the use of low-resolution reconnaissance using satellite imagery to generate plans for rovers that reduce the overall uncertainty in the various geological classes. This becomes pivotal when exploration to collect diverse samples is resource constrained through exploration budgets and transmission bandwidths. We put forward two major contributions- a science-aware planner that uses information gain and a novel method of estimating this information gain. We propose an exploration strategy, based on the Multi-Heuristic A*, to solve the trade-off between optimizing path lengths and geological exploration through Pareto-optimal solutions. We show that our algorithm, which explicitly uses projected entropy-reduction in planning, significantly outperforms science-agnostic approaches and other science-aware strategies like greedy best-first searches. We further propose a feature-space based entropy formulation in contrast to the frequently used differential entropy formulation and show superior results when reconstructing the unsampled data from the set of sampled points.**

I. INTRODUCTION

Space exploration has been aided greatly by the development of extraterrestrial rovers which employ domain dependent task planners. One of the specific domain specific tasks includes spectroscopic analysis of geological samples. The frontiers of such exploration include not only other planets like Mars but also terrestrial environments like the Atacama desert [1], [2]. Spectral analysis of the red planet has been carried out in the past [3], [4] using high-resolution equipment like the Alpha Particle X-ray Spectrometer and the Sample Analysis at Mars (SAM) instrument suite [5]. The advancement in scientific payload allows rovers to perform scientific exploration efficiently and reliably over larger distances. However, low bandwidth and high latency lead to large feedback times during operation which proves to be a bottleneck in operation and planning. Shared autonomy provides a way to mitigate this by communicating higher-level objectives to the rover to execute autonomously through delegation of low-level controls. However, planning for high-level objectives is challenging due to the competing nature of the objectives for the exploration itself.

This work aims at increasing the effectiveness of the collaboration between robots and researchers by proposing a new science-aware approach to exploration. The proposed approach utilizes the low-resolution reconnaissance data to measure the information gain by visiting a point on the map and then generating budget-constrained plans which maximize the information gained. This low-resolution data is relatively cheaper to obtain through interplanetary communications when employing extraterrestrial rovers to sample high-resolution spectroscopic data.

All authors are with the Robotics Institute, Carnegie Mellon University.
{sgautam, bsinharo, albertoc, dw0s}@andrew.cmu.edu

We pose the problem of exploring a space as a multi-objective optimization problem wherein we try to maximize the information gained through map traversal whilst minimizing the distance traveled to collect that information. These competing objectives often force researchers to use heuristics and intuition to guide the rover to an overall global objective. We propose a structured way to resolve this ambiguity using a Pareto-optimal solution obtained as a result of implementing an evolutionary multi-objective algorithm [6]. We further present our results: the effectiveness of computing information gain in feature-space versus calculating differential entropy [7]. Differences in these approaches are pivotal for science-based approaches and this is discussed in detail in section IV.

To solve the problem of generating science-aware paths, we incorporate the projected information gain of a location explicitly in a search-based planning paradigm using the Multi-Heuristic A* (MHA*) [8]. This projected information gain is used as an additional inadmissible heuristic guided by the Euclidean heuristic which anchors the search towards the global goal. Paths generated henceforth are bounded in length due to the proven sub-optimality bounds for the MHA*. The quality of the paths is quantifiable using the reconstruction error of the sampling points when reconstructing the ground truth. Additionally, this approach scales well with the larger sizes of maps than greedy or direct approaches. Paths obtained from the Multi-Heuristic approach can be seen in Figure 1.

II. RELATED WORK

Optimization techniques based on reconstruction error and differential entropy have been well researched and analyzed in the past [9]. However, graph-based search techniques, which have proven optimality guarantees, are often overlooked in information-based path planning. Nevertheless, path planning for maximizing mutual information [10] has been explored in underwater exploration which testifies a strong need to develop such techniques for extra-terrestrial exploration. Multi-robot exploration provides faster coverage and robustness to failures but deploying such a system proves to be prohibitive in extra-terrestrial environments [11][12]. Additionally, while there has been much investigation into evaluating information gain using the differential entropy

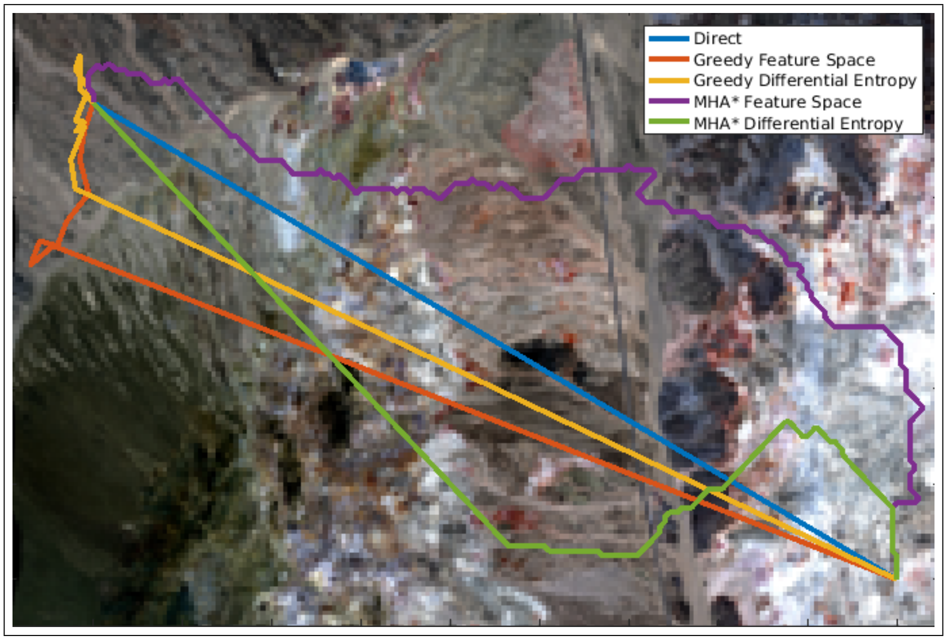


Fig. 1: Planned paths using the greedy, direct and MHA* approaches on spectroscopic data from Cuprite, Nevada. The start position is on the upper left and the goal position is at the bottom right of the map. Whilst the greedy approach visits high-information areas initially, it fails to explore the area near the goal due to path constraints. The direct approach, being science-blind, yields poor information returns. The science-aware planning algorithm, MHA* with information gain as an inadmissible heuristic, deviates from the canonical straight line path to maximize information gain.

of the sample set, there exist few approaches that estimate information gain directly in the feature space of images. Our search-based Multi-Heuristic A* approach using information gain in the feature space as the inadmissible heuristic underpins our novel approach to solving the problem of rover exploration.

Solving the problem of adaptive exploration requires multiple steps. First, a classification algorithm is run on the key features extracted from the rover and orbital imagery to train a model. Once a model is learned, sampling reward is predicted for each feasible sampling point. Following that, a path is planned through the terrain using the sampling reward as a metric. Finally, sampling along the path can be used to adaptively improve the model on the planned transect. A wide body of Artificial Intelligence research exists in each of these activities.

Like most classification problems, features are extracted from imagery collected either through orbital flybys or through rover traversal. Thompson et al. extract rock-classes from the rover imagery using a modified Viola/Jones rock classifier [13], which can then be used as features for terrain classification in a spatial model [14]. Classification in the pixel space, using robust machine learning techniques like random forests, yields much more precise classification [15]. Classification methods in pixel-space first detect *superpixels* and later defining a similarity graph over the superpixels by applying a graph cut algorithm [16], [17]. Working in the pixel-space is at risk to the *curse of dimensionality* as not all wavelength bands in a spectroscopic image are of importance. As such, PCA-based dimensionality reduction [18] can be applied prior to classification, thereby reducing computation requirements.

Once classified into geologically relevant classes, the reward calculation proceeds by estimating the gain from each sampling opportunity. Both Foil and Thompson consider iteratively modeling the sample space to assess the reward of future samples [19], [13]. Additionally, Thompson et al. provide a spatial model that predicts future readings based on previous observations using a Gaussian process reliant on information gain based sampling. Foil et al. models the sample space using modifications to Gaussian mixture model and relies on different metrics to calculate rewards. For instance, [19] uses an adaptive Gaussian mixture model in sample space to assign rewards to points where sampling is feasible. Most methods in exploration robotics assign an entropy or information gain metric to every exploration opportunity and define an objective function which usually serves to maximize the total information gain. Several approaches exist to delineate transects using the estimated reward of a location. Such planning involves formulating an objective function and optimizing it with distance constraints. Maximum-Entropy Sampling (MES) strategies select the next sampling point based on the sole criterion of maximizing the entropy within the set of sampled points. Such a method is therefore restricted by selecting when a single sample is being selected. However, the problem at hand needs to select a subset of all available sampling opportunities. Spatial design is applied by Thompson et al., which consists of selecting samples that maximize the differential entropy of the sample space [14].

On the other hand, instead of optimizing over the differential entropy, Thompson et al. optimizes the reconstruction error and analyzes greedy Least Squares and non-negative Least Squares optimization techniques. For optimization to work, a path needs to be evaluated. Therefore, it becomes necessary to select sampling points within a transect, which can then provide the quality of the path. Bayesian experimental design techniques and optimal foraging based on multi-armed bandits are applied by Matikainen et al. for sampling opportunity selection on a given transect [20].

III. Problem Definition

The problem of exploring unknown geological terrains draws equivalence to the problem of mapping an area. Maps encapsulate diverse information ranging from terrain characteristics to geological properties. During mapping, the agent creates a comprehensive representation of the environment by visiting locations in the environment and taking samples of the environment [13] which in effect allows for comparison to the *Travelling Salesman Problem* [21].

The specific problem of budgeted adaptive exploration is that of planning for maximizing information gain given an orbital spectroscopic image of the terrain and a sampling budget. Planning is constrained by the number of samples the rover can take as well as the total distance it can travel. We assume that the there exists a global goal location where the rover intends to be at the end of the planning cycle. This goal location is generally derived from higher level exploration goals like visiting specific sites. The problem, therefore, reduces to generating plans that deviate from the shortest path to this global goal at the cost of collecting more information about the geology of the terrain. The sampling budget is imposed as a hard constraint and is affected by the time taken to sample and transmit to the ground station. This sampling budget represents the constraints imposed by time, bandwidth and power requirements [14].

The problem is motivated by the fact that spatial distribution of minerals is pivotal in reasoning about the geological history of the region and serves as a precursor to planning terrain-aware paths for advancing further exploration. Moreover, a geological map is a more compact representation of visual or spectroscopic data, which is required for efficient data transfer over constrained high-bandwidth links between planetary missions and the Earth [19].

One of the most widely used sensors for geological classification are spectrometers, which measure the reflectance in different wavelength bands. While orbital flybys can provide an overview of a large region, they are limited by their limited resolution. Atmospheric effects and inherent measurement noise further lead to this spectral data being noisy. Instead of relying on this data for accurate classification, this low-resolution data when used in conjunction with high-resolution spectrometric observation from rovers leads to the formulation of better sampling strategies. Rovers, equipped with more accurate spectrometers, can be commanded on the ground to take more informative measurements, guiding their sampling based on the intelligence received from orbital flybys.In contrast, the complete teleoperation of the rover to guide sampling based on human intuition allows greater customization for visiting areas of interest. However, such an unstructured approach generally yields poorer results.

IV. Approach

A. Multi-Heuristic A*

Search-based planners form a potent method to explore the state space from a start location to a goal location, using heuristics to guide the search. Heuristics traditionally have been used to better guide the search towards the goal configuration. However, innovative heuristics, such as those which use cached plans [22] to estimate *cost-to-goal* have been used. Until recently, a common way to combine multiple admissible heuristics would be to take their maximum, i.e, for two admissible heuristics H_1 and H_2 at state s:

$$H(s) = \max\left(H_1(s), H_2(s)\right). \tag{1}$$

This combination, however, dilutes the effectiveness of an individual heuristic in situations where it could be more informative. For example, planning in constrained spaces, a heuristic that measures the distance to obstacles could be more informative than simple Euclidean distance. Combining multiple heuristics can be effectively done through the Multi-Heuristic A* (MHA*) algorithm [8], even if one of the heuristics is inadmissible. Heuristically guided search based techniques are optimal at a resolution and weighted search techniques ensure bounded sub-optimality. Despite incorporating inadmissible heuristics, the MHA* algorithm has bounded sub-optimality guarantees. Proofs for bounded sub-optimality and guarantees on state expansions for the MHA* can be found in [23]. One of the main theorems for the algorithm provides sub-optimal bounds on the path cost. Given that the weight on the anchor heuristic is w and the weight on n arbitrary inadmissible heuristics is w_1, $w_2 ... w_n$, it is given that

$$g(s_{goal}) \leq w * w_1 * w_2 * ... w_n * g^*(s_{goal}) \tag{2}$$

where g^* represents the optimal g-value. This in effect ensures that the solution is bounded by the sub-optimality factor of:

$$w * w_1 * w_2 * .... w_n. \tag{3}$$

Due to the guarantees and a well-formulated way to combine inadmissible heuristics, we use the MHA* as the search based technique. The euclidean distance to goal is the consistent and admissible heuristic. Information gain is an inadmissible heuristic which has no guarantees on the optimality of the path. Information gain is the only inadmissible heuristic, with a weight α, which reduces the sub-optimality equation to be bounded by

$$w * \alpha. \tag{4}$$

Information gain, despite being inconsistent and inadmissible, can guide the search into high areas of information, while the Euclidean distance to the goal anchors the search to the goal. As such, Euclidean distance was used as the anchor heuristic, and the inverse information gain was used as the inadmissible heuristic to model it as a cost.

The state space for the graph search is defined to be Euclidean and the edge costs are modeled as such. Explicitly including the uncertainty reduction in the state space and formulating this composite cost function, which encapsulates the notion of a short path and that of maximizing information gain, is challenging. Such a cost function can possibly be learned by learning the cost function through expert demonstration [24], [25]. However, learning from expert demonstrations does not lend well to the problem of geological exploration due to their scarcity in a budgeted domain. As a result, these two inherently competing objectives lend the problem to be of finding Pareto-optimal solutions.

Once the heuristics have been designed, parameters w and α are usually tuned to achieve the desired behavior in the given environment. Instead of hand-tuning these parameters, we implement an evolutionary algorithm based Pareto optimization. This is essential in a budgeted domain where scientists are constrained by the distance budgets and try to maximize the information gain. For the sake of not sacrificing optimality in Euclidean space, we set w to 1 and optimize the weight for the information gain, α.

B. Evolutionary algorithm based Pareto-optimization

The optimization problem can be formulated as maximizing information gain while minimizing the cost incurred while traversing the path. In its most primitive form, the path cost can be modeled simply as the length of the path traversed. These two competing objectives form an arbitrary optimization manifold where gradient-based optimization methods are generally slower and are prone to local minima. Exploring such arbitrary decision surfaces is a problem well suited for evolutionary algorithms such as the Multi-Objective Genetic Algorithm (MOGA) and the Non-dominated Sorting Genetic Algorithm-II (NSGA-II) [26].

We use the NSGA-II algorithm as it scales better with more decision points and preserves diversity in the population when compared to the MOGA. We optimize on the design parameter α, the weight given to the information-gain heuristic during planning. Intuitively, this also relates to the relation between the importance given to following the shortest path as compared to exploring areas with more information gain. The Pareto-optimizer framework we follow is depicted in Algorithm 1.

An example of a Pareto-optimal solution is given in Figure 2. Points in the Pareto frontier are non-dominated, i.e. improving any one objective is always at the cost of making the other objective worse. For budgeted exploration, the problem of picking a plan for the rover to execute reduces to picking a parameter from this optimization that satisfies the plan budget while simultaneously maximizing the information gain.

C. Information Gain

Information theory formalizes entropy as the expected information in a random variable. Discrete random variables entropy can be calculated using Equation 5, while continuous variables entropy can be calculated using Equation 6.

$$H(x) = -\sum_{x \in X} P(x) \log P(x) \tag{5}$$

$$H(x) = -\int_{x \in X} P(x) \log P(x) \tag{6}$$

For stationary distributions, optimal experimental design can be applied to select the sampling point that maximizes the entropy of the sample space [13], [14]. Spatial design is an experimental design technique that chooses the next sampling point in addition to a subset of all sampling opportunities. Solving for the entropy of the sample space yields the following entropy formulation:

$$H(x) = -\frac{1}{2} \sum_{b \in B} log(2\pi\sigma_b), \tag{7}$$

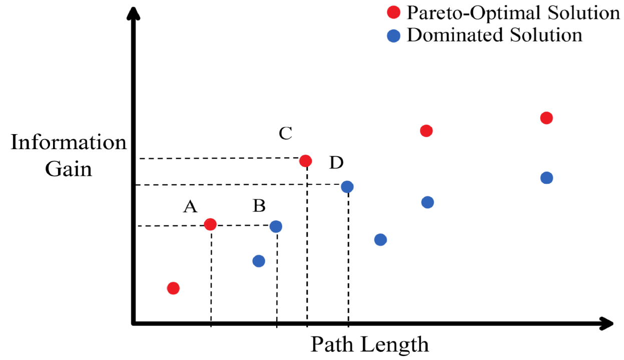


Fig. 2: Concept of state dominance in the space of rover exploration. Point A dominates point B as the length of the path to gain the same amount of information is greater for point B. Point C dominates point D as the amount of information gained is more for a shorter path length as compared to point D.

Algorithm 1 Pareto Optimizer

procedure PARETOOPTIMIZER
Input: *PopSize, InitAlpha, MHAplanner, MaxGenerations*
Output: *ParetoFrontier*
FitnessFunc ← *MHAplanner*
CurrentPopulation ← *InitAlpha*
CurrentPopulation ← *CurrentPopulation* + *RandomPopulationGenerator((PopSize-1)*
while $NumGeneration \leq MaxGeneration$ **do**
Offsprings ← *Crossover(CurrentPopulation)*
Offsprings ← *Mutation(Offspring)*
EvaluablePop ← *CurrentPopulation* + *OffSpring*
Fitnesses ← *FitnessFunc(EvaluablePop)*
CurrentPopulation ← *BestPopulation(CurrentPopulation,Fitnesses)*
return *NonDominatedPopulation (CurrentPopulation)*

which is the differential entropy for a Gaussian random variable. As demonstrated in [19], if the sample space is a mixture of Gaussians, it can be shown that the differential entropy for Gaussian random variables can be expressed as Equation 7. Spatial design is then effectively an entropy calculation of a Gaussian continuous random variable. A scheme that samples points based on differential entropy maximizes the variance in the set of sampled points, σ_b. This is an effective measure assuming that the greater variation within the samples that the rover carries, the more information about the map is gained.

Traditionally, exploration robotics calculates entropy in sample space using the aforementioned theoretical backing. However, entropy can also be calculated in image space, where each channel in the image is a feature. We present this new approach of calculating entropy in the feature space as a way of minimizing the reconstruction error. As such, an RGB image will have 3 dimensions, and a n-wavelength band spectral image will have an n-dimensional feature space, where each pixel will be a single point and image will form a point cloud in feature space, a map M. It is quite intuitive that a spectral image will form clusters in the feature space, which can then be used to formulate discrete entropy based exploration strategy.

The objective is to minimize $H\left(M|X\right)$, which is the entropy of the complete map given the sampled points. The assumption is that choosing points according to Equation 8 is the optimized way to minimize $H\left(M|X\right)$. Note that each pixel in the spectroscopic image is considered a random variable and the whole image/map is also considered a random variable. This effectively means choosing the point that has the maximum entropy given the previously collected samples X and the satellite-provided spectroscopic map M of the region with high uncertainty/noise.

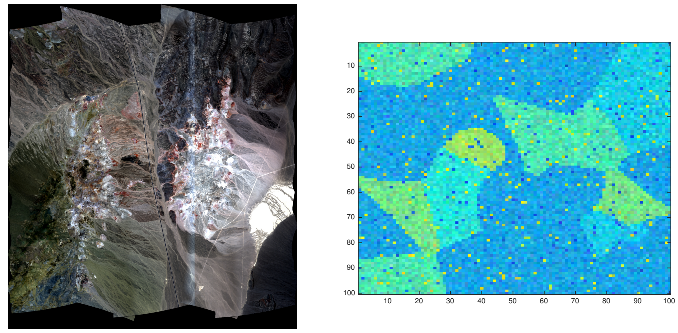
$$x_{next} = \underset{x \in M - X}{\operatorname{argmax}} H(x|M, X \cap x) \tag{8}$$

To calculate the entropy of a point, k-means clustering is used to attain C clusters. The covariance of each cluster, assuming Gaussian densities, can be approximated from the sample covariance from cluster members. Using the cluster center as the mean and the calculated covariance, the probability of a map point being sampled from the cluster can then be computed. Shannon discrete entropy can then be applied to work out the entropy of the map point as shown in Equation 9.

$$p_x^{c_i} = \frac{1}{(2\pi)^f |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2}||x-\mu_{c_i}||^2_{\Sigma_i}} \tag{9}$$

Inherently, this formulation pushes the rover to sample points that make the clusters tighter. This can be attributed to the points of maximum entropy not being classified with high confidence in any of the clusters due to factors such as sensor noise. The underlying assumption is that the data in spectroscopic feature space forms a mixture of Gaussians.

To ensure a comprehensive comparison, we implement both entropy formulations. While the differential entropy prioritizes variance of samples, the feature-space entropy prioritizes samples with low clustering confidence.



(a) Data obtained from the AVIRIS-NG spectroscopic sensor.

(b) Generated simulated data.

Fig. 3: Datasets encapsulating the geological environment as seen by the orbiting satellites.

V. Experimental Results

A. Experimental Setup

The experimental setup involved using simulated and real satellite data to compare different exploration algorithms. This ensured that the algorithms were comparable not only in real world data, but the claims would hold to much wider and complex scenarios that could be recreated in simulation.

We use high-resolution spectroscopic data collected from the Airborne Visible Infrared Imaging Spectrometer Next Generation (AVIRIS-NG) [27], [28]. The AVIRIS-NG sensor mapped the Cuprite mining district of Nevada at high spatial resolution (3.9 m per pixel) with radiance measurements from 380-2510 nm [29]. The data were acquired during overflights in 2014 and converted from measured at-sensor radiance to surface reflectance using the procedure described in [30]. The Cuprite mining dataset and our simulated dataset can be seen in Figure 3. The dataset contains 20-channel reflectance measurements which we use in its entirety to compute information content and generate plans. The simulated data consisted of 10 random cluster centers, each center representing a spectral class. To account for the sensor noise for the rover and the orbiter, a zero-mean Gaussian noise was added to the ground truth data. The data available to the orbiter is of much higher variance than that available to the rover. This is consistent with available real-world data where the rover carries much more sophisticated sampling equipment with low sensor noise than orbiters.

The metric we use to compare planning approaches is the mean reconstruction error as explained in [19]. This metric quantifies how well the sampled points, X, can recreate the unsampled data, Y, as shown in Equation 10.

$$\begin{gathered} W^* = \underset{W}{\operatorname{argmin}} ||W^T X - Y||_2 \\ w_i > 0, \forall w_i \in W \end{gathered} \tag{10}$$

The Pareto-optimal frontier was generated using a fixed start and end point for a fixed planning environment. Comparison between different planning approaches was performed over a set of X maps, each with 100 randomized starting and ending locations.

B. Results

The Pareto-optimal frontier, generated using the multi-objective genetic algorithm approach, can be seen in Figure 4. Each point in the frontier depicts a planning configuration with a fixed α. Each point in the frontier is non-dominated by any other point in the frontier where state dominance is defined in terms of lower information gain whilst having a higher path cost. Instead of optimizing for a single objective like the reconstruction error [14], our genetic algorithm based multi-objective optimization optimizes for both the path length and the information gain simultaneously. The non-dominated frontier provides decision points for estimating the entropy reduction given a constrained path length which provides a structured way in reducing ambiguity in decision making.

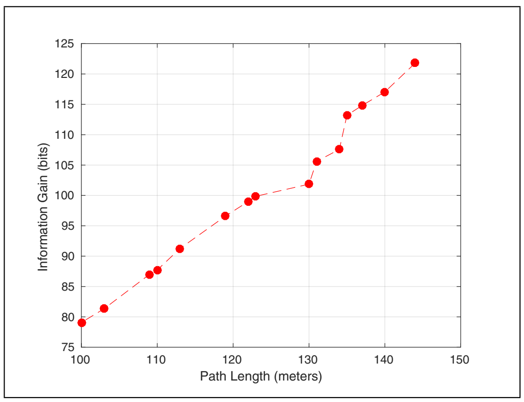
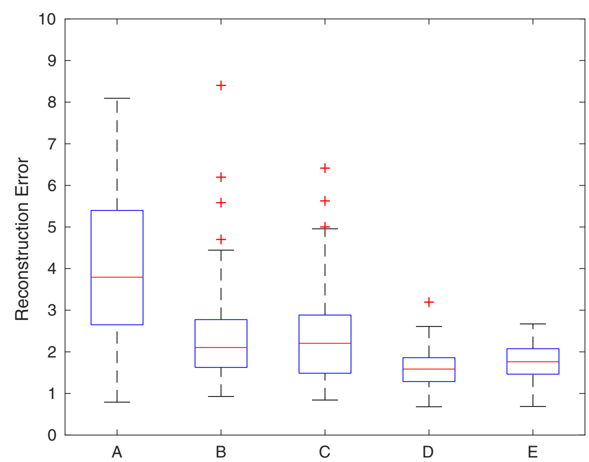


Fig. 4: The Pareto frontier for the multi-objective optimization. Each point in the frontier represents a trajectory generated for a Pareto-optimal value of alpha.

We compared our approach of using the MHA* with other science-aware and science-agnostic approaches. The direct planner is a science-agnostic approach to exploration which simply tries to plan a path between the start and the goal using classical planning algorithms. We use the A* algorithm as a baseline science-agnostic approach due to its strong guarantees on optimality. This comparison provides a baseline for any science-aware method as a method which explicitly takes geological information into account while planning should not be any worse than a planner that is agnostic to this.

Further, to compare to a science-aware approach, we compare the results obtained by a short-sighted greedy planner. The greedy planner simulates a limited-horizon best-first search based on the information gain. The greedy planner samples at the highest entropy point in a window around its current location. Once the planner has deviated from the goal enough to just satisfy the path constraint, it plans a direct path to the goal. This closely approximates a rover being teleoperated to the nearest waypoint of interest and being directed to the goal once the exploration budget has been exhausted. Such a planner serves as a baseline when comparing science-aware approaches.

Whilst, it might seem ideal to compare to a planner which uses dynamic programming to recursively compute the most informative path from the goal, such an approach is not suitable when operating with practical applications. Running a dynamic programming planner would result in an unbounded computation if not restricted to a map and even in the bounded case, the planning resources in terms of time and memory are exponential in the number of states. We implemented a simple dynamic programming based approach to solving a 100x100 grid which took nearly 5 hours to solve on a Core-i7 machine, thereby proving its severe limitation in scaling to real-world scenarios.



(a) Mean and variance plots for the simulated dataset

Fig. 5: Comparison for planning approaches based on their mean reconstruction error. (A) Direct, (B) Greedy with Differential Entropy, (C) Greedy with Feature-Space Entropy, (D) MHA* with Differential Entropy, (E) MHA* with Feature-Space Entropy.

The results for our comparison are depicted in table I. The science-aware greedy and MHA* methods are able to reconstruct the unsampled points better, proving that using differential and feature-space entropies allow the rover to collect more meaningful samples. Within the information metrics, we see superior performance for the MHA* when coupled with our proposed feature space formulation. Even for the case when differential entropy is used, the MHA* planner performs much better than both the greedy and direct approach. The differences are more glaring in the simulated data but are also reflected in the Cuprite dataset, albeit on a smaller scale, as shown in Figure 5. As a consequence of exploration, the path cost for the MHA* planner lies between that of greedy and direct approaches. Additionally, the planning times for the MHA* are comparable to other approaches and are not egregiously prohibitive like dynamic programming. Additionally, the type of paths provided by

TABLE I: Reconstruction error for planning approaches using different information metrics and datasets.

Algorithm	Information Metric	Dataset	Mean Reconstruction Error	Mean Path Length (m)	Mean Planning Time (s)
Direct	—	Cuprite	1.0865	166.16	0.0026
Direct	—	Simulated	8.3016	68.05	0.0024
Greedy	Differential	Cuprite	1.133	295.38	0.1073
Greedy	Differential	Simulated	2.3150	651.57	0.0990
Greedy	Feature-Space	Cuprite	0.975	287.82	1.262
Greedy	Feature-Space	Simulated	2.416	649.23	0.749
MHA*	Differential	Cuprite	1.076	165.231	11.39
MHA*	**Differential**	**Simulated**	**1.6015**	91.2819	4.1843
MHA*	**Feature-Space**	**Cuprite**	**0.9428**	170.891	4.411
MHA*	Feature-Space	Simulated	1.7623	76.7723	0.4375

each planner is also shown. As can be predicted, the direct planner takes the rover directly to the goal position. On the other hand, the greedy planner deviates from this direct path till a budget runs out and then follows a direct path to the goal. This isn't a good approach because it uses up it's budget pretty much near the start position. The MHA*, however, provides a goal oriented path that widens the exploration horizon as it doesn't use it's budget near the start-location. Paths generated using the various planning approaches can be visualized on the Cuprite dataset in Figure 1.

VI. Conclusion and Future Work

This paper describes a science-aware search-based planning approach using information gain as a heuristic. We propose measuring information gain in the feature-space in addition to conventional differential entropy approaches. We present a structured approach to resolving the constrained exploration problem by optimizing for the information and the path length in a constrained multi-objective optimization. This results in supplementing decision making for rover exploration with a Pareto-optimal decision frontier as compared to a single decision point. Additionally, we demonstrate the effectiveness of using the MHA* with information as an inadmissible heuristic when compared when compared to greedy and science-blind approaches and show how our approach can better reconstruct the entire map through the points in its set of sampled points in the trajectory.

After proving the effectiveness of our algorithm against non-science and greedy methods, we wish to draw comparisons to other science-based methods in the future. Additionally, having tackled the problem of a budgeted path, we wish to investigate developing a sampling strategy on the planned path to optimize the information gain whilst being confined to a sampling and path budget. Further, we aim to use our sampling strategy to develop efficient re-planning strategies that best use the dynamic information gained through the sampling strategy.

References

[1] D. Wettergreen *et al.*, "Robotic technologies for surveying habitats and seeking evidence of life: Results from the 2004 field experiments of the" life in the atacama" project," in *36th Annual Lunar and Planetary Science Conf.*, vol. 36, 2005.

[2] ——, "Long-distance autonomous survey and mapping in the robotic investigation of life in the atacama desert," *Int. Symposium on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS)*.

[3] E. Stolper *et al.*, "The petrochemistry of Jake_M: a Martian mugearite," *Science*, vol. 341, no. 6153, p. 1239463, 2013.

[4] R. Rieder *et al.*, "Chemistry of rocks and soils at meridiani planum from the alpha particle x-ray spectrometer," *Science*, vol. 306, no. 5702, pp. 1746–1749, 2004.

[5] P. Mahaffy *et al.*, "Sample analysis at mars (sam) instrument suite for the 2011 mars science laboratory," in *Lunar and Planetary Science Conf.*, vol. 40, 2009, p. 1088.

[6] K. Deb *et al.*, "Multi-objective optimization," in *Decision Sciences: Theory and Practice*. CRC Press, 2016, pp. 145–184.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, New Jersey: Wiley-Interscience, 2006.

[8] S. Aine *et al.*, "Multi-heuristic A*," *The Int. J. Robotics Research*, vol. 35, no. 1-3, pp. 224–243, 2016.

[9] D. Thompson *et al.*, "Autonomous science during large-scale robotic survey," *J. Field Robotics*, vol. 28, no. 4, pp. 542–564, 2011.

[10] J. Binney *et al.*, "Informative path planning for an autonomous underwater vehicle," in *Proc. IEEE ICRA*, 2010, pp. 4791–4796.

[11] K. H. Low *et al.*, "Adaptive multi-robot wide-area exploration and mapping," in *Proc. 7th Int. Joint Conf. Autonomous Agents and Multiagent Systems - Volume 1*, 2008, pp. 23–30.

[12] A. Singh *et al.*, "Efficient planning of informative paths for multiple robots," vol. 7, 2007, pp. 2204–2211.

[13] D. R. Thompson, "Intelligent mapping for autonomous robotic survey," Ph.D. dissertation, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August 2008.

[14] D. R. Thompson *et al.*, "Spatio-spectral exploration combining in situ and remote measurements," in *AAAI*, 2015, pp. 3679–3685.

[15] G. Foil *et al.*, "Probabilistic surface classification for rover instrument targeting," in *Proc. IEEE/RSJ IROS*, 2013, pp. 775–782.

[16] R. Achanta *et al.*, "Slic superpixels compared to state-of-the-art superpixel methods," in *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, 2012, pp. 2274–2282.

[17] M. Ester *et al.*, "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining*, ser. KDD'96. AAAI Press, 1996, pp. 226–231.

[18] S. Wold *et al.*, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[19] G. Foil, "Efficiently sampling from underlying physical models," Ph.D. dissertation, Robotics Institute, Carnegie Mellon University, October 2016.

[20] P. K. Matikainen *et al.*, "Multi-armed recommendation bandits for selecting state machine policies for robotic systems," in *Proc. IEEE ICRA*, 2013, pp. 4545–4551.

[21] M. Dorigo and L. M. Gambardella, "Ant colonies for the travelling salesman problem," *Biosystems*, vol. 43, no. 2, pp. 73–81, 1997.

[22] M. Phillips *et al.*, "E-graphs: Bootstrapping planning with experience graphs," in *Robotics: Science and Systems*, vol. 5, no. 1, 2012.

[23] S. Aine *et al.*, "Mha*: The proofs," Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-13-18, 2013.

[24] B. D. Ziebart *et al.*, "Maximum entropy inverse reinforcement learning." in *AAAI*, vol. 8, 2008, pp. 1433–1438.

[25] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proc. 21st Int. Conf. Machine Learning*, ser. ICML '04. ACM, 2004, pp. 1–.

[26] K. Deb *et al.*, "A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: Nsga-ii," in *Int. Conf. Parallel Problem Solving From Nature*. Springer, 2000, pp. 849–858.

[27] R. O. Green *et al.*, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris)," *Remote Sensing of Environment*, vol. 65, no. 3, pp. 227–248, 1998.

[28] L. Hamlin *et al.*, "Imaging spectrometer science measurements for terrestrial ecology: Aviris and new developments," in *Aerospace Conference*. IEEE, 2011, pp. 1–7.

[29] G. A. Swayze *et al.*, "Mapping advanced argillic alteration at cuprite, nevada, using imaging spectroscopy," *Economic Geology*, vol. 109, no. 5, pp. 1179–1221, 2014.

[30] D. R. Thompson *et al.*, "Atmospheric correction for global mapping spectroscopy: Atrem advances for the hyspiri preparatory campaign," *Remote Sensing of Environment*, vol. 167, pp. 64–77, 2015.