DCT Maps:
Compact Differentiable Lidar Maps Based on the Cosine Transform

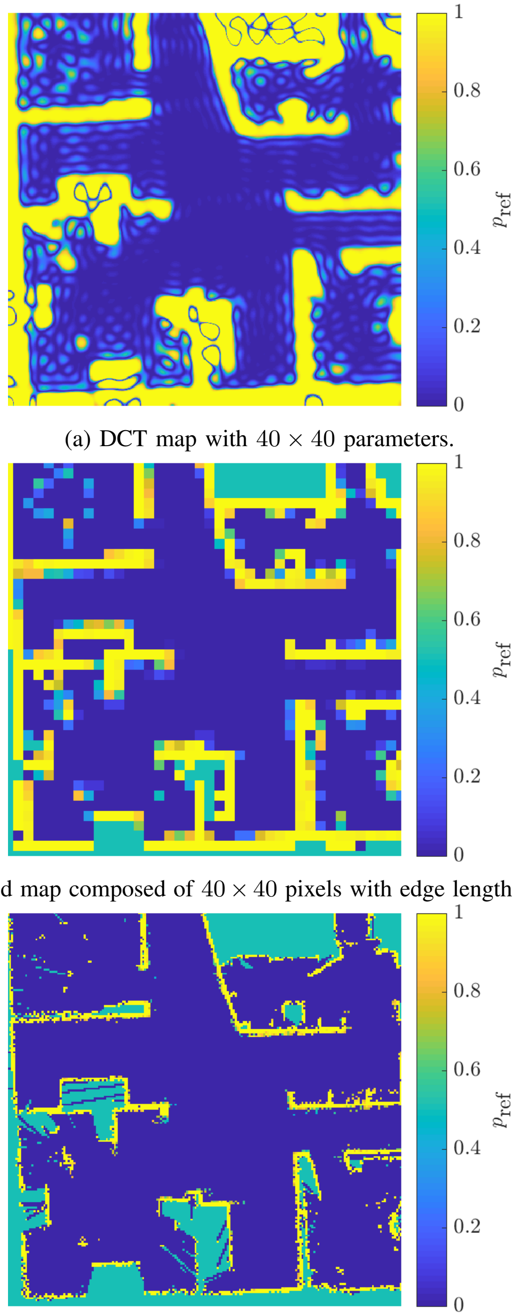
Alexander Schaefer, Lukas Luft, Wolfram Burgard

***Abstract*— Most robot mapping techniques for lidar sensors tessellate the environment into pixels or voxels and assume uniformity of the environment within them. Although intuitive, this representation entails disadvantages: The resulting grid maps exhibit aliasing effects and are not differentiable. In the present paper, we address these drawbacks by introducing a novel mapping technique that does neither rely on tessellation nor on the assumption of piecewise uniformity of the space, without increasing memory requirements. Instead of representing the map in the position domain, we store the map parameters in the discrete frequency domain and leverage the continuous extension of the inverse discrete cosine transform to convert them to a continuously differentiable scalar field in the position domain, which we call DCT map. A DCT map assigns to each point in space a lidar decay rate, which models the local permeability of the space for laser rays. In this way, the map can describe objects of different laser permeabilities, from completely opaque to completely transparent. DCT maps represent lidar measurements significantly more accurate than grid maps, Gaussian process occupancy maps, and Hilbert maps, all with the same memory requirements, as demonstrated in our real-world experiments.**

I. Introduction

Mapping and localization are at the heart of almost every mobile robotic system. In this context, lidar is a popular sensor modality, as lidar sensors produce relatively accurate, low-noise signals. Using these signals for mapping and localization requires an inverse and a forward sensor model. The inverse sensor model converts recorded measurements to a map. The forward model uses this map to assess the probability of incoming sensor readings given the robot pose. The maps produced by the inverse pass are often grid maps: They tessellate the physical space into square pixels or cubic voxels. Each pixel or voxel contains a value that is assumed to be constant within it. This value characterizes the statistical optical properties of the corresponding portion of space. Fig. 1b shows an example of such a grid map built from 2-D lidar scans recorded in an office environment.

Although tessellation is intuitive, grid maps bring with them several drawbacks. First, they can only coarsely approximate the true spatial distribution of the optical properties of interest. Aliasing effects occur whenever the optical characteristics of the environment change, as these transitions are never perfectly aligned with the raster of the grid. The grid map in fig. 1b exhibits the resulting characteristic staircase patterns. Although increasing the map resolution



(a) DCT map with 40×40 parameters.

(b) Grid map composed of 40×40 pixels with edge length $25\,\text{cm}$.

(c) Grid map composed of 200×200 pixels with edge length $5\,\text{cm}$.

Fig. 1: Decay rate maps of the same $10\,\text{m} \times 10\,\text{m}$ patch of the Intel Research Lab dataset [1] generated from the identical set of planar lidar measurements. The colors encode the reflection probability $p_{\text{ref}} := 1 - \exp(-\lambda)$, where λ denotes the local laser decay rate.

This work was partially supported by the European Commission in the Horizon 2020 framework program under grant agreement 644227-Flourish.

All authors are with the Department of Computer Science, University of Freiburg, Germany.
{aschaef,luft,burgard}@cs.uni-freiburg.de

can theoretically alleviate this problem (see fig. 1c), quadratic or cubic memory complexity quickly renders this approach prohibitive. Depending on the use case, non-cubic voxels may mitigate the errors induced by tessellation [2]. Second, grid maps are not continuously differentiable, although this is a desirable property of any map. Continuous differentiability would allow to localize the robot by maximizing the measurement likelihood over the robot poses, and even to perform SLAM by maximizing the measurement likelihood over the robot poses and the map parameters.

In the present paper, we choose a different approach to avoid the aforementioned detrimental effects without increasing the memory demands. Inspired by well-established digital image compression algorithms like JPEG, we store the map parameters in the discrete frequency domain and use the so-called continuous extension of the inverse discrete cosine transform [3] to obtain a continuously differentiable scalar field in the position domain. In addition to the regular inverse discrete cosine transform, its continuous extension not only computes the function values at discrete grid points in the spatial domain, but also closely approximates them in between. We combine this map representation with the recently developed decay rate model for lidar sensors [4]. The resulting DCT maps model the local permeability of the space for laser rays. Fig. 1a depicts such a DCT map. It was built from the identical information as the grid map in fig. 1b and has the same memory footprint, but it does not exhibit staircase patterns and better preserves the map contours. Indeed, our experiments show that DCT maps represent lidar data with higher accuracy than other approaches. Moreover, the continuous derivatives of DCT maps can be calculated in closed form, a fact that enables optimization-based SLAM.

In the following, we first survey different map representations. Then, we describe the mathematics behind DCT maps in detail. Finally, the findings of experiments conducted with publicly available real-world 2-D lidar datasets are presented.

II. Related Work

Occupancy grid maps [5] were among the first probabilistic map representations used in robotics and are still widely used today. They tessellate the space into independent cells and assign each cell the posterior probability of being occupied. Occupancy grid maps cannot model semi-transparent objects; they assume that each cell is either completely free or completely occupied. In contrast, the decay rate model, which we employ to formulate DCT maps, explicitly models the permeability of each cell for a laser ray. If used in conjunction with grid maps, it even allows to calculate posterior distributions over the decay rate values without additional computational effort [6].

While 2-D occupancy maps are able to model large areas, even moderately-sized 3-D occupancy grids quickly outgrow the memory limitations of modern computers. For this reason, several research projects target compressed map representations. Elevation maps [7] assume that the environment can be represented by a 2-D grid map whose cells contain not only occupancy values, but also one height coordinate per cell. To relax the assumption that the world is a single surface, [8] extends elevation maps to multi-level surface maps. Multi-volume occupancy grids [9] manage volumetric data as 2-D arrays, too, but in contrast to multi-level surface maps, each cell contains a list of occupied height regions and one of free height regions. Octrees [10] approach the memory limitation problem by hierarchically partitioning the space using an octal tree data structure. They have found broad application in robotics to model the spatial distribution of the occupancy value [11], [12], [13]. The authors of [14] present an octree-based data structure that is efficient to update and to copy, so it can be used in particle filter-based SLAM, where hundreds of maps must be maintained in parallel. To model the dynamics of the environment, [15] assumes that the occupancy values in an octree are subject to periodic changes. For each cell, the authors record the occupancy value over time and transform the resulting function to the frequency domain to predict the occupancy value at a later point in time. In this way, they achieve high compression ratios compared to storing one occupancy map per time step. Multi-resolution occupied voxel lists [16] differ from traditional occupancy mapping in that they store only the positions of the voxels that have been observed more frequently as occupied than as free. They are neither able to differentiate between unoccupied and unknown volumes, nor to account for semi-transparent voxels.

The normal distributions transform [17] was initially conceived in the context of scan matching. Based on this work, [18] introduces the so-called normal distributions transform occupancy map. Basically, this map is a grid map, but instead of a single scalar, every cell contains a normally distributed occupancy probability density function, which is cropped at the voxel bounds. In this way, it drops the assumption that the space is uniform within each voxel. As opposed to DCT maps, however, normal distributions occupancy maps achieve higher accuracy at the cost of increased memory consumption. Like all other occupancy-based approaches, they are not able to model semi-permeable objects, either. Normal distributions occupancy maps are extended and advanced in [19], [20], [21].

Other approaches completely abandon the notion of voxels. For example, [22] uses Haar wavelets to represent 3-D occupancy data. The authors of [23] drop the restriction that the elementary volumes of a map shall fill the space without gaps. Instead, they model the environment by non-overlapping spheres of equal sizes. In this way, they are able to more closely represent curved surfaces.

Point clouds are a simple and convenient way to represent lidar sensor data. However, in contrast to occupancy maps or decay rate maps, they are lossy in the sense that they store only the endpoints of the rays. They discard the ray path information of both reflected rays and rays that are not reflected. When point clouds are used for mapping, they accumulate memory for every incoming measurement, which limits their suitability for long-term navigation. Despite their drawbacks, many SLAM systems [24], [25], [26] represent lidar data in the form of point clouds.

In object reconstruction in computer graphics, objects are modeled as line segments in 2-D [27] or as polygon meshes in 3-D [28]. The resulting models can achieve an astonishing level of detail [29]. However, similar to mapping approaches based on implicit shape potentials like KinectFusion [30], they are not perfectly suited for lidar-based robot localization due to their sheer memory footprint and their inability to deal with semi-transparent materials.

Recently, machine learning techniques have completely relaxed the independence assumption between grid cells and produce continuous occupancy maps. Gaussian process occupancy maps (GPOM), for example, learn the environment of a robot and predict future states [31], [32]. Building on the latter, the authors of [33] present an incremental GPOM formulation that enables online mapping. Gaussian processes have also been applied to other map representations like implicit shape potentials [34]. Hilbert maps [35] are continuous occupancy maps built by projecting the lidar measurements in a Hilbert space, learning a logistic regression classifier, and then classifying each point in space as free or occupied.

III. Approach

In this section, we shortly revisit how the decay-rate model computes measurement probabilities conditioned on any kind of map, then we define the map using the continuous extension of the inverse discrete cosine transform. With these prerequisites, we derive the forward model, which computes the probability of a lidar measurement given the spectral parameters of the DCT map. In the last step, we formulate the inverse model as an optimization problem: We estimate the map parameters by maximizing the joint likelihood of all measurements collected during mapping.

For brevity and without loss of generality, the following derivation is performed for 2-D space. The derivation of the forward and inverse sensor model in 3-D exactly parallels the 2-D case.

A. The Decay Rate Model

The decay-rate model [4] models the probability that a lidar ray traverses a uniform medium as exponential decay process. The corresponding map assigns a decay rate to each point in space. This decay rate is a non-negative real number that describes the interaction between the laser ray and the environment completely.

To formulate the forward model mathematically, we introduce the following definitions. A lidar measurement $z := \{s, v, r\}$ describes a ray that originates at the sensor position s, travels in direction v, and ends after having traveled distance r. Assuming that the sensor provides its true position s, the true ray direction v, and that we are given a specific map $\mathcal{M}$, we model the non-deterministic interaction between the ray and the environment by the measurement probability density with respect to the radius

$$p(r) := p(r \mid s, v, \mathcal{M}). \tag{1}$$

Consequently, the absolute probability that the ray covers at least distance r is

$$\mathcal{N}(r) := 1 - \int_0^r p(r')\, dr'. \tag{2}$$

Alternatively, we can express equation (2) in form of the differential equation

$$p(r) = -\frac{d\mathcal{N}(r)}{dr}. \tag{3}$$

The essential idea of the decay rate model consists in the assumption that $\mathcal{N}(r)$ obeys an exponential decay process

$$\frac{d\mathcal{N}(r)}{dr} = -\lambda\left(r\right)\mathcal{N}\left(r\right), \tag{4}$$

where $\lambda(r)$ denotes the decay rate at a specific radius r along the ray. By combining this model assumption with differential equation (3), we obtain the measurement probability density

$$p(r) = \lambda(r)\,\mathcal{N}(r). \tag{5}$$

In (4) and (5), $\lambda(r)$ is obtained by evaluating the map $\lambda(x, y)$ along the trajectory of the ray.

The above formulation of the decay rate model is independent of any specific map representation. To use it as forward model in combination with DCT maps, we need to define the map function $\lambda(r)$ and solve the differential equation. In order to do so, we describe the spatial representation of DCT maps in the next section in detail. After that, we have all prerequisites at hand to solve the differential equation. The solution enables us to express the measurement probability of a lidar measurement given the map in closed form.

B. Transforming the Spectral Map Representation to the Spatial Domain

To avoid the disadvantages related to tessellation, DCT maps represent the map parameters in the discrete frequency domain instead of the position domain. Calculating the measurement likelihood from such a representation requires the definition of the transformation from the frequency domain to the spatial domain. We employ the continuous extension of the inverse discrete cosine transform (CEIDCT) [3]. Like other continuous extensions of Fourier-related transforms, it converts a discrete signal in the frequency domain to a continuous signal in the spatial domain. However, it differs from its relatives in that the continuous signal converges to the continuous function from which it was sampled for an increasing number of parameters (see [3], pp. 11–12). Moreover, its parameters are purely real-valued. For these reasons, it is particularly suited for our use case.

If we assume the spectral map parameters to be given by a matrix $\mathcal{A}$ with L rows and M columns, and if we denote the elements of $\mathcal{A}$ by a_{lm} with $l \in \{0, 1, \ldots, L-1\}$ and $m \in \{0, 1, \ldots, M-1\}$, the CEIDCT transforms them to a continuously differentiable decay rate map defined for each point (x, y) in the spatial domain

$$\lambda(x, y) = \left(\sum_{l=0}^{L-1}\sum_{m=0}^{M-1} a_{lm}\cos\left(l\tilde{x}\right)\cos\left(m\tilde{y}\right)\right)^2 \tag{6}$$

$$
\begin{aligned}
&= \left(\sum_{i=0}^{I-1} a_i \cos(l_i \tilde{x}) \cos(m_i \tilde{y}) \right)^2 \\
&= \sum_{i=0}^{I-1} \sum_{j=0}^{I-1} a_i \cos\left(l_i \tilde{x}\right) \cos\left(m_i \tilde{y}\right) \\
&\qquad a_j \cos\left(l_j \tilde{x}\right) \cos\left(m_j \tilde{y}\right) \\
&= \frac{1}{8} \sum_{i=0}^{I-1} \sum_{j=0}^{I-1} a_i \, a_j \sum_{\alpha \in Q} \sum_{\beta \in Q} \sum_{\gamma \in Q} \\
&\qquad \cos\Big((l_i + \alpha l_j) \tilde{x} + \beta (m_i + \gamma m_j) \tilde{y} \Big)
\end{aligned}
\tag{7}
$$

with $I := L\,M$ and $Q := \{-1, +1\}$. The tildes denote the π-normalization of the map coordinates: $\tilde{x} := \frac{\pi x}{X}$, $\tilde{y} := \frac{\pi y}{Y}$, where X and Y indicate the extent of the map. The variables l_i and m_i are the row and column indices into the matrix $\mathcal{A}$ that correspond to its i-th element a_i.

The original formulation of the CEIDCT does not square the double sum in (6). We employ this variant, however, because it ensures that the decay rate is non-negative for every point in the map. Negative decay rates would cause problems, as we cannot interpret the negative measurement probabilities in which they might result.

To solve equation (4), we still need to transition from $\lambda(x, y)$ to $\lambda(r) := \lambda(r, s, v)$. To that end, we apply the ray equation $[x, y]^\intercal = s + v\,r$ to (7) and obtain

$$
\begin{aligned}
\lambda(r) = \frac{1}{8} &\sum_{i=0}^{I-1} \sum_{j=0}^{I-1} a_i \, a_j \sum_{\alpha \in Q} \sum_{\beta \in Q} \sum_{\gamma \in Q} \\
&\cos\Big((l_i + \alpha l_j) \left[\tilde{s}_x + \tilde{v}_x r\right] \\
&\qquad + \beta (m_i + \gamma m_j) \left[\tilde{s}_y + \tilde{v}_y r\right] \Big).
\end{aligned}
\tag{8}
$$

C. Computing the Measurement Likelihood

Now we express the measurement probability of a lidar ray as a function of the measurement z and the spectral representation of the map $\mathcal{A}$ by solving the differential equation (4). The solution is

$$
\mathcal{N}(r) = \exp\left\{ -\, \mathcal{S}\left(s, v, r\right) \right\}
\tag{9}
$$

with

$$
\begin{aligned}
\mathcal{S}\left(s, v, r\right) &= \int_0^r \lambda(r') \, dr' \\
&= \frac{1}{8} \sum_{i=0}^{I-1} \sum_{j=0}^{I-1} a_i \, a_j \sum_{\alpha \in Q} \sum_{\beta \in Q} \sum_{\gamma \in Q} A_{ij}
\end{aligned}
$$

where

$$
\begin{aligned}
A_{ij} :=& A(i, j, \alpha, \beta, \gamma) \\
= &\begin{cases}
\frac{\left[\sin\left((l_i + \alpha l_j)\left[\tilde{s}_x + \tilde{v}_x r'\right] + \beta(m_i + \gamma m_j)\left[\tilde{s}_y + \tilde{v}_y r'\right]\right)\right]_0^r}{(l_i + \alpha l_j)\tilde{v}_x + \beta(m_i + \gamma m_j)\tilde{v}_y}, \\
\qquad \text{if } (l_i + \alpha l_j)\tilde{v}_x + \beta(m_i + \gamma m_j)\tilde{v}_y \neq 0 \\
\\
r \cos\left(\left(l_i + \alpha l_j\right) \tilde{s}_x + \beta \left(m_i + \gamma m_j\right) \tilde{s}_y\right), \\
\qquad \text{if } (l_i + \alpha l_j)\tilde{v}_x + \beta(m_i + \gamma m_j)\tilde{v}_y = 0
\end{cases}
\end{aligned}
\tag{10}
$$

Note that out of the infinite number of solutions to (4), we chose the one that satisfies the boundary condition $\mathcal{N}(0) \overset{!}{=} 1$.

By plugging equations (8) and (9) in (5), we finally obtain the closed-form solution of the measurement likelihood $p(r)$ for rays with real-valued radius r.

Not all lidar measurements are real-valued, though. In practice, the range of every lidar scanner is limited to a finite interval $[r_{\min}, r_{\max}]$. We call the rays reflected outside this interval no-return rays. In the following, we assume that the sensor identifies rays falling below $r_{\min}$ by the tag sub and rays that exceed $r_{\max}$ by the tag super. Consequently, the space of all possible measurements r is the mixed discrete-continuous set $\mathcal{D} := \{\mathrm{sub}, \mathrm{super}, r' : r' \in [r_{\min}, r_{\max}]\}$.

Fortunately, the decay-rate model easily accommodates both sorts of no-return rays:

$$
P(\mathrm{sub}) = \int_0^{r_{\min}} p(r) \, dr = 1 - \mathcal{N}(r_{\min}),
\tag{11}
$$

$$
P(\mathrm{super}) = \int_{r_{\max}}^{\infty} p(r) \, dr = \mathcal{N}(r_{\max}).
\tag{12}
$$

Supporting no-return rays is an important feature of the model. In a typical outdoor setting, no-return rays represent a considerable fraction of all measurements. If the model is unable to incorporate the information they convey, which is the case for the endpoint model, for example, it will inevitably loose accuracy.

During mapping and localization, one does not need to evaluate the measurement probability of a single ray, but of a whole laser scan $Z := \{z_k\}$ consisting of K rays, both with real-valued radius and without detected reflection. To obtain this probability, we first formulate the probability density function for each ray over the mixed space $\mathcal{D}$ by combining equations (5), (11), and (12) to

$$
p\left(z \mid \mathcal{M}\right) := \begin{cases}
P(\mathrm{sub}), & \text{if } r = \mathrm{sub} \\
p(r), & \text{if } r \in [r_{\min}, r_{\max}] \\
P(\mathrm{super}), & \text{if } r = \mathrm{super}
\end{cases}
$$

The above result, which we call a mixed probability density, is positive, real, and integrates to unity. Using the independence assumption, we then compute the joint probability density of all rays as the product

$$
p\left(Z \mid \mathcal{M}\right) = \prod_{k=0}^{K-1} p\left(z_k \mid \mathcal{M}\right).
$$

D. Building the Decay Rate Map

During the inverse pass, we want to determine the map parameters $\mathcal{A}$ that best explain the lidar measurements collected in the mapping run:

$$
\mathcal{A} = \underset{\mathcal{A}}{\operatorname{argmax}}\, p(Z \mid \mathcal{A}) = \underset{\mathcal{A}}{\operatorname{argmax}} \log\Big\{ p(Z \mid \mathcal{A}) \Big\}.
$$

This non-linear multivariate optimization problem can be solved by iterative computational optimization techniques like stochastic gradient descent or trust-region methods. These methods work considerably more reliable and faster when provided with first-order and second-order analytical logarithmic derivatives with respect to the spectral map parameters. To calculate the derivatives, we introduce the following shortcut notation:

$$\frac{\partial \lambda(x,y)}{\partial a_i} =: \sum_{j=0}^{I-1} a_j \, B_{ij} =: B_i,$$

with

$$B_{ij} := 2\cos(l_i\tilde{x})\cos(m_i\tilde{y})\cos(l_j\tilde{x})\cos(m_j\tilde{y})$$

and

$$\frac{\partial \mathcal{N}}{\partial a_i} = -\mathcal{N}\,\frac{\partial \mathcal{S}\,(s,v,r)}{\partial a_i} =: -\mathcal{N}\,C_i$$

where

$$C_i = \frac{1}{8}\sum_{j=0}^{I-1} a_j \sum_{\alpha\in Q}\sum_{\beta\in Q}\sum_{\gamma\in Q} A_{ij} + A_{ji} =: \sum_{j=0}^{I-1} a_j \, C_{ij}$$

with A_{ij} as defined in (10). Using this notation, we can express the first-order logarithmic derivative of the absolute probability $P(z \mid \mathcal{A})$ of a single measurement in a compact way:

$$\frac{\partial \log\{p(z \mid \mathcal{A})\}}{\partial a_i} = \begin{cases} \frac{\mathcal{N} C_i}{1-\mathcal{N}}, & \text{if } r = \text{sub} \\ \frac{B_i}{\lambda} - C_i, & \text{if } r \in [r_{\min}, r_{\max}] \\ -C_i, & \text{if } r = \text{super} \end{cases}$$

The derivative of the joint absolute measurement probability is then simply the sum of the derivatives of the individual measurement likelihoods

$$\frac{\partial \log\{p(Z \mid \mathcal{A})\}}{\partial a_i} = \sum_{k=0}^{K-1} \frac{\partial \log\{p(z \mid \mathcal{A})\}}{\partial a_i}.$$

The second-order derivatives of the measurement likelihood with respect to the map parameters are given by

$$\frac{\partial^2 \log\{p(z \mid \mathcal{A})\}}{\partial a_j \partial a_i} = \begin{cases} \frac{N(C_{ij}-C_iC_j)}{1-N} + \frac{N^2 C_i C_j}{(1-N)^2}, & \text{if } r = \text{sub} \\ \frac{B_{ij}}{\lambda} - \frac{B_i B_j}{\lambda^2} - C_{ij}, & \text{if } r \in [r_{\min}, r_{\max}] \\ -C_{ij}, & \text{if } r = \text{super} \end{cases}$$

As before, the second-order derivative of the joint measurement log-likelihood is the sum of the second-order derivatives of all measurements.

IV. Experiments

To assess how well DCT maps represent lidar data in comparison to existing mapping approaches, we conduct three series of experiments. In the first series, we compare the spatial map values of DCT maps and grid maps with identical memory requirements to a ground truth map and use the resulting error as a measure of map accuracy. In the second series, we evaluate the likelihoods that DCT maps, grid maps, Gaussian process occupancy maps, and Hilbert maps assign to measurements that were used to build them. The higher this likelihood, the better the respective map explains the underlying data. We conclude this section with a comparison of the empirical execution times of the different approaches.

The data at the basis of our experiments stems from rich planar lidar datasets recorded in spacious indoor environments. Each set contains the corresponding robot poses computed by SLAM, which we use as ground truth poses to build all maps. The data is publicly available from the Robotics Data Set Repository [1]. Table I shows which datasets were used in our experiments.

A. Map Value Comparison

In this experiment series, we compare the map values of DCT maps and decay rate grid maps of different resolutions to the values of a decay rate ground truth grid map. All grid maps are computed according to the algorithm described in [4]. At the beginning, for each dataset, we create a fine-grained ground truth grid map that covers a $10\,\mathrm{m} \times 10\,\mathrm{m}$ patch densely filled with 10^4 lidar measurements. It consists of 200×200 pixels with an edge length of $0.05\,\mathrm{m}$. Then, we use the same sets of measurements to build pairs of one DCT map and one grid map, respectively, for each dataset and each resolution. The maps in these pairs possess the same number of parameters and require the identical amount of memory. We use five different map resolutions: 10×10, 13×13, 20×20, 29×29, and 40×40. For grid maps, they correspond to pixel edge lengths of $1.00\,\mathrm{m}$, $0.75\,\mathrm{m}$, $0.50\,\mathrm{m}$, $0.35\,\mathrm{m}$, and $0.25\,\mathrm{m}$. To give an intuition of what these maps look like, fig. 1 exemplarily shows the 40×40 DCT map of the Intel Research Lab dataset, the corresponding grid map, and the ground truth grid map.

Having created the maps, we sample the ground truth map at the midpoints of all cells that were observed at least once and look up the corresponding values in the DCT map and in the grid map. The resulting map values are hard to compare: As the decay rate λ is defined over the half-open interval $[0, \infty)$, the map values might be infinite. In order to make them comparable, we employ the strictly increasing monotonic transformation function $p_{\text{ref}} = 1 - \exp(-\lambda)$, which maps every decay value to a finite value $p_{\text{ref}} \in [0, 1]$. The value p_{ref} can be interpreted as the absolute probability that a ray is reflected before having traveled a distance of $1\,\mathrm{m}$ in a hypothetical homogeneous medium of decay rate λ. Please note that the distance $1\,\mathrm{m}$ is an arbitrarily chosen parameter. However, while surveying different distance values, we found out that varying this parameter has little effect on the quality of the results. We compute the root mean squared error (RMSE) in p_{ref} between the DCT maps and the ground truth map and between the grid maps and the ground truth map. Table II condenses the corresponding results by determining the mean and the standard deviation of the RMSE values. Additionally, it indicates the p-values of the one-sided paired-sample t-test. Small p-values indicate that the null hypothesis is unlikely and that the alternative hypothesis – the mean RMSE of DCT maps is smaller than the one of grid maps – becomes likely.

While at a resolution of 10×10, both map modalities hardly differ in terms of accuracy, for all finer resolutions, DCT maps outperform grid maps at a confidence level of at least $99\,\%$. Note that the maximum gain in accuracy is located at a resolution of 29×29; at 40×40, DCT maps are still significantly more accurate than grid maps, but the gain is not as large as for 29×29. We attribute this to the fact that with increasing resolution, grid maps converge to the ground truth map, which itself is a grid map.

B. Measurement Probability Comparison

The maps computed in the first experiment series are maximum likelihood maps. Maximum likelihood maps shall maximize the measurement probability of the data that was used to create them. The higher this likelihood, the better the map represents the underlying data. Consequently, in the second experiment series, we interpret the likelihood a map assigns to its underlying data as a measure of its quality. We compare four different approaches: DCT maps, decay rate grid maps, Gaussian process occupancy maps (GPOM), and Hilbert maps, which also model the spatial occupancy probability as a continuous scalar field. More specifically, we use GPOMs with Matérn kernel functions as described in [33] and Hilbert maps with Nyström features, which, according to [35], give the most accurate map results. All hyperparameters are set as described in [33] and [35], respectively. The data at the basis of the experiments is the same as in the previous experiment series, but the number of lidar measurements is reduced to 500.

The comparison is designed to guarantee that all maps have the same memory requirements in terms of numbers of real-valued parameters. For grid maps and DCT maps, we can ensure that by comparing maps with the same number of pixels and spectral parameters, respectively. For GPOM, we randomly downsample the design matrix and the target vector so that the length of the Gaussian process parameter vector matches the number of grid pixels and spectral parameters, respectively. For Hilbert maps, we set the number of components of the Nyström features to the number of parameters corresponding to the specific resolution.

Now, we compute the joint measurement likelihood of all lidar measurements for each map modality. For grid maps, we calculate the measurement likelihood according to [4]. For DCT maps, we follow the equations given in section III-C. For GPOMs and Hilbert maps, we rasterize their continuous occupancy fields with a pixel edge length of $0.05\,\mathrm{m}$, perform ray tracing, and cumulate the occupancy probabilities along the rays.

Table III displays the resulting findings: the mean and standard deviation of the log-likelihood differences between DCT maps and the other approaches over all datasets. After having performed Anderson-Darling tests to ensure that the measurement probability quotients are indeed log-normally distributed, we perform one-sided paired-sample t-tests. For all resolutions, they indicate that DCT maps yield significantly higher measurement log-likelihoods at a confidence level of at least $98.56\,\%$.

The results show that the log-differences between DCT maps and grid maps are two magnitudes smaller than those between DCT maps and GPOM or Hilbert maps, respectively. The level of the difference is influenced by the raster size chosen when computing the measurement likelihood for GPOMs and Hilbert maps. But the main reason for these large differences is the fact that both GPOMs and Hilbert maps have comparatively high memory requirements. GPOMs store the map information in the parameter vector. The number of training points processed is proportional to the length of the parameter vector. As we restricted this length, only a limited number of training points could be processed; as a consequence, the classification results of GPOMs remain rather vague. As far as Hilbert maps are concerned, in [35], the authors recommend to use 1000 components for mapping with Nyström features. In our experiments, we use numbers as small as 100 parameters. Additionally, both GPOMs and Hilbert maps suffer from the fact that they need to sample a limited number of free and occupied training points along the laser rays, whereas DCT maps and decay rate grid maps incorporate the full path information of an arbitrary number of rays. Fig. 2 illustrates the resulting differences in accuracy between the maps produced by the four approaches for 13×13 parameters.

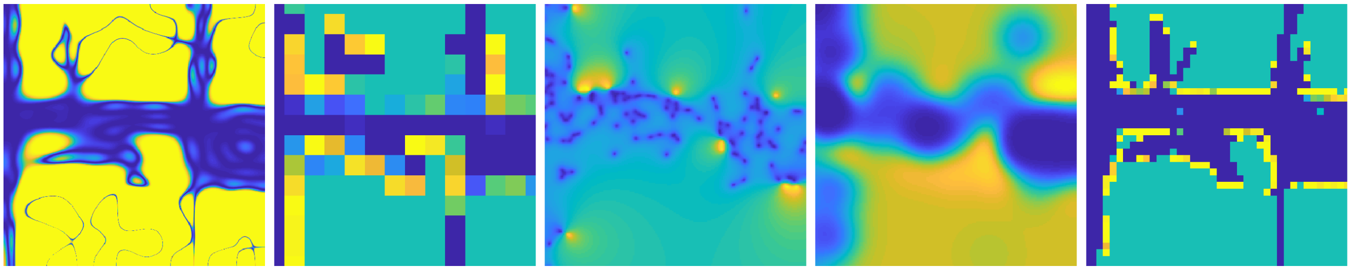
C. Execution Times

To give an intuition of the empirical runtime requirements of each of the methods used in the previous section, we average over ten mapping runs performed per method, dataset, and resolution. Table IV lists the medians of these averages over all datasets. The measurements are collected on an Intel i7-2600K CPU running at $3.40\,\mathrm{GHz}$. Grid maps, DCT maps, and GPOM are implemented in MATLAB R2017b. The GPOM implementation is based on the publicly available GPML toolbox [36]. To time Hilbert maps, we customized the Python implementation provided by [35]. The DCT optimization process is stopped once the relative change in the measurement log-likelihood is smaller than $1 \cdot 10^{-3}$.

Table IV indicates that grid maps are by far the fastest mapping technique. DCT maps and GPOMs are approximately two orders of magnitude slower. This is due to the fact that during the optimization phase, DCT maps and GPOMs need to consider all parameters, which leads to quadratic computational complexity in the number of parameters. The most expensive operation in grid mapping, however, is ray casting, resulting in approximately linear computational complexity in the map resolution. Hilbert maps are three to four magnitudes slower than grid maps, the reason for this probably being the non-differentiable nature of the objective function, which needs to be approximated using finite differences.

V. Conclusion and Future Work

We presented a novel map representation based on the recently introduced decay rate model for lidar sensors [4]. In contrast to most conventional maps, our so-called DCT maps store the map parameters in the discrete frequency domain. We applied the continuous extension of the inverse discrete cosine transform to the spectral parameters to obtain



(a) DCT map. (b) Grid map. (c) GPOM. (d) Hilbert map. (e) Ground truth grid map.

Fig. 2: Maps of different modalities created in the experiment series described in section IV-B for the University of Washington dataset. The four maps to the left all have the same memory requirements of a mere 169 real-valued parameters. The grid map (e) is given as ground truth with a resolution of 40×40. The decay rate maps (a), (b), (e) show p_{ref} values as described in section IV-A, the other maps show occupancy probabilities. Blue means 0, green means 0.5, yellow means 1.

Acapulco Conv. Ctr.	Edmonton Conv. Ctr.	Uni. Freiburg, 101
Uni. Texas, ACES3	FHW museum	Infinite corridors
Belgioioso Castle	Uni. Washington, Seattle	Intel Research Lab
MIT, CSAIL	Uni. Freiburg, 079	Örebro University

TABLE I: The 12 datasets taken from the Robotics Data Set Repository [1] and used in all three experiment series.

	DCT		GM			
l [m]	μ	σ	μ	σ	p [%]	Δ_μ [%]
1.00	0.3314	0.0679	0.3330	0.0708	39.36	0.48
0.75	0.3146	0.0675	0.3319	0.0752	1.01	0.52
0.50	0.2932	0.0645	0.3093	0.0690	0.03	5.21
0.35	0.2571	0.0611	0.2822	0.0672	0.05	8.89
0.25	0.2370	0.0563	0.2543	0.0583	0.07	6.80

TABLE II: Mean and standard deviation of the absolute RMSE values of DCT maps and grid maps with respect to the ground truth map, computed over all datasets. GM denotes grid maps, l is the pixel edge length, μ and σ denote the mean and the standard deviation of the RMSE, respectively, and p is the p-value of the one-sided paired-sample t-test. The variable $\Delta_\mu := 1 - \frac{\mu_{\text{DCT}}}{\mu_{\text{GM}}}$ indicates the improvement in RMSE of DCT maps relative to grid maps.

	$lp_{\text{DCT}} - lp_{\text{GM}}$		$lp_{\text{DCT}} - lp_{\text{GPOM}}$		$lp_{\text{DCT}} - lp_{\text{HM}}$	
l [m]	μ	σ	μ [10^4]	σ [10^4]	μ [10^4]	σ [10^4]
1.00	88.5	74.1	4.21	3.00	4.30	3.42
0.75	150.6	146.5	3.89	2.81	4.19	3.46
0.50	135.7	63.9	3.53	2.60	4.06	3.42
0.35	196.1	101.5	2.97	2.15	4.12	3.32
0.25	96.8	132.0	2.68	1.90	4.12	3.31

TABLE III: Mean and standard deviation of the log-likelihood differences between DCT maps and the other mapping approaches, computed over all datasets. The variable lp denotes the cumulated log-likelihood of all measurements in one dataset, GM denotes grid maps, HM means Hilbert maps, l is the pixel edge length, and μ and σ are the mean and the standard deviation of the log-likelihood differences, respectively.

l [m]	t_{DCT} [s]	t_{GM} [s]	t_{GPOM} [s]	t_{HM} [s]
1.00	3.52	0.0917	1.12	22.8
0.75	4.69	0.0915	1.98	40.6
0.50	3.70	0.0923	3.25	108.4
0.35	18.25	0.0926	6.45	328.3
0.25	39.84	0.0942	14.04	949.3

TABLE IV: Empirical execution time measurements collected during map creation. The variable t denotes the median of the mapping times over all datasets.

a continuously differentiable scalar field in the position domain.

Compared to other mapping approaches like decay rate grid maps, Gaussian process occupancy maps (GPOM), and Hilbert maps, the proposed approach results in significantly improved map accuracy, as demonstrated in extensive real-world experiments. Moreover, DCT maps provide a ray tracing-based forward sensor model that allows to infer measurement probabilities directly from the spectral map representation in closed form, whereas the computation of ray tracing-based measurement probabilities based on continuous occupancy maps like GPOM and Hilbert maps necessitates the rasterization of the map and hence the introduction of a rasterization parameter. As opposed to GPOM and Hilbert maps, DCT maps use the full ray path information when building the map instead of sampling points along the ray.

Due to the promising experimental results, we plan improvements and extensions of DCT maps. First, we will address the issue that the computational complexity of DCT maps is higher than the one of grid maps, and that incremental updates require the repeated optimization of the map parameters. More specifically, we will develop a hybrid approach that locally optimizes the map and that makes use of massive parallelization. Second, we will extend the method by explicitly representing unexplored areas in the map. Currently, DCT maps are not able to distinguish between observed and unobserved map regions. Third, we will investigate how well DCT maps are suited for lossy compression. Finally, we plan to present a complete SLAM system based on DCT maps in the near future.

Acknowledgments

We thank Maani Ghaffari Jadidi for kindly supporting us with his GPOM implementation, Lionel Ott for releasing his Python implementation of Hilbert maps, and Patrick Beeson, Mike Bosse et al., Dieter Fox, Dirk Hähnel, Nick Roy, and Cyrill Stachniss for providing the datasets.

References

[1] A. Howard and N. Roy, "The robotics data set repository (Radish)," 2003. [Online]. Available: http://radish.sourceforge.net/

[2] J. Ryde and M. Brünig, "Non-cubic occupied voxel lists for robot maps," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2009, pp. 4771–4776.

[3] A. Atoyan and J. Patera, "Properties of continuous Fourier extension of the discrete cosine transform and its multidimensional generalization," *Journal of Mathematical Physics*, vol. 45, no. 6, pp. 2468–2491, 2004.

[4] A. Schaefer, L. Luft, and W. Burgard, "An analytical lidar sensor model based on ray path information," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1405–1412, July 2017.

[5] H. Moravec and A. Elfes, "High resolution maps from wide angle sonar," in *Proceedings, 1985 IEEE International Conference on Robotics and Automation*, vol. 2, Mar 1985, pp. 116–121.

[6] L. Luft, A. Schaefer, T. Schubert, and W. Burgard, "Closed-form full map posteriors for robot localization with lidar sensors," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept 2017, pp. 6678–6684.

[7] M. Herbert, C. Caillas, E. Krotkov, I. S. Kweon, and T. Kanade, "Terrain mapping for a roving planetary explorer," in *Proceedings, 1989 International Conference on Robotics and Automation*, vol. 2, May 1989, pp. 997–1002.

[8] R. Triebel, P. Pfaff, and W. Burgard, "Multi-level surface maps for outdoor terrain mapping and loop closing," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2006, pp. 2276–2282.

[9] I. Dryanovski, W. Morris, and J. Xiao, "Multi-volume occupancy grids: an efficient probabilistic 3D mapping model for micro aerial vehicles," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2010, pp. 1553–1559.

[10] D. Meagher, "Geometric modeling using octree encoding," *Computer Graphics and Image Processing*, vol. 19, no. 2, pp. 129 – 147, 1982.

[11] K. M. Wurm, A. Hornung, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: a probabilistic, flexible, and compact 3D map representation for robotic systems," in *Proceedings of the ICRA 2010 workshop on best practice in 3D perception and modeling for mobile manipulation*, vol. 2, 2010.

[12] P. Payeur, P. Hebert, D. Laurendeau, and C. M. Gosselin, "Probabilistic octree modeling of a 3D dynamic environment," in *Proceedings of the International Conference on Robotics and Automation*, vol. 2, Apr 1997, pp. 1289–1296 vol.2.

[13] K. Pathak, A. Birk, J. Poppinga, and S. Schwertfeger, "3D forward sensor modeling and application to occupancy grid based sensor fusion," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2007, pp. 2059–2064.

[14] N. Fairfield, G. Kantor, and D. Wettergreen, "Real-time SLAM with octree evidence grids for exploration in underwater tunnels," *Journal of Field Robotics*, vol. 24, no. 1-2, pp. 03–21, 2007.

[15] T. Krajnik, J. P. Fentanes, G. Cielniak, C. Dondrup, and T. Duckett, "Spectral analysis for long-term robotic mapping," in *2014 IEEE International Conference on Robotics and Automation*, May 2014, pp. 3706–3711.

[16] J. Ryde and H. Hu, "3D mapping with multi-resolution occupied voxel lists," *Autonomous Robots*, vol. 28, no. 2, p. 169, Sep 2009.

[17] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, Oct 2003, pp. 2743–2748 vol.3.

[18] J. Saarinen, H. Andreasson, T. Stoyanov, J. Ala-Luhtala, and A. J. Lilienthal, "Normal distributions transform occupancy maps: application to large-scale online 3D mapping," in *2013 IEEE International Conference on Robotics and Automation*, May 2013, pp. 2233–2238.

[19] J. Saarinen, H. Andreasson, T. Stoyanov, and A. J. Lilienthal, "Normal distributions transform Monte-Carlo localization (NDT-MCL)," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov 2013, pp. 382–389.

[20] T. Stoyanov, J. Saarinen, H. Andreasson, and A. J. Lilienthal, "Normal distributions transform occupancy map fusion: simultaneous mapping and tracking in large scale dynamic environments," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov 2013, pp. 4702–4708.

[21] R. Valencia, J. Saarinen, H. Andreasson, J. Vallv, J. Andrade-Cetto, and A. J. Lilienthal, "Localization in highly dynamic environments using dual-timescale NDT-MCL," in *2014 IEEE International Conference on Robotics and Automation*, May 2014, pp. 3956–3962.

[22] M. Yguel, C. T. M. Keat, C. Braillon, C. Laugier, and O. Aycard, "Dense mapping for range sensors: efficient algorithms and sparse representations," in *Robotics: Science and Systems III, June 27-30, 2007, Georgia Institute of Technology, Atlanta, Georgia, USA*, 2007.

[23] D. Fridovich-Keil, E. Nelson, and A. Zakhor, "AtomMap: a probabilistic amorphous 3D map representation for robotics and surface reconstruction," in *2017 IEEE International Conference on Robotics and Automation*, May 2017, pp. 3110–3117.

[24] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann, "6D SLAM – 3D mapping outdoor environments," *Journal of Field Robotics*, vol. 24, no. 8-9, pp. 699–722, 2007.

[25] H. Surmann, A. Nüchter, and J. Hertzberg, "An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments," *Robotics and Autonomous Systems*, vol. 45, no. 3, pp. 181 – 198, 2003.

[26] D. M. Cole and P. M. Newman, "Using laser range data for 3D SLAM in outdoor environments," in *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, 2006, pp. 1556–1563.

[27] R. Chatila and J. Laumond, "Position referencing and consistent world modeling for mobile robots," in *Proceedings. 1985 IEEE International Conference on Robotics and Automation*, vol. 2, Mar 1985, pp. 138–145.

[28] M. Montemerlo, D. Hähnel, D. Ferguson, R. Triebel, W. Burgard, S. Thayer, W. Whittaker, and S. Thrun, "A system for three-dimensional robotic mapping of underground mines," Carnegie-Mellon University Pittsburgh, PA, School of Computer Science, Tech. Rep. 1, 2002.

[29] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk, "The Digital Michelangelo Project: 3D scanning of large statues," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '00. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 2000, pp. 131–144.

[30] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon, "KinectFusion: Real-time dense surface mapping and tracking," in *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, Oct 2011, pp. 127–136.

[31] S. O'Callaghan, F. T. Ramos, and H. Durrant-Whyte, "Contextual occupancy maps using Gaussian processes," in *2009 IEEE International Conference on Robotics and Automation*, May 2009, pp. 1054–1060.

[32] S. T. O'Callaghan and F. T. Ramos, "Gaussian process occupancy maps," *The International Journal of Robotics Research*, vol. 31, no. 1, pp. 42–62, 2012.

[33] M. Ghaffari Jadidi, J. Valls Miro, and G. Dissanayake, "Gaussian processes autonomous mapping and exploration for range-sensing mobile robots," *Autonomous Robots*, Nov 2017.

[34] S. Dragiev, M. Toussaint, and M. Gienger, "Gaussian process implicit surfaces for shape estimation and grasping," in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 2845–2850.

[35] F. Ramos and L. Ott, "Hilbert maps: Scalable continuous occupancy mapping with stochastic gradient descent," *The International Journal of Robotics Research*, vol. 35, no. 14, pp. 1717–1730, 2016.

[36] C. E. Rasmussen and H. Nickisch, "Gaussian processes for machine learning (gpml) toolbox," *J. Mach. Learn. Res.*, vol. 11, pp. 3011–3015, Dec. 2010.