On-Manifold GMM Registration

Wennie Tabib, Cormac O'Meadhra, and Nathan Michael

Abstract—This paper presents a robust Gaussian Mixture Model (GMM) registration technique to enable mapping and navigation in dark, complex, unstructured domains such as caves and mines. Subterranean environments are often unmapped and challenged by low-lighting conditions and communication constraints. Prior works that leverage direct and image-based techniques with active illumination fail when taking tight turns due to sensor washout, dense methods are sensitive to initialization, and state-of-the-art registration methods that leverage probabilistic models are neither real-time viable nor thoroughly evaluated with real-world data. The proposed approach minimizes the squared L2 norm between two distributions through an on-manifold parameterization of the objective function. The contribution of this paper is a robust, real-time viable distribution-to-distribution registration methodology that considers all possible correspondences between mixture components. The approach is evaluated in a feature-scarce mine, unstructured cave, and on open-source data of an office environment with both LIDAR and depth sensors. The results demonstrate superior performance as compared to the state of the art. Beyond cave and mine environments, the method readily extends to solving the problem of registration in cluttered domains.

Index Terms—Localization, range sensing, mining robotics.

I. INTRODUCTION

SUBTERRANEAN environments are challenged by low-lighting conditions, repetitive structure, and are often unmapped. For example, of the 4378 documented caves in Virginia, only 1348 are mapped [24], which compounds the dangers of cave rescues to extricate injured cavers as no prior maps exists to understand the structure of the terrain. Sinkholes are catastrophic and sudden events that turn buildings to rubble in an instant. Without complete maps, rescue teams must work with limited information to find survivors [28]. Subterranean nuclear waste storage facilities become inaccessible to humans after a radiation leak without risking exposure. However, deploying robots into these environments to locate the source of the leak is challenged due to the repetitive nature of the storage facility [5].

Autonomous robotic mapping in these environments is predicated on accurate pose estimation. The contribution of this paper is a methodology to robustly register laser or depth observations represented as probabilistic generative models by minimizing the squared L2 norm between the distributions. Because the GMM is a compact representation of

The authors are with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA {wtabib,comeadhr,nmichael}@cmu.edu

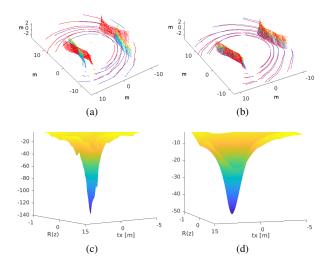Digital Object Identifier (DOI): see top of this page.

Fig. 1: Natural environments, which are the domain of interest, typically have sparse features, rendering feature based matching difficult. Target (red) and source (blue-yellow) point clouds are initially misaligned in (a) and the goal is to align them as in (b). When the pointclouds are represented as GMMs with fully anisotropic covariances the resulting cost function has the steep and non-convex form of (c). Converting the anisotropic covariances into isotropic-planar covariances yields a cost function that is almost completely convex as shown in (d) and enables the accurate registration shown in (b).

the sensor observation, the cost for considering all possible correspondences between the mixture components in the two distributions becomes much more tractable as compared to operating directly on points. The paper proceeds with related work in Section II followed by the proposed approach that formulates an objective function to minimize the squared L2 norm between the two distributions, derives a closed-form gradient and Hessian, and details the optimization framework to register the distributions (see Fig. 1). Section IV evaluates the proposed approach. The conclusion is presented in Section V and the mathematical derivations for the rotation parameterization and the full derivation of the gradient and Hessian are included in Appendices A and B, respectively.

II. RELATED WORK

The state-of-the-art navigation and mapping algorithms can be segregated into image-based and depth-based pose estimation techniques. Feature-based image processing algorithms fail in environments that do not enable feature extraction and matching, particularly, in low-lighting conditions and in the presence of repetitive structure. Dense methods such as [19, 22] operate directly on pointclouds, but are typically more sensitive to initialization.

Direct methods may be leveraged to estimate pose from high-frame rate depth or RGB-D sensor observations, but require motions between frames to be small and rely on the brightness constancy assumption which does not hold in
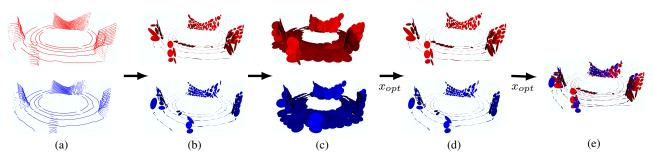
Fig. 2: Overview of methodology: (a) Pointclouds taken at two different poses (shown in red and blue) are converted into (b) GMMs with anisotropic covariances. The covariances are converted to (c) isotropic-planar (isoplanar) and an optimization is run that minimizes the squared L2 norm between the two distributions. The output of the optimization x_{opt} is used to seed a second optimization with anisotropic covariances. The results are rigid transformation parameters that align the GMMs (shown in (e)).

subterranean environments [13, 17]. Alismail et al. [1] seek to overcome the limitations of feature-based and direct methods by developing a hybrid method that employs bit-plane features within a direct alignment framework, but the approach fails in subterranean environments when taking tight turns due to washout when the light source output is constrained to be very close to the image sensor.

The Iterative Closest Point (ICP) algorithm [2] computes point correspondences and identifies the optimal alignment of corresponding pairs in the least-squares sense. While fast, it is sensitive to outliers and initialization, and many variants have been proposed to overcome these limitations [6, 18]. Generalized ICP [21] increases the robustness of ICP by modeling the local structure of the data as a disk centered on each point and normal and matching via a plane-to-plane distance metric. A significant limitation of standard ICP approaches is the use of hard correspondences, but while soft correspondence methodologies have been shown to yield improved performance, these methods come at increased computational cost [27]. While one could argue that downsampling the pointcloud is a viable method for increasing computational speed, this operation may result in loss of information [14].

The Normal Distribution Transform (NDT) framework learns anisotropic multivariate normal distributions over points assigned to a discrete set of bins in the observed space and performs registration by minimizing the L2 norm between two NDT maps [25]. This approach discretizes the space, which limits the representational power of the distributions and increases the computational burden of the framework. The proposed approach is distinct from this work in that the role of the determinant is explicitly accounted for and all possible correspondences between the two distributions are considered.

Approaches have been proposed that employ GMMs to represent pointclouds [8–10, 15]. Alignment is performed by minimizing a distance metric as in [15], or integrating into the model learning stage through an expectation-maximization (EM) framework [8, 10]. EM-based approaches assume that the pointclouds to be matched are both observations of the same scene, which is not realistic in practice as measurements often consist of large non-overlapping regions.

Methods that leverage probabilistic generative models to register point sets have been demonstrated to be more robust to outliers and noise [7, 10]. However, state-of-the-art methods are either not real-time viable [10] or tested on a single scan with varying rotations, translations and rates of downsampling

so that the data is virtually identical for every test [7]. Prior work has demonstrated the viability of real-time GMM computation through hardware acceleration [9]. Building on this work, all measurements are assumed to be represented by compact anisotropic GMMs and the gap in the state of art is bridged with a real-time viable registration technique that is tested on real-world data of subterranean and cluttered environments.

III. METHODOLOGY

Figure 2 illustrates the proposed approach. First, pointclouds (Fig. 2a) are converted to full anisotropic GMMs (Fig. 2b). The mathematical definition of GMMs is provided in Section III-A. Next, Section III-B formulates an objective function for registration from the correlation integral that maximizes the overlap between two distributions. Section III-C derives the closed-form equations for the gradient and Hessian. Section III-D details the optimization framework that optimizes first with GMMs with isoplanar covariances (Fig. 2c) that serves to smooth the cost function and reduce local minima followed by a refinement optimization using the original anisotropic covariances (Fig. 2d) to obtain accurate registration parameters and align the distributions (Fig. 2e).

A. Preliminaries

A Gaussian Mixture Model (GMM) is a probabilistic model composed of a weighted combination of M Gaussian densities, with the form

$$p(\boldsymbol{x}) = \sum_{m=1}^{M} \pi_m \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_m, \boldsymbol{\Lambda}_m) \qquad (1)$$

Each Gaussian density $\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_m, \boldsymbol{\Lambda}_m)$ is referred to as a component of the GMM and parameterized by a mean $\boldsymbol{\mu}_m$ and covariance $\boldsymbol{\Lambda}_m$. The components are combined through a set of normalized mixing coefficients $\{\pi_1, \ldots, \pi_m\}$ that represent the prior probability of selecting the mth component. The Expectation Maximization (EM) algorithm computes the GMM parameters by introducing latent correspondences to factorize the joint likelihood so that closed form maximization of parameters of individual components is possible.

B. Registration

GMM to GMM registration, commonly referred to as distribution to distribution (D2D) registration, is achieved by aligning modeled probability densities. Given two GMMs, $\mathcal{G}_0(\boldsymbol{x})$

and $\mathcal{G}_1(\boldsymbol{x})$, with potentially different numbers of components, the transformation parameters $\boldsymbol{\theta}$, are sought that minimize the *distance* between $\mathcal{G}_0(\boldsymbol{x})$ and $T(\mathcal{G}_1(\boldsymbol{x}), \boldsymbol{\theta})$. $\mathcal{G}_0(\boldsymbol{x})$ and $\mathcal{G}_1(\boldsymbol{x})$ are defined as:

$$\mathcal{G}_0(\boldsymbol{x}) = \sum_m^M \pi_m \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_m, \boldsymbol{\Lambda}_m)$$

$$\mathcal{G}_1(\boldsymbol{x}) = \sum_k^K \tau_k \mathcal{N}(\boldsymbol{x}|\boldsymbol{\nu}_k, \boldsymbol{\Omega}_k)$$

and $T(\cdot, \boldsymbol{\theta})$ is the rigid transformation function consisting of a rotation $\boldsymbol{R}$ and translation $\boldsymbol{t}$. A transformed GMM with parameters $\{\tau_k, \boldsymbol{\nu}_k, \boldsymbol{\Omega}_k\}$ is expressed as

$$T(\mathcal{G}_1(\boldsymbol{x}), \boldsymbol{\theta}) = \sum_{k=1}^K \tau_k \mathcal{N}(\boldsymbol{x}|\boldsymbol{R}\boldsymbol{\nu}_k + \boldsymbol{t}, \boldsymbol{R}\boldsymbol{\Omega}_k\boldsymbol{R}^T) \qquad (2)$$

Following the success of D2D alignment for NDT-map [25] and in the original work demonstrating GMMs for point set alignment [15], the squared L2 norm is employed as the measure of distance between the target GMM $\mathcal{G}_0$ and the transformed source GMM $T(\mathcal{G}_1(\boldsymbol{x}), \boldsymbol{\theta})$. The cost function to minimize may be written as:

$$\begin{aligned}\boldsymbol{\theta}^* &= \arg\min_{\boldsymbol{\theta}} \int \|\mathcal{G}_0(\boldsymbol{x}) - T(\mathcal{G}_1(\boldsymbol{x}), \boldsymbol{\theta})\|_2^2 \mathrm{d}\boldsymbol{x} \\ &= \arg\min_{\boldsymbol{\theta}} \int \left(\|\mathcal{G}_0\|_2^2 + \|T(\mathcal{G}_1, \boldsymbol{\theta})\|_2^2 - 2\mathcal{G}_0 T(\mathcal{G}_1, \boldsymbol{\theta}) \right) \mathrm{d}\boldsymbol{x} \end{aligned} \qquad (3)$$

where dependence on $\boldsymbol{x}$ is omitted for brevity. Eqn. (3) may be simplified by noting that the first term does not depend on the transformation parameters $\boldsymbol{\theta}$ and that energy in a distribution (which is given by the second term) is invariant under rigid transformation. Intuitively, this means that application of the transformation parameters does not change the shape of the distribution. Thus, the minimization of the squared L2 distance corresponds to maximization of the correlation of the two distributions:

$$\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} \int 2\mathcal{G}_0(\boldsymbol{x}) T(\mathcal{G}_1(\boldsymbol{x}), \boldsymbol{\theta}) \mathrm{d}\boldsymbol{x} \qquad (4)$$

The objective function (4) has the same extrema as that of the Cauchy-Schwarz divergence [16], due to the monotonicity of the log function. In the case of two GMMs, the correlation integral of (4) has a closed-form solution given by [16]

$$\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} \sum_{m=1}^M \sum_{k=1}^K \pi_m \tau_k \mathcal{N}(\boldsymbol{\mu}_m | \boldsymbol{R}\boldsymbol{\nu}_k + \boldsymbol{t}, \boldsymbol{\Lambda}_m + \boldsymbol{R}\boldsymbol{\Omega}_k\boldsymbol{R}^T) \quad (5)$$

which may be written as

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} -F, \qquad (6)$$

where

$$F = \sum_{m=1}^M \sum_{k=1}^K f_{mk}$$

$$f_{mk} = \pi_m \tau_k \frac{|\boldsymbol{\Sigma}_{mk}|^{-1/2}}{(2\pi)^{3/2}} \exp\left(-\frac{1}{2}\boldsymbol{y}_{mk}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{y}_{mk}\right)$$

$$\boldsymbol{y}_{mk} = \boldsymbol{\mu}_m - \boldsymbol{R}\boldsymbol{\nu}_k - \boldsymbol{t}$$

$$\boldsymbol{\Sigma}_{mk} = \boldsymbol{\Lambda}_m + \boldsymbol{R}\boldsymbol{\Omega}_k\boldsymbol{R}^T$$

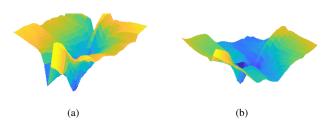|	
(a)	(b)

Fig. 3: The non-convexity of the cost function is evident. The cost function with the determinant (a) exhibits a wider basin of attraction than the cost function without (b).

The goal is to maximize the correlation, which is an integral whose value depends on the transformation parameters $\boldsymbol{\theta}$. Using [15, 16], the correlation integral may be cast as the minimization problem in (6) that finds the alignment that maximizes the overlap between the distributions. The optimal rigid transformation parameters may be solved for by setting the gradient of (6) to zero. The Hessian contains information about the local curvature of the function, so it increases the rate of convergence of the optimization. The gradient and Hessian are derived in closed-form in Section III-C. An unconstrained optimization problem is cast that optimizes objective (6) via a trust-region method (Section III-D).

The cost function is distinct from that of [25], in that the role of the determinant is explicitly accounted for in the cost function. That is, f_{mk} becomes

$$\pi_m \tau_k \frac{1}{(2\pi)^{3/2}} \exp\left(-\frac{1}{2}\boldsymbol{y}_{mk}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{y}_{mk}\right) \qquad (7)$$

when the determinant is removed. The determinant acts to increase the weight applied to components that have high certainty such that these components have greater attractive forces. The effect of the determinant on the shape of the cost function is shown in Fig. 3. Retaining the determinant in the cost function increases the complexity of the expression derived for both the gradient and the Hessian, however, empirical results suggest that inclusion of the determinant increases robustness.

The objective (6) is highly non-convex due the the nature of the GMM, as is demonstrated empirically in Fig. 3. Furthermore, the optimization domain SE(3), which is the product manifold SO(3) $\times \mathbb{R}^3$, is non-convex due to the non-convexity of SO(3). The structure of the manifold can be better captured by leveraging a minimal axis-angle parameterization that uniquely defines a rotation through the exponential map. This reparameterization enables the use of more straightforward, unconstrained optimization techniques. Details of the rotation parameterization may be found in Appendix A.

C. Gradient and Hessian

The gradient on the manifold $\mathbb{R}^3 \times SO(3)$ is derived and has coordinates $\boldsymbol{\xi} = (\boldsymbol{t}, \boldsymbol{u})$. For brevity, the partial derivative is expressed as $\boldsymbol{R}_{\xi_a} = \frac{\partial \boldsymbol{R}}{\partial \xi_a}$. The derivatives $\boldsymbol{R}_{u_a}$ and $\boldsymbol{R}_{u_b u_a}$ are given in Appendix A. For the remaining derivatives, $\boldsymbol{R}_{t_a} = \boldsymbol{0}$, $\boldsymbol{t}_{u_a} = \boldsymbol{0}$, $\boldsymbol{t}_{t_a} = \boldsymbol{e}_a$, $\boldsymbol{t}_{t_b t_a} = \boldsymbol{0}$ and both mixed second derivatives are zero. The gradient is given by

$$\frac{\partial F}{\partial \xi_a} = \sum_{m=1}^M \sum_{k=1}^K f_{mk}(d_{mk}^{(a)} + q_{mk}^{(a)}) \qquad (8)$$

where

$$d_{mk}^{(a)} = -Tr\left\{\boldsymbol{\Omega}_k \boldsymbol{R}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{R}_{\xi_a}\right\} \qquad (9)$$

$$q_{mk}^{(a)} = \boldsymbol{y}_{mk}^T \boldsymbol{\Sigma}_{mk}^{-1} (\boldsymbol{R}_{\xi_a} \boldsymbol{\nu}_k + \boldsymbol{t}_{\xi_a}) + \boldsymbol{y}_{mk}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{R} \boldsymbol{\Omega}_k \boldsymbol{R}_{\xi_a}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{y}_{mk} \quad (10)$$

Using the compact representation of the gradient (8), the Hessian is expressed as

$$\frac{\partial^2 F}{\partial \xi_b \partial \xi_a} = \sum_{m=1}^{M} \sum_{k=1}^{K} f_{mk} \left[(d_{mk}^{(a)} + q_{mk}^{(a)})(d_{mk}^{(b)} + q_{mk}^{(b)}) + (D_{mk}^{(ba)} + Q_{mk}^{(ba)}) \right]$$

$$(11)$$

where $D_{mk}^{(ba)} = \frac{\partial}{\partial \xi_b} d_{mk}^{(a)}$, $Q_{mk}^{(ba)} = \frac{\partial}{\partial \xi_b} q_{mk}^{(a)}$, and $d_{mk}^{(b)}$ and $q_{mk}^{(b)}$ are the same as $d_{mk}^{(a)}$ and $q_{mk}^{(a)}$, but refer to elements of ξ_b instead of ξ_a. The expressions are expanded as

$$D_{mk}^{(ba)} = -Tr\Bigg\{\boldsymbol{\Omega}_k \bigg(\boldsymbol{R}_{\xi_b}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{R}_{\xi_a}$$

$$- \boldsymbol{R}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{Z}_{mk}^{(b)} \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{R}_{\xi_a} + \boldsymbol{R}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{R}_{\xi_b \xi_a} \bigg) \Bigg\} \qquad (12)$$

$$Q_{mk}^{(ba)} = -\Bigg(\boldsymbol{j}_{mk}^{(b)T} \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{j}_{mk}^{(a)} - \boldsymbol{y}_{mk}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{Z}_{mk}^{(b)} \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{j}_{mk}^{(a)}$$

$$+ \boldsymbol{y}_{mk}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{H}_{mk}^{(ba)} - \boldsymbol{y}_{mk}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{Z}_{mk}^{(a)} \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{j}_{mk}^{(b)}$$

$$+ \boldsymbol{y}_{mk}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{Z}_{mk}^{(b)} \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{Z}_{mk}^{(a)} \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{y}_{mk}$$

$$- \frac{1}{2} \boldsymbol{y}_{mk}^T \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{Z}_{mk}^{(ba)} \boldsymbol{\Sigma}_{mk}^{-1} \boldsymbol{y}_{mk} \Bigg) \qquad (13)$$

where

$$\boldsymbol{j}_{mk}^{(a)} = -\boldsymbol{R}_{\xi_a} \boldsymbol{\nu}_k - \boldsymbol{t}_{\xi_a}$$

$$\boldsymbol{Z}_{mk}^{(a)} = \boldsymbol{R}_{\xi_a} \boldsymbol{\Omega}_k \boldsymbol{R}^T + \boldsymbol{R} \boldsymbol{\Omega}_k \boldsymbol{R}_{\xi_a}^T$$

$$\boldsymbol{H}_{mk}^{(ba)} = \frac{\partial \boldsymbol{j}_a^{mk}}{\partial \xi_b} = -\boldsymbol{R}_{\xi_b \xi_a} \boldsymbol{\nu}_k$$

$$\boldsymbol{Z}_{mk}^{(ba)} = \boldsymbol{R}_{\xi_b \xi_a} \boldsymbol{\Omega}_k \boldsymbol{R}^T + \boldsymbol{R}_{\xi_a} \boldsymbol{\Omega}_k \boldsymbol{R}_{\xi_b}^T + \boldsymbol{R}_{\xi_b} \boldsymbol{\Omega}_k \boldsymbol{R}_{\xi_a}^T + \boldsymbol{R} \boldsymbol{\Omega}_k \boldsymbol{R}_{\xi_b \xi_a}^T$$

Please see Appendix B for the detailed derivation of both the gradient and the Hessian.

D. Optimization Framework

As stated in Section III-B, the cost function to minimize is highly non-convex, which may result in a poor quality registration unless the initialization is close to the actual solution. A technique similar to that of GICP [21] is employed to smooth the cost function and prevent the optimizer from becoming trapped in local minima. Segal et al. [21] note that pointclouds taken from two different positions sample a two-dimensional manifold in three-dimensional space, so each measured point only provides a constraint along its surface normal. When the sampled points are represented as a probabilistic generative model, the covariance about the mean is modified to model the high certainty about the surface normal direction and low certainty along the local plane. The modified covariance is computed as the eigen decomposition $\boldsymbol{\Lambda}_m = \boldsymbol{U}_m \boldsymbol{D}_m \boldsymbol{U}_m^T$ and $\boldsymbol{\Lambda}_m$ is replaced with

$$\boldsymbol{\Lambda}_m = \boldsymbol{U}_m \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \epsilon \end{bmatrix} \boldsymbol{U}_m^T \qquad (14)$$

where ϵ is a small constant representing the covariance along the normal. This has the effect of smoothing the cost function

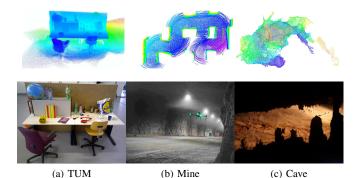(a) TUM   (b) Mine   (c) Cave

Fig. 4: (a) An RGB-D dataset of a cluttered office environment and laser datasets of a (b) mine (Image credit: Near Earth Autonomy) and (c) unstructured cave environment are employed to evaluate the proposed approach. The top row of images are from left-to-right a reconstruction of depth observations from the TUM RGB-D SLAM dataset and benchmark, laser observations of a mine environment, and Faro scans of the undeveloped cave. The bottom row consists of images of the environments.

(see Fig. 1d for a representative example). However, this change in cost function is not guaranteed to preserve the minimum at the same location. Therefore, the optimization is run twice: once with the modified isotropic-planar (isoplanar) covariances specified in [21] to smooth the cost function followed by a second run with the original cost function seeded with the results from the first run. The second optimization quickly converges.

The cost function proposed by Stoyanov et al. [25] to remove the contribution of the determinant (see (7)) is compared to the proposed approach, called the Isoplanar Hybrid approach, in Section IV-A.

An unconstrained optimization problem is realized through the on-manifold parameterization of the objective function (6). The Riemannian trust-region method with conjugate gradients is used to optimize the objective, which is implemented in the MATLAB manifold optimization toolbox, manopt [4].

IV. RESULTS

The selection of the Isoplanar Hybrid approach as the proposed approach is first motivated by comparing to variants of the GMM registration in Section IV-A and then by comparing to the state-of-art registration methods, NDT D2D[1] [25], GICP[2] [21], and NICP[3] [22], in Section IV-B. The performance is evaluated with three datasets consisting of RGB-D and Velodyne VLP-16 laser observations illustrated in Fig. 4.

The GMM registration algorithms consider all pairs of components instead of just the closest pairs as is done in [25]. Preliminary tests that incorporate a d_{max} parameter analogous to GICP's to limit the number of component pairs considered during registration suggests a considerable increase in speed; however, the d_{max} parameter is highly dependent on the observed environment and correct selection of this parameter is a challenge in its own right so the timing results are not reported.

The metrics used to evaluate the proposed method are the Root Mean Square Error (RMSE) as detailed in [26] and the

---

[1]https://github.com/OrebroUniversity/perception_oru-release

[2]https://github.com/PointCloudLibrary/pcl

[3]https://github.com/yorsh87/nicp

odometric error between consecutive frames. The RMSE is defined as the relative pose error at time step i:

$$\mathbf{E}_i := \left(\mathbf{Q}_i^{-1}\mathbf{Q}_{i+1}\right)^{-1}\left(\mathbf{P}_i^{-1}\mathbf{P}_{i+1}\right) \tag{15}$$

$$RMSE(\mathbf{E}_{1:n}) := \left(\frac{1}{n-1}\sum_{i=1}^{n-1}\|trans(\mathbf{E}_i)\|^2\right)^{1/2} \tag{16}$$

where $trans(\mathbf{E}_i)$ refers to the translational components of the relative pose error $\mathbf{E}_i$, the estimated trajectory $\mathbf{P}_1,\dots\mathbf{P}_n \in SE(3)$ and the ground truth trajectory $\mathbf{Q}_1,\dots,\mathbf{Q}_n \in SE(3)$. The odometric error is computed as the translation and rotation error between frames 1 and j where $j \in [1,n]$:

$$\mathbf{E}_j := \left(\mathbf{Q}_1^{-1}\mathbf{Q}_j\right)^{-1}\left(\mathbf{P}_1^{-1}\mathbf{P}_j\right) \tag{17}$$

$$OE(\mathbf{E}_{1:n}) := \|trans(\mathbf{E}_j)\| \tag{18}$$

For both RMSE and odometric errors, the rotation errors are similarly computed.

A. Evaluation of GMM Registration Variants

Two evaluations are conducted that motivate the selection of the Isoplanar Hybrid approach as the proposed approach. Each evaluation highlights the strengths and weaknesses of the GMM registration variants. There are five variations considered: the *m-GMM Isoplanar* method leverages the isoplanar covariances of (14) with cost function (6); the proposed *m-GMM Isoplanar Hybrid* approach performs an optimization with cost function (6) and isoplanar covariances (14) followed by a refinement optimization using (6) and anisotropic covariances seeded with the results from the first optimization; the *m-GMM No Det.* variant performs a single optimization using the cost function from (7) and anisotropic covariances; the *m-GMM No Det. Hybrid* approach first performs an optimization using the cost function from (7) and anisotropic covariances followed by a second optimization with the cost function from (6) and anisotropic covariances seeded with the results from the first optimization; and the *m-GMM Anisotropic* approach performs a single optimization using the cost function from (6) and anisotropic covariances. In each case the m stands for the number of components in the GMMs.

1) TUM Dataset: The first evaluation is performed with a complete TUM RGB-D SLAM benchmark dataset consisting of over 2500 depth images shown in Fig. 4a. The RMSE results are shown in Fig. 5. The Isoplanar approach yields poorer registration results as compared to the other approaches because the cost function changes the shape of the covariances to bias certainty along the direction of the normal, which has the effect of making the optimization more robust to large pose differences between consecutive scans but introduces some error in the final registration results for this dataset where the difference in pose is small (approximately 1-2 cm and 1-2 degrees between consecutive scans). Note that the Isoplanar Hybrid approach overcomes this problem by running a second optimization with the anisotropic covariances.

2) Mine Dataset: The second evaluation is performed with more than 300 laser observations from a mine environment shown in Fig. 4b. The RMS errors are shown in Fig. 6. The translation and rotation differences between successive

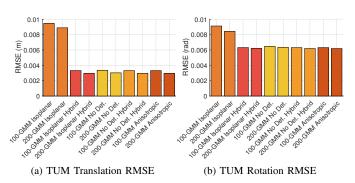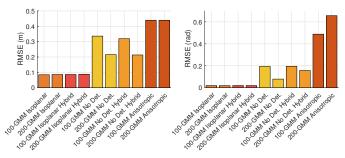(a) TUM Translation RMSE   (b) TUM Rotation RMSE

Fig. 5: RMS errors for translation and rotation for each pair of consecutive observations in the TUM dataset consisting of more than 2500 depth images. The Isoplanar Hybrid approach overcomes the poor registration results of the Isoplanar variant by running a refinement optimization with anisotropic covariances.

(a) Mine Frame-to-frame Translation RMSE   (b) Mine Frame-to-frame Rotation RMSE

Fig. 6: RMS errors for translation and rotation for each pair of consecutive observations in the Mine dataset consisting of more than 300 laser scans. The Isoplanar approaches have the best performance because biasing the covariances along the normal direction makes makes the optimization more robust to large translations and rotations.

observations is much larger in this dataset than the TUM dataset (20-60 cm and 2-10 degrees, respectively). In this case the Isoplanar approaches perform best because the effect of biasing the covariances to have higher certainty in the direction of the normal is to make the optimization more robust to larger translations and rotations between successive observations. The Anisotropic registration is unable to overcome local minima (representative example shown in Fig. 1) which results in poor registration performance and the approaches that remove the determinant have a much narrower basin of attraction (representative example shown in Fig. 3) as compared to the Isoplanar approaches, which yields an optimization that is much more sensitive to initialization.

Because the Isoplanar Hybrid registration approach obtains accurate registration results in both evaluations, it is used to evaluate against state-of-art registration approaches.

B. Evaluation against State-of-art Methods

The Isoplanar Hybrid GMM approach is compared to NDT D2D, GICP, and NICP. The cell size of the NDT D2D registration and Isoplanar Hybrid GMM registration approaches is varied in the evaluations to determine the cell sizes and numbers of components that produce the best performance for each dataset. Varying the d_{max} parameter of GICP, which is the maximum matching threshold, does not significantly affect accuracy or timing performance so the parameter is

Dataset	GICP (s)	NDT (s)	NICP (s)	GMM (s)
TUM	3.3565	0.0357	0.3703	0.2959
Mine	0.3238	0.0045	0.3754	0.3810
Cave	0.5536	0.0030	0.4904	0.3905

TABLE I: Timing results for each method. The 100-GMM Isoplanar Hybrid approach remains within reasonable bounds of timing as compared to the state of the art.

conservatively set to be 100 for all tests. NICP provides optimized parameters for depth and laser sensors so the appropriate configuration file is used depending on the sensor used in the dataset.

Timing results are reported in Table I. All of the tests are run with a single thread on an Intel i7-6700K CPU with 32GB of RAM. The evaluations assume that GMMs and NDT maps are pre-computed so only the time taken to register the distributions is counted. Hardware accelerated methods exist to create GMMs (e.g., [9, 23]) so the time to create these distributions is not counted. In addition, the NICP algorithm operates in image space, so to test the laser pointcloud datasets, the pointclouds are projected to a spherical 1000×1500 depth image per the parameters checked into the NICP codebase. One may note that the reported timing results are much larger than what is advertised in [22], and the reason for this is this work operates on the full-resolution 480×640 depth images to analyze the effectiveness of the registration algorithms, whereas the images in [22] are 120×160—a factor of $16\times$ fewer points—which accounts for the difference in timing results.

1) TUM Dataset: Each method is evaluated with the approximately 2500 sensor observations from the TUM dataset shown in Fig. 4a. The RMSE values shown in Fig. 7 are used to determine the optimal parameters for the NDT cell size and number of GMM components to use to compute the odometric error.

Figure 7 illustrates that for the TUM dataset, an NDT cell size of 0.05 m produces the best performance and 100 components in the Isoplanar Hybrid registration method is sufficient. The odometric error is shown in Fig. 8 and two views of the reconstructed paths are shown in Fig. 9. GICP produces the poorest performance for this dataset, which is similar to the behavior of the m-GMM Isoplanar variant from Fig. 5 that also uses the isoplanar covariances without a refinement optimization. The NDT approach fares slightly better in terms of translation performance which aligns with the behavior of the m-GMM No Det. registration variant from Fig. 5. NICP performs better than the NDT approach but has about 0.9 m more error in translation and 0.26 rad. error than the 100-GMM Isoplanar Hybrid approach by the end of the trajectory.

2) Mine Dataset: The next dataset is the Mine dataset consisting of more than 300 laser scans, and exhibits significantly larger rotations and translations as compared to the TUM dataset (20-60 cm and 2-10 degree differences between consecutive observations). The RMS errors for each method are shown in Fig. 10. The NDT with cell size equal to 0.5 m performs best and 100 components for the GMM registration are sufficient to outperform the state of the art in terms of
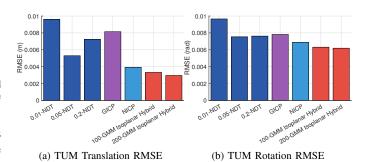

(a) TUM Translation RMSE   (b) TUM Rotation RMSE

Fig. 7: The RMS error for (a) translation and (b) rotation demonstrate the 100-GMM Isoplanar Hybrid approach is sufficient to outperform the other approaches for frame-to-frame RMSE.
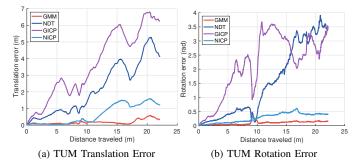

(a) TUM Translation Error   (b) TUM Rotation Error

Fig. 8: The (a) translation and (b) rotation error statistics are plotted as a function of distance traveled. The proposed approach has lower odometric error compared to the state of the art.
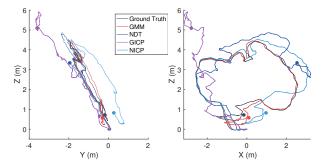

Fig. 9: Two views of the reconstructed paths for the TUM dataset. The trajectories begin at (0,0,0) and end at the colored dots. Note the 100-GMM Isoplanar Hybrid approach closely matches the ground truth trajectory as compared to the other approaches.

translation RMSE. While NDT outperformed the GICP results for the TUM dataset, GICP outperforms NDT for the mine dataset as shown in Fig. 11. The estimated trajectories using each of the methods are shown in Fig. 12.

3) Cave Dataset: The last dataset is the Cave dataset consisting of more than 350 laser scans from Rapps Cave in Greenbrier County, WV. The terrain exhibits complex concavities and disjoint objects, which are visible in the bottom image of Fig. 4c. An aerial robot equipped with a VLP-16 laser scanner was used to obtain data in the cave. A FARO[4] laser scanner was used to collect ground truth.

The translation and rotation RMS errors may be found in Fig. 13, which illustrates that the 100-Isoplanar Hybrid GMM approach is sufficient for this dataset and NDT D2D performs best with a cell size of 2 m. The odometric errors

---

[4]https://www.faro.com
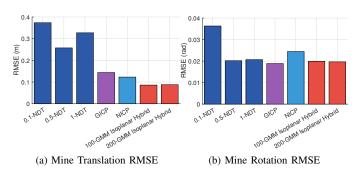
(a) Mine Translation RMSE

(b) Mine Rotation RMSE

Fig. 10: The (a) translation and (b) rotation RMSE for the Mine dataset are shown. A cell size of 0.5m performs best for the NDT map. The GMM registration approach performs well with 100 components.
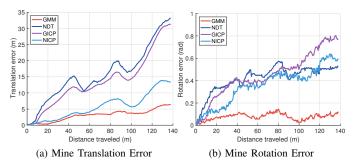

(a) Mine Translation Error

(b) Mine Rotation Error

Fig. 11: The (a) translation and (b) rotation odometric error statistics for the Mine dataset demonstrate that the 100-GMM Isoplanar Hybrid approach significantly outperforms the other approaches.
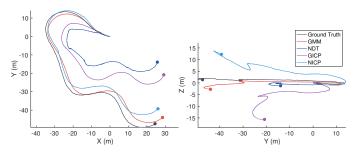

Fig. 12: Two views of the estimated trajectories for the Mine dataset. The trajectories begin at (0,0,0) and end at the colored dots. Note the 100-GMM Isoplanar Hybrid approach closely matches the ground truth trajectory.

(a) Cave Translation RMSE

(b) Cave Rotation RMSE

Fig. 13: The (a) translation and (b) rotation RMSE for each approach with the Cave dataset.

(a) Cave Translation Error

(b) Cave Rotation Error

Fig. 14: The (a) translation and (b) rotation error statistics demonstrate that the 100-GMM Isoplanar Hybrid approach outperforms the other approaches.

Fig. 15: Two views of the estimated trajectories for the Cave dataset. The trajectories begin at (0,0,0) and end at the colored dots. Note the 100-GMM Isoplanar Hybrid approach outperforms the NICP, 2 m NDT Map and GICP approaches, particularly towards the end of the trajectory.

are shown in Fig. 14. While the NICP approach initially does well, the performance deteriorates towards the end of the trajectory. The 100-GMM Isoplanar Hybrid approach most closely follows the ground truth trajectory for this dataset. The trajectories of each of the approaches is shown in Fig. 15.

V. CONCLUSION AND FUTURE WORK

This paper demonstrated a real-time viable method for registering GMMs in feature-deprived and dark environments such as mines and caves. The approach minimizes the squared L2 norm between the two distributions and considers all possible correspondences between mixture components. The compactness of the GMM representation is leveraged to decrease the runtime of the algorithm. Superior results compared to the state of the art are obtained in a degraded mine, unstructured cave, and cluttered office environments. Future work will consist of leveraging the approach in multi-robot contexts to enable accurate registration between robots mapping subterranean voids.
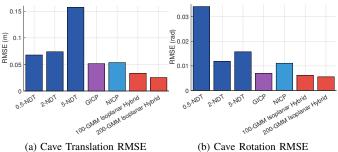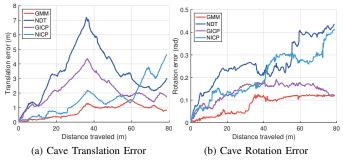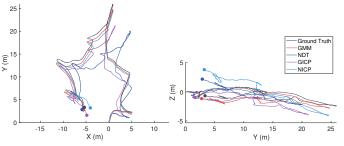
APPENDIX A
ROTATION PARAMETERIZATION

To remove the need for the explicit SO(3) constraint, the minimal axis-angle parameterization is employed that uniquely defines a rotation through the exponential map [3]. In the axis-angle formulation, a vector $\boldsymbol{u}$ is constructed from a rotation angle $\alpha = |\boldsymbol{u}|$ and a rotation axis $\bar{\boldsymbol{u}} = \alpha^{-1}\boldsymbol{u}$. Care must be taken with this representation as α approaches π, since the cost-function wraps around at this point. In the applications discussed in this paper, the rotation angles observed are typically sufficiently small so as to avoid this issue, but in other applications explicit reasoning about this edge case may be necessary [12].

The exponential map utilized to reconstruct a rotation matrix from the axis-angle representation matrix is defined in terms of the matrix exponential as

$$\boldsymbol{R}(\boldsymbol{u}) = e^{[\boldsymbol{u}]_\times}, \quad [\boldsymbol{u}]_\times = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \quad (19)$$

The compact form of the Jacobian of the exponential map is given by [11] from which the Hessian is derived as

$$\boldsymbol{R}_{u_i} = \frac{u_i[\boldsymbol{u}]_\times + [\boldsymbol{u} \times (\boldsymbol{I_3} - \boldsymbol{R})\boldsymbol{e}_i]_\times}{\|\boldsymbol{u}\|^2}\boldsymbol{R} \tag{20}$$

$$\boldsymbol{R}_{u_j u_i} = \Bigg[\delta_{ij}\boldsymbol{u} + u_i\boldsymbol{e}_j + [\boldsymbol{e}_j]_\times(\boldsymbol{I_3} - \boldsymbol{R})\boldsymbol{e}_i$$
$$- [\boldsymbol{u}]_\times \boldsymbol{R}_{u_j}\boldsymbol{e}_i\Bigg]_\times \frac{\boldsymbol{R}}{\|\boldsymbol{u}\|^2} + \boldsymbol{R}_{u_i}\left(\boldsymbol{R}^T\boldsymbol{R}_{u_j} - \frac{2u_j\boldsymbol{I_3}}{\|\boldsymbol{u}\|^2}\right) \tag{21}$$

APPENDIX B
DERIVATION OF GRADIENT AND HESSIAN

Starting from the cost function $F = \sum_m \sum_k f_{mk}$, the terms $d_{mk}^{(a)}$ and $q_{mk}^{(a)}$ are derived from $d_{mk}^{(a)} = |\boldsymbol{\Sigma}_{mk}|^{1/2}\frac{\partial}{\partial \xi_a}|\boldsymbol{\Sigma}_{mk}^{-1}|^{1/2}$, which is the partial derivative of the determinant of the precision matrix and $q_{mk}^{(a)} = -\frac{1}{2}\frac{\partial}{\partial \xi_a}\boldsymbol{y}_{mk}^T\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{y}_{mk}$, which is the partial derivative of the scaled Mahalanobis distance in the Gaussian. In order to expand these expressions further, we make use of the determinant identity $det(\boldsymbol{A}^n) = det(\boldsymbol{A})^n$ [20, Eq. (23)], the determinant derivative $\frac{\partial \det(\boldsymbol{X}^k)}{\partial \boldsymbol{X}} = k\det(\boldsymbol{X}^k)\boldsymbol{X}^{-T}$ [20, Eq. (58)], and the matrix chain rule $\frac{\partial g(\boldsymbol{U})}{\partial v_i} = Tr\left[\left(\frac{\partial g(\boldsymbol{U})}{\partial \boldsymbol{U}}\right)^T\frac{\partial \boldsymbol{U}}{\partial v_i}\right]$ [20, Eq. (137)]. Beginning with $d_{mk}^{(a)}$

$$d_{mk}^{(a)} = |\boldsymbol{\Sigma}_{mk}|^{1/2}Tr\left\{\left(\frac{\partial}{\partial \boldsymbol{\Sigma}_{mk}}|\boldsymbol{\Sigma}_{mk}^{-1/2}|\right)^T\frac{\partial \boldsymbol{\Sigma}_{mk}}{\partial \xi_a}\right\}$$
$$= -\frac{1}{2}Tr\left\{|\boldsymbol{\Sigma}_{mk}|^{-1/2}\boldsymbol{\Sigma}_{mk}^{-1/2}(\boldsymbol{R}_{\xi_a}\boldsymbol{\Omega}_k\boldsymbol{R}^T + \boldsymbol{R}\boldsymbol{\Omega}_k\boldsymbol{R}_{\xi_a}^T)\right\}$$
$$= -Tr\left\{\boldsymbol{\Sigma}_{mk}^{-1/2}\boldsymbol{R}_{\xi_a}\boldsymbol{\Omega}_k\boldsymbol{R}^T\right\}$$

For $q_{mk}^{(a)}$ we need the derivative of the matrix inverse $\frac{\partial \boldsymbol{A}^{-1}(x)}{\partial x} = -\boldsymbol{A}^{-1}\frac{\partial \boldsymbol{A}}{\partial x}\boldsymbol{A}^{-1}$ [20, Eq. (59)].

$$q_{mk}^{(a)} = \frac{\partial}{\partial \xi_a}\boldsymbol{y}_{mk}^T\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{y}_{mk}$$
$$= -\boldsymbol{y}_{mk}^T\boldsymbol{\Sigma}_{mk}^{-1}\frac{\partial \boldsymbol{y}_{mk}}{\partial \xi_a} - \frac{1}{2}\boldsymbol{y}_{mk}^T\frac{\partial \boldsymbol{\Sigma}_{mk}^{-1}}{\partial \xi_a}\boldsymbol{y}_{mk}$$
$$= -\boldsymbol{y}_{mk}^T\boldsymbol{\Sigma}_{mk}^{-1}\frac{\partial \boldsymbol{y}_{mk}}{\partial \xi_a} - \frac{1}{2}\boldsymbol{y}_{mk}^T\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{Z}_a\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{y}_{mk}$$

Where $\boldsymbol{j}_{mk}^{(a)} = -\frac{\partial R}{\partial \xi_a}\nu_k - \frac{\partial t}{\partial \xi_a}$ and $\boldsymbol{Z}_{mk}^{(a)} = \frac{\partial}{\partial \xi_a}\boldsymbol{\Sigma}_{mk} = \frac{\partial}{\partial \xi_a}\boldsymbol{R}\boldsymbol{\Omega}_k\boldsymbol{R}^T = \boldsymbol{R}_{\xi_a}\boldsymbol{\Omega}_k\boldsymbol{R}^T + \boldsymbol{R}\boldsymbol{\Omega}_k\boldsymbol{R}_{\xi_a}^T$. Using these identities and noting that the second term in $q_{mk}^{(a)}$ is of the form $\boldsymbol{x}^T(\boldsymbol{P} + \boldsymbol{P}^T)\boldsymbol{x} = 2\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}$, we arrive at (10)

$$q_{mk}^{(a)} = \boldsymbol{y}_{mk}^T\boldsymbol{\Sigma}_{mk}^{-1}(\boldsymbol{R}_{\xi_a}\nu_k + \boldsymbol{t}_{\xi_a}) + \boldsymbol{y}_{mk}\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{R}\boldsymbol{\Omega}_k\boldsymbol{R}_{\xi_a}^T\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{y}_{mk}$$

In order to derive $D_{mk}^{(ba)}$, we require the trace identity $\frac{\partial Tr(\boldsymbol{X}\boldsymbol{A})}{\partial \boldsymbol{X}} = \boldsymbol{A}^T$ [20, Eq. (100)]

$$D_{mk}^{(ba)} = -\frac{\partial}{\partial \xi_b}Tr\left\{\boldsymbol{\Omega}_k\boldsymbol{R}^T\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{R}_{\xi_a}\right\}$$
$$= -Tr\left\{\left(\frac{\partial Tr\{\boldsymbol{\Omega}_k\boldsymbol{R}^T\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{R}_{\xi_a}\}}{\partial \boldsymbol{R}^T\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{R}_{\xi_a}}\right)^T\frac{\partial \boldsymbol{R}^T\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{R}_{\xi_a}}{\partial \xi_b}\right\}$$
$$= -Tr\{\boldsymbol{\Omega}_k(\boldsymbol{R}_{\xi_b}^T\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{R}_{\xi_a} - \boldsymbol{R}\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{Z}_b\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{R}_{\xi_a}$$
$$+ \boldsymbol{R}\boldsymbol{\Sigma}_{mk}^{-1}\boldsymbol{R}_{\xi_b\xi_a})\}$$

$Q_{mk}^{(ba)}$ follows from $q_{mk}^{(a)}$ using the same rules employed to derive $q_{mk}^{(a)}$.

ACKNOWLEDGMENTS

REFERENCES

[1] Hatem Alismail, Brett Browning, and Simon Lucey. Direct visual odometry using bit-planes. *CoRR*, abs/1604.00990, 2016. URL http://arxiv.org/abs/1604.00990.

[2] Paul J. Besl and Neil D. McKay. Method for registration of 3-d shapes. *IEEE Trans. Pattern Analysis and Machine Intell.*, 14:239–259, February 1992.

[3] J. Blanco. A tutorial on se (3) transformation parameterizations and on-manifold optimization. (012010), October 2010.

[4] N. Boumal, B. Mishra, P. A. Absil, and R. Sepulchre. Manopt, a Matlab toolbox for optimization on manifolds. *Journal of Machine Learning Research*, 15(1): 1455–1459, 2014.

[5] James Conca. Environmental monitoring programs and public engagement for siting and operation of geological repository systems: Experience at the waste isolation pilot plant. In *Geological Repository Systems for Safe Disposal of Spent Nuclear Fuels and Radioactive Waste (Second Edition)*, pages 667–709. Elsevier, 2017.

[6] S. Du, N. Zheng, S. Ying, and J. Liu. Affine iterative closest point algorithm for point set registration. *Pattern Recognition Letters*, 31(9):791–799, 2010.

[7] B. Eckart and A. Kelly. Rem-seg: A robust em algorithm for parallel segmentation and registration of point clouds. In *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.*, pages 4355–4362, Tokyo, Japan, November 2013.

[8] B. Eckart, K. Kim, A. Troccoli, A. Kelly, and J. Kautz. Mlmd: Maximum likelihood mixture decoupling for fast and accurate point cloud registration. In *3D Vision (3DV), 2015 Intl. Conf. on*, pages 241–249, Lyon, France, October 2015.

[9] B. Eckart, K. Kim, A. J Troccoli, A. Kelly, and J. Kautz. Accelerated generative models for 3d point cloud data. In *Proc. of the Intl. Conf. on Computer Vision*, pages 5497–5505, Las Vegas, NV, June 2016.

[10] G. D Evangelidis and R. Horaud. Joint alignment of multiple point sets with batch and incremental expectation-maximization. *IEEE Trans. Pattern Analysis and Machine Intell.*, 40(6):1397–1410, June 2017.

[11] G. Gallego and A. Yezzi. A compact formula for the derivative of a 3-d rotation in exponential coordinates. *Journal of Mathematical Imaging and Vision*, 51(3): 378–384, 2015.

[12] F. S. Grassia. Practical parameterization of rotations using the exponential map. *Journal of graphics tools*, 3:29–48, March 1998.

[13] M. Irani and P. Anandan. About direct methods. In Bill Triggs, Andrew Zisserman, and Richard Szeliski, editors, *Vision Algorithms: Theory and Practice*, pages 267–277, Berlin, Heidelberg, 2000. Springer Berlin Heidelberg.

[14] Mariano Jaimez and Javier Gonzalez-Jimenez. Fast visual odometry for 3-d range sensors. *IEEE Transactions on Robotics*, 31(4):809–822, 2015.

[15] B. Jian and B. C. Vemuri. Robust point set registration using gaussian mixture models. *IEEE Trans. Pattern Analysis and Machine Intell.*, 33(8):1633–1645, December 2010.

[16] K. Kampa, E. Hasanbelliu, and J. C. Principe. Closed-form cauchy-schwarz pdf divergence for mixture of gaussians. In *Proc. of the IEEE Intl. Joint Conf. on Neural Networks*, pages 2578–2585, San Jose, CA, July 31 – August 5 2011.

[17] C. Kerl, J. Sturm, and D. Cremers. Dense visual slam for rgb-d cameras. In *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.*, pages 2100–2106, Tokyo, Japan, November 2013.

[18] J. Minguez, L. Montesano, and F. Lamiraux. Metric-based iterative closest point scan matching for sensor displacement estimation. *IEEE Trans. on Robotics*, 22 (5):1047–1054, October 2006.

[19] 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*, pages 127–136, Basel, Switzerland, October 2011.

[20] K. B. Petersen, M. S. Pedersen, et al. The matrix cookbook, November 2012. URL http://www2.imm.dtu.dk/pubdb/p.php?3274. Version 20121115.

[21] A. Segal, D. Haehnel, and S. Thrun. Generalized-icp. In *Proc. of Robot.: Sci. and Syst.*, page 435, Seattle, WA, USA, June 2009.

[22] Jacopo Serafin and Giorgio Grisetti. Using extended measurements and scene merging for efficient and robust point cloud registration. *Robot. Auton. Syst.*, 92: 91–106, June 2017.

[23] Shobhit Srivastava. Efficient, multi-fidelity perceptual representations via hierarchical gaussian mixture models. Technical Report CMU-RI-TR-17-44, Pittsburgh, PA, August 2017.

[24] Alejandro C Stella-Watts, Christopher P Holstege, Jae K Lee, and Nathan P Charlton. The epidemiology of caving injuries in the united states. *Wilderness & environmental medicine*, 23(3):215–222, 2012.

[25] T. Stoyanov, M. Magnusson, H. Andreasson, and A. J. Lilienthal. Fast and accurate scan registration through minimization of the distance between compact 3d ndt representations. *J. Intl. & Robot. Research*, 31:1377–1393, September 2012.

[26] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers. A benchmark for the evaluation of rgb-d slam systems. In *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.*, Oct. 2012.

[27] T. Tamaki, M. Abe, B. Raytchev, and K. Kaneda. Softassign and em-icp on gpu. In *Proc. of the First Intl. Conf. on Networking and Computing*, pages 179–183, Higashi-Hiroshima, Japan, November 2010.

[28] T Waltham. Sinkhole hazard case histories in karst terrains. *Quarterly Journal of Engineering Geology and Hydrogeology*, 41(3):291–300, 2008.