Journal of Electronic Imaging 28(1), 013026 (Jan/Feb 2019)

Scan registration for underwater mechanical scanning imaging sonar using symmetrical Kullback–Leibler divergence

Min Jiang,[a,b] Sanming Song,[a,*] Fengzhen Tang,[a] Yiping Li,[a] Jian Liu,[a] and Xisheng Feng[a]
[a]Chinese Academy of Sciences, Shenyang Institute of Automation, State Key Laboratory of Robotics, Shenyang, China
[b]University of Chinese Academy of Sciences, Beijing, China

Abstract. Due to its advantages in size and energy consumption, mechanical scanning imaging sonar (MSIS) has been widely used in portable and economic underwater robots to observe the turbid and noisy underwater environment. However, handicapped by the coarseness in spatial and temporal resolution, it is difficult to stitch the scan pieces together into a panoramic map for global understanding. A registration method named symmetrical Kullback–Leibler divergence (SKLD)-distribution-to-distribution (D2D), which models each scan as a Gaussian mixture model (GMM) and evaluates the similarity between two GMMs in a D2D way with the measure defined by SKLD, is proposed to register the scans collected by MSIS. SKLD not only weights the difference between distributions with the prior probability but also increases the numerical stability with the symmetrical constraint in distance measure. Moreover, an approximation strategy is designed to derive a tractable solution for the KLD between two GMMs. Experimental results on the scans that were collected from the realistic underwater environment demonstrate that SKLD-D2D dramatically reduces the computational cost without compromising the estimation precision. © *2019 SPIE and IS&T* [DOI: 10.1117/1.JEI.28.1.013026]

Keywords: underwater robots; mechanical scanning imaging sonar; scan registration; symmetrical Kullback–Leibler divergence.

Paper 180403 received May 9, 2018; accepted for publication Jan. 8, 2019; published online Feb. 7, 2019.

1 Introduction

Exploration and exploitation of the ocean are an everlasting research interest for human beings. To this end, autonomous underwater vehicles (AUV) embarked with different types of sensors have been developed as their assistant.

In the case of clear water, optical sensors (e.g., camera) may be the ideal tools for underwater structure maintenance, topographic reconnaissance, and oceanographic investigations, such as ship hull inspection,[1] fish assemblage monitoring,[2] and underwater archaeological sites recording.[3] However, in many situations, the sea water is filled with a great quantity of plankton, mud, or sand, severely shortening the visual distance of the optical sensors. Therefore, acoustic sensors [e.g., the forward-looking scanning (FLS) sonar], which perceive the surrounding environments by transmitting the ultrasonic wave and receiving the reverberations, are more suitable in this case. They can penetrate the turbid water to observe the immediate environment efficiently.

The mechanical scanning imaging sonar (MSIS)—a type of acoustic sensor—is chosen for this study. It works as follows. The transducer emits a beam in a certain direction. The beam will travel along this direction. If there is an object in its path, the beam would be bounced back. Then, the transducer would hear an echoic response, which will be subsequently used to perceive the object. The information of the resulting echoic response will be quantized into a series of discrete intensity values named "bins," with each bin corresponding to a predefined distance. After that, MSIS changes the sending direction by a predetermined angle, emits another beam, and waits for the corresponding response. This process is repeated until the entire scan sector is covered.[4] The MSIS sonar is compact, low cost, and low energy-consumption, making it suitable for portable robots (e.g., small AUVs) and energy-saving robots (e.g., long-endurance underwater vehicles). For example, Chen et al.[5] installed an MSIS in an AUV to generate a map of the underwater environment with improved localization accuracy. Dong et al.[6] employed an MSIS to localize the remote operated vehicle in the reactor pool of the nuclear power plant by comparing the geometric features extracted from sonar point cloud with the priori map. The manned submersible—"Jiaolong," which is well known for exploring deep sea, takes advantage of the low-frequency MSIS instead of the high-frequency FLS sonar to avoid obstacles.[7]

A large map is often required for the underwater robot to understand the environment, perform the self-localization, and optimize the motion trajectory. However, as the acoustic wave is attenuated rapidly in water, the acoustic sonar can only obtain the scan of a small area in each frame. Therefore, to construct a large map, the scans collected by the sonar at different times need to be stitched up appropriately. This leads to the problem of scan registration, a kind of data-association technique. The scan registration can correct the accumulated error from the dead-reckoning system composed of Doppler velocity log and gyro. It plays a vital role in the simultaneous localization and mapping (SLAM) problem, which is used to construct a spatial map of an unknown environment while simultaneously determining the mobile robot's position relative to this map in the field of mobile robot navigation research. SLAM has been successfully applied to terrestrial robots and aerial robots. However, due to the difficulties in extracting discriminative features

*Address all correspondence to Sanming Song, E-mail: songsanming@sia.cn

from low-resolution acoustic images, little progress has been achieved in building up the underwater SLAM system.

Generally, MSIS sonar possesses two drawbacks. On the one hand, the height information of its scan is irreversibly lost in the imaging process because the three-dimensional (3-D) structure information contained in the propagation space of each fan-shaped acoustic wave is squashed into a one-dimensional signal. On the other hand, the acoustic image will be distorted if the location and pose of the underwater vehicle alter during the scanning process because the mechanically actuated transducer sequentially scans the underwater environment. Therefore, the acoustic scan collected by MSIS has a lower resolution in both spatial range and temporal dimension than other types of sonar, discouraging the application of the traditional registration methods that rely on feature matching,[8,9] especially when MSIS works in the natural underwater environment.

Hence, methods independent of feature extraction have attracted the attention of the underwater community. Castellani et al.[10] adopted the iterative closest point (ICP) method to register 3-D acoustic point clouds. Hernández et al.[11] proposed a probabilistic variant of ICP to solve the distortion problem in registering the point clouds from the MSIS sonar. ICP has been regarded as the golden metric in point cloud registration. However, it has two major limitations: (1) the computation in finding the point correspondences is intensive and (2) it does not consider the uncertainty in point cloud collection.

To conquer the above limitations, people propose to model the points with a probability distribution and estimate the registration parameters analytically. The normal distributions transform (NDT) method proposed by Biber and Strasser[12] divides the reference scan into equal-sized blocks and describes each block as a Gaussian distribution. In the registration step, a likelihood score is obtained when each point in the floating scan is mapped to the nearest Gaussian component in the reference scan. An optimal or suboptimal transformation parameter set can be obtained by maximizing the accumulated likelihood score with a selected optimization strategy. This work is later extended to 3-D case by Magnusson et al.[13] However, NDT still suffers a high computational burden because the Mahalanobis distance has to be calculated between every point in the floating scan and every Gaussian component in the reference scan. Such a registration mode is also named as point-to-distribution (P2D) for short, in the following sections.

Beyond P2D, the distribution-to-distribution (D2D) mode where both the floating and reference scans are modeled as probability functions has been proposed to speed up the registration process further. Tsin and Kanade[14] proposed a kernel-correlation (KC) method to depress the outliers in registering the noisy point clouds. After summing up the kernel correlation between every transformed model point and all the scene points, the cost function is minimized over the transformation space to yield the desired transformation parameters. However, KC method scales poorly with the number of points.[15] Jian and Vemuri[16] modeled each scan with a series of fixed-width spherical Gaussian components and evaluated the discrepancy between Gaussian components with the $\mathcal{L}_2$ norm. Its extension to a variable-width Gaussian kernel function can be found in Ref. 17. However, the $\mathcal{L}_2$ norm is theoretically unjustified for measuring the dissimilarities between probability distribution functions.[18]

In this paper, we propose an MSIS sonar scan registration method in the mode of D2D. First, the sonar scan is filtered with a threshold to get rid of the background information that corresponds to the hollow waters. The remaining point clouds are then grouped into semiequal blocks by the K-means clustering algorithm. Second, each block is modeled with a Gaussian function and a mixture model is obtained. The symmetrical Kullback–Leibler divergence (SKLD) is adopted to measure the similarity between two GMMs. Lastly, the SKLD-based cost function is approximated by a practicable solution and optimized by the Newton gradient descent procedure.

Main contributions of our work are summarized as follows. First, we set a lower segmentation threshold value to keep as much environment information as possible, while traditional MSIS sonar image segmentation method only keeps the highest echo intensity bin in each beam. The advantage of our method is that one can judge whether the registered point clouds are sampled from the same place or not by the overlapping ratio of those point clouds. Our method can thus be used to solve the data association task using the MSIS sonar in an unstructured underwater environment, such as the loop closure detection of underwater SLAM or underwater global localization. Second, our method can solve the registration problem with a large number of points, while traditional featureless registration methods[11,19,20] that are dedicated to MSIS point cloud registration may fail. Since they are mostly the variants of ICP, they also suffer the first limitation of ICP as stated above, despite using either Euclidean distance or Mahalanobis distance. However, we model each point cloud as a GMM as in Ref. 17. Thus, a large number of points are represented by a few compact Gaussian components. Instead of using L2 norm to measure the similarity between two GMMs as in Refs. 16 and 17, we propose to use approximative SKLD, where a closed-form solution to the KL-divergence between two GMMs is approximated by the summation of the KL-divergence of each Gaussian component correspondence. SKLD not only weights the difference between distributions with the prior probability but also increases the numerical stability with the symmetrical constraint in distance measure. Compared to other state-of-the-art registration methods, our method can decrease computational cost while keeping the precision in the case of numerous points being contained in a point cloud.

Two hypotheses are assumed in this paper:

(1) The information of the inertial navigation system (INS) is assumed to be accurate during each scan period. The scan registration is used to correct or alleviate the long-term accumulated error of INS. Note that scans will be discarded if the vehicle rotates with a U-turn, where a large drift error occurs in the INS.

(2) The underwater robot is assumed to travel in a horizontal plane, where only the translation parameters t_x, t_y and rotation parameter θ have to be estimated. The extension of our method to the 3-D case is straightforward.

The paper is organized as follows: the scan registration problem, including scan modeling and similarity measure,

is introduced in Sec. 2. The proposed SKLD-D2D algorithm is presented in Sec. 3. Experimental results validating the SKLD-D2D approach are presented in Sec. 4. Section 5 concludes the paper and suggests potential future works with regard to the SKLD-D2D method.

2 Scan Registration

In this section, we give a brief definition of scan registration problem, a thorough introduction of point cloud modeling, and a concrete description of existing distance measures in scan registration.

2.1 *Scan Registration Problem*

Consider two overlapping scans, floating scan denoted by I_f and reference scan denoted by I_r, are taken by the MSIS at different times. Suppose there exists an unknown transformation matrix T parameterized by the parameter set $\Psi = (t_x, t_y, \theta)^T$, mapping each point in I_f to its counterpart in I_r consistently. Scan registration is to find the optimal transformation parameter set Ψ, which minimizes the objective function $\mathcal{J}$ that measures the differences between two scans modeled by two distributions, respectively,

$$\Theta^* = \arg\min_{\Psi} \mathcal{J}\{\mathcal{G}[\mathbf{x}|T(I_f;\Psi)], \mathcal{G}(\mathbf{x}|I_r)\}, \tag{1}$$

where $\mathcal{G}(\mathbf{x}|I)$ represents the probability distribution of the point cloud extracted from each scan, and $\mathbf{x} = (x, y)^T$ where x, y represent coordinates of a point in Cartesian coordinate system. The global minimum of the objective function is expected to be obtained if the floating scan and the reference scan are perfectly aligned.

2.2 *Sonar Scan Modeling*

In the modeling of sonar scan, the choice of $\mathcal{G}(\mathbf{x}|I)$ is tricky. Generally, it is better to choose a function that is continuously differentiable of at least order two, for example, Gaussian distribution. In this paper, we model the point cloud of a scan as a GMM:

$$\mathcal{G}(\mathbf{x}|I) = \sum_{i=1}^{m} w_i \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_i, \Sigma_i), \tag{2}$$

where m is the number of Gaussian components, w_i is the prior probability, and $\mathcal{N}(\cdot)$ denotes Gaussian distribution. The mean μ_i and covariance matrix Σ_i for each component are

$$\boldsymbol{\mu}_i = \frac{1}{|\mathcal{G}_i|} \sum_{j=1}^{|\mathcal{G}_i|} \mathbf{x}_j, \tag{3}$$

$$\Sigma_i = \frac{1}{|\mathcal{G}_i| - 1} \sum_{j=1}^{|\mathcal{G}_i|} (\mathbf{x}_j - \boldsymbol{\mu}_i)(\mathbf{x}_j - \boldsymbol{\mu}_i)^T, \tag{4}$$

where $|\mathcal{G}_i|$ is the number of points contained in the i'th component $\mathcal{G}_i$. By modeling point cloud as a GMM, the discrete point set can be seen as statistical samples drawn from a piecewise continuous and differentiable compound function.

Traditional methods such as NDT[12] and D2D 2-D NDT[17] divide point cloud into equal-sized blocks and describe each block with a Gaussian function. However, the normal distribution assumption may be severely violated for the blocks along the regional boundaries. To better accommodate the Gaussian assumption, we group the point cloud extracted from each scan into more compact clusters with the K-means clustering method.[21] Note that the prior probability w_i can be considered as a constant because the clusters have very similar size. It will be omitted in the latter discussions.

An example of GMM modeling of the MSIS sonar scans is shown in Fig. 1. To facilitate visualization, each Gaussian component is represented as an ellipse, with the center determined by the mean of the cluster and the radiuses determined by the eigenvalues of the covariance matrix.

2.3 *Similarity Measure*

An intuitive way to measure the distance between two distributions is to sum up the differences of function values in the overlapping domain, as the $\mathcal{L}_2$ norm does. It has been used in Refs. 16, and 17 to measure the dissimilarity of the two probability density functions $\mathcal{G}(\mathbf{x}|I_f)$ and $\mathcal{G}(\mathbf{x}|I_r)$:

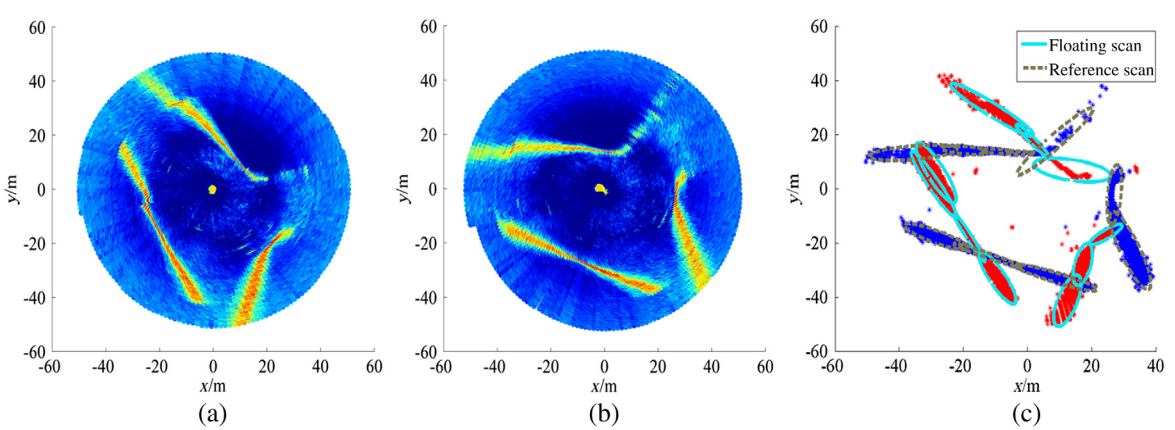

Fig. 1 Gaussian mixture modeling to the MSIS scans. The extracted point clouds from (a) and (b) are presented in red and blue in (c). The compact clusters, generated by the K-means clustering algorithm, are drawn with the ellipses.

$$\mathcal{L}_2[\mathcal{G}(\mathbf{x}|I_f), \mathcal{G}(\mathbf{x}|I_r)] = \int [\mathcal{G}(\mathbf{x}|I_f) - \mathcal{G}(\mathbf{x}|I_r)]^2 \mathrm{d}\mathbf{x}. \tag{5}$$

However, KLD,[22] also known as relative entropy:

$$KL[\mathcal{G}(\mathbf{x}|I_r)||\mathcal{G}(\mathbf{x}|I_f)] = \int \mathcal{G}(\mathbf{x}|I_r) \log \frac{\mathcal{G}(\mathbf{x}|I_r)}{\mathcal{G}(\mathbf{x}|I_f)} \mathrm{d}\mathbf{x}, \tag{6}$$

has been considered to be more appropriate for measuring the difference between the two probability distributions because it weights the difference according to the probability distribution.[23]

In fact, it has been proved in Ref. 24 that KLD is bounded by $\mathcal{L}_2$ norm, indicating that optimizing the objective function based on $\mathcal{L}_2$ norm is equivalent to optimizing the upper and lower bounds of the objective function based on the KLD. Therefore, we propose to construct a sonar scan registration algorithm in the framework of KLD.

3 Symmetric Kullback–Leibler Divergence-Distribution To Distribution

It has been recognized by Said et al.[25] that the symmetric version of KLD, known as Jeffrey divergence

$$\begin{aligned}\mathcal{J}[\mathcal{G}(\mathbf{x}|I_f), \mathcal{G}(\mathbf{x}|I_r)] &= KL[\mathcal{G}(\mathbf{x}|I_f)||\mathcal{G}(\mathbf{x}|I_r)] \\ &\quad + KL[\mathcal{G}(\mathbf{x}|I_r)||\mathcal{G}(\mathbf{x}|I_f)],\end{aligned} \tag{7}$$

can increase the stability and robustness of registration algorithms. In this paper, we propose to register sonar scans by minimizing the symmetrical KL-divergence between their corresponding mixture probability distributions.

3.1 *Approximate Kullback–Leibler Divergence*

There is no closed-form expression of symmetrical KL-divergence between two GMMs. Now that symmetrical KL-divergence is essentially composed of two KLDs, refer to Eq. (7), we derive an approximative closed-form expression for the KLD between two GMMs.

Let $\mathcal{G}(\mathbf{x}|I_f) = \sum_{i=1}^{m} \alpha_i \mathcal{N}_i$ and $\mathcal{G}(\mathbf{x}|I_r) = \sum_{j=1}^{m'} \beta_j \mathcal{N}_j$ be two GMMs, where α_i and β_j are the prior probabilities. The KLD between $\mathcal{G}(\mathbf{x}|I_f)$ and $\mathcal{G}(\mathbf{x}|I_r)$ could be rewritten as

$$\begin{aligned}&KL[\mathcal{G}(\mathbf{x}|I_r)||\mathcal{G}(\mathbf{x}|I_f)] \\ &= \sum_{j=1}^{m'} \beta_j \int \mathcal{N}_j \log \mathcal{G}(\mathbf{x}|I_r) - \sum_{j=1}^{m'} \beta_j \int \mathcal{N}_j \log \mathcal{G}(\mathbf{x}|I_f) \\ &\approx \sum_{j=1}^{m'} \beta_j \int \mathcal{N}_j \log \beta_j \mathcal{N}_j - \sum_{j=1}^{m'} \beta_j \int \mathcal{N}_j \log \alpha_{i_n} \mathcal{N}_{i_n} \\ &= \sum_{j=1}^{m'} \beta_j \left[KL(\mathcal{N}_j|\mathcal{N}_{i_n}) + \log \frac{\beta_j}{\alpha_{i_n}} \right] \\ &\stackrel{\mathrm{def}}{=} KL_{\mathrm{match}}[\mathcal{G}(\mathbf{x}|I_r)||\mathcal{G}(\mathbf{x}|I_f)],\end{aligned} \tag{8}$$

where the approximation assumes that the integral $\int \mathcal{N}_j \log \mathcal{G}(\mathbf{x}|I_f)$ is dominated by the term $\alpha_{i_n} \mathcal{N}_{i_n}$, with i_n labeling the Gaussian component in I_f that is most similar to $\mathcal{N}_j$. The underlying condition is valid when there is little overlap between any two Gaussian components in I_f. Actually, the point clusters, which are generated by the K-means clustering algorithm, have a relatively convex contour and can be well modeled by the semicompact-support Gaussian function. Similar logic can be found in Ref. 26.

Substituting the closed-form expression of the KLD between two Gaussian distributions $\mathcal{N}_i$, $\mathcal{N}_j$ in Eq. (8), the approximative KLD between two GMMs can be given as follows:

$$\begin{aligned}KL_{\mathrm{match}}[\mathcal{G}(\mathbf{x}|I_r)||\mathcal{G}(\mathbf{x}|I_f)] &= \sum_{j=1}^{m'} \frac{1}{2} \left\{ \mathrm{trace}\left(\Sigma_{i_n}^{-1} \Sigma_j \right) \right. \\ &\quad \left. + (\boldsymbol{\mu}_{i_n} - \boldsymbol{\mu}_j)^T \Sigma_{i_n}^{-1} (\boldsymbol{\mu}_{i_n} - \boldsymbol{\mu}_j) - k + \ln \frac{|\Sigma_{i_n}|}{|\Sigma_j|} \right\},\end{aligned} \tag{9}$$

where k is the dimension of vector $\mathbf{x}$.

3.2 *Symmetrical Kullback–Leibler Divergence-Distribution-to-Distribution Minimization*

Substituting Eq. (9) in Eq. (7), we can obtain the objective function as follows:

$$\begin{aligned}\mathcal{F}[\mathcal{G}(\mathbf{x}|I_f), \mathcal{G}(\mathbf{x}|I_r), \Psi] &= \sum_{i=1}^{m} \frac{1}{2} \left\{ \mathrm{trace}\left(\Sigma_{j_n}^{-1} R^T \Sigma_i R \right) - k \right. \\ &\quad \left. + (R\boldsymbol{\mu}_i + \mathbf{t} - \boldsymbol{\mu}_{j_n})^T \Sigma_{j_n}^{-1} (R\boldsymbol{\mu}_i + \mathbf{t} - \boldsymbol{\mu}_{j_n}) + \ln \frac{|\Sigma_{j_n}|}{|R^T \Sigma_i R|} \right\} \\ &\quad + \sum_{j=1}^{m'} \frac{1}{2} \left\{ \mathrm{trace}\left[\left(R^T \Sigma_{i_n} R \right)^{-1} \Sigma_j \right] - k \right. \\ &\quad + (R\boldsymbol{\mu}_{i_n} + \mathbf{t} - \boldsymbol{\mu}_j)^T (R^T \Sigma_{i_n} R)^{-1} (R\boldsymbol{\mu}_{i_n} + \mathbf{t} - \boldsymbol{\mu}_j) \\ &\quad \left. + \ln \frac{|R^T \Sigma_{i_n} R|}{|\Sigma_j|} \right\}.\end{aligned} \tag{10}$$

According to the second assumption in Sec. 1, only the translation vector $\mathbf{t} = (t_x, t_y)^T$ and rotation matrix R have to be estimated, where t_x and t_y are the horizontal and vertical translation, respectively, and R is determined by rotation θ, i.e.,

$$R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{11}$$

Taking the computational cost in searching the global optimum into consideration, we resort to the Newtonian gradient descend method for a suboptimal estimation.

The optimization procedure of the proposed SKLD-D2D method is depicted in Algorithm 1. After the parameter initialization, the algorithm iterates until stopping criteria meets, e.g., $|g| < \delta$, where δ is a very small threshold. In each iteration, every Gaussian component in $\mathcal{G}(\mathbf{x}|I_f)$ is queried against all components in $\mathcal{G}(\mathbf{x}|I_r)$ to determine the Gaussian correspondence with the minimum KLD measure and vice versa. Driven by the increment of gradient vector g and Hessian matrix H, i.e., $W_g(\Psi, \mathcal{N}_i, \mathcal{N}_j)$ and $W_H(\Psi, \mathcal{N}_i, \mathcal{N}_j)$, the parameters move gradually toward a local optimum. The gradient vector and Hessian matrix can be found in Sec. 6 and Appendix A.

Algorithm 1 Scan registration by the SKLD-D2D algorithm.

Input:

Floating scan I_f, reference scan I_r, initial parameter set Ψ_0, learning rate η

Output:

Transformation parameter Ψ

Filter scans with an intensity threshold γ.
Remove outliers and group the point clouds with K-means clustering.
Fit each cluster with a Gaussian distribution, and obtain $\mathcal{G}(\boldsymbol{x}|I_f)$ and $\mathcal{G}(\boldsymbol{x}|I_r)$.
$\Psi \leftarrow \Psi_0$, $\text{iter}_{\text{cur}} \leftarrow 0$
while $|g| > \delta$ and $\text{iter}_{\text{cur}} < \text{iter}_{\text{MAX}}$ **do**
  $g \leftarrow 0$, $H \leftarrow 0$
  for $\mathcal{N}_i \in \mathcal{G}(\boldsymbol{x}|I_f)$ **do**
    $\mathcal{N}_i' \leftarrow T(\Psi, \mathcal{N}_i)$, $d = [\,]$
    for $\mathcal{N}_j \in \mathcal{G}(\boldsymbol{x}|I_r)$ **do**
      $d = [d; \text{KLD}(\mathcal{N}_i' || \mathcal{N}_j)]$
    end for
    $d' = \text{sort}(d)$
    $j_n = d'(1)$
    $g \leftarrow g + W_g(\Psi, \mathcal{N}_i, \mathcal{N}_{j_n})$
    $H \leftarrow H + W_H(\Psi, \mathcal{N}_i, \mathcal{N}_{j_n})$
  end for
  for $\mathcal{N}_j \in \mathcal{G}(\boldsymbol{x}|I_r)$ **do**
    $d = []$
    for $\mathcal{N}_i \in \mathcal{G}(\boldsymbol{x}|I_f)$ **do**
      $\mathcal{N}_i' \leftarrow T(\Psi, \mathcal{N}_i)$
      $d = [d; \text{KLD}(\mathcal{N}_j || \mathcal{N}_i')]$
    end for
    $d' = \text{sort}(d)$
    $i_n = d'(1)$
    $g \leftarrow g + W_g(\Psi, \mathcal{N}_j, \mathcal{N}_{i_n})$
    $H \leftarrow H + W_H(\Psi, \mathcal{N}_j, \mathcal{N}_{i_n})$
  end for
  $\Delta\Psi = -\eta \cdot H^{-1} \cdot g$
  $\Psi = \Psi + \Delta\Psi$
  $\text{iter}_{\text{cur}} = \text{iter}_{\text{cur}} + 1$
end while
return Ψ

4 Experiments

In this section, we study the performance of the proposed registration method using the sonar dataset provided by Ribas et al.[9] This dataset was collected using a Tritech Miniking sonar when the AUV traversed through the Fluvia Nautic abandoned marina near St. Pere Pescador on the Costa Brava. It includes ~230 scans.

In general, this dataset meets the assumption (2) presented in Sec. 1. Although there is no valid depth information contained in the dataset, there are three reasons we can safely make this assumption for the dataset. First, a differential global positioning system (DGPS) unit mounted in a buoy was attached at the top of the vehicle during the experiments. It is well known to the underwater robot community that the GPS signal cannot penetrate the water surface. However, the signals outputted by the DGPS were valid all the time during the data collection, suggesting that the underwater robot was traveling roughly parallel to the surface of the sea. Second, the vertical beamwidth of the used sonar is 40 deg. The AUV could have some vertical displacements during the data collection while not strongly violating the assumption (2). Finally, from the experimental setting described in Ref. 27, it can be seen that the dataset was collected from planar environment or those composed only of vertical walls that had a constant section independent of the vehicle's depth.

According to the assumption (1) in Sec. 1, the location and pose of each scan line can be derived from the dead-reckoning system. Then, all the scan lines during a cycle can be projected from the polar coordinate into the same Cartesian coordinate with the corresponding transformation matrix, forming a circular sonar image.[19] Refer to Fig. 1, for example.

The registration methods compared with our method include ICP, KC, NDT, and D2D 2D-NDT. To validate the feasibility of utilizing the intuitive KLD for the MSIS sonar scans, we also estimated the transformation parameters with a method named KLD-D2D, where symmetrical KL-divergence in the proposed SKLD-D2D method was replaced by the KLD. All the algorithms involved in this paper were implemented in MATLAB 2014a on a 3.2 GHz Intel Core i5 processor with 4 GB of RAM. No acceleration strategies were used here for all methods.

The parameters of SKLD-D2D were set as follows: the transformation parameters, including t_x, t_y, and θ were initialized to be zeros, respectively. The maximum number of iterations iter_{MAX} and the stopping criterion δ were set as 30 and 10^{-6}, respectively. The learning rate was 1.1. In the image filtering step, as stated in Sec. 1, we high-pass filtered the image using a lower threshold, i.e., $\gamma = 80$. The influence of the threshold γ on the performance of SKLD-D2D is studied and explained in Sec. 4.1. Twenty bins closest to the transducer were discarded because they were false echoes produced by a transient "ringing" of the transducer head.[28] The blobs that contained less than three points were considered as scattering noises and removed directly. The clustering number K in the K-means clustering method was proportional to the point number N:

$$K = \left\lceil \frac{N}{C} \right\rceil, \tag{12}$$

where C was empirically set to be 120, and the performance of SKLD-D2D was not very sensitive to this parameter. Note that the number of Gaussian components is a tradeoff between the computational cost and the estimation precision. A larger number of Gaussian components will lead to improved registration accuracy but at the cost of increased computational complexity.

Multiple candidate values of each hyperparameter of other methods were examined, and the one with the best performance was selected. The strategy that is gradually dividing point cloud into finer grids was adopted for the NDT and D2D 2-D NDT methods to ensure the convergence.[13] These hyperparameters were kept the same for all the experiments.

4.1 *Influence of the Parameter γ*

We first evaluated the influence of the parameter γ on the performance of our proposed SKLD-D2D registration method. To quantitatively study the performance of our method, we artificially created ground-truth transformation parameters. The translation parameters were randomly generated from a uniform distribution in $[-4,4]$ m. The rotation angle was randomly sampled from a uniform distribution in $[-10,10]$ deg.

The floating scans were generated by using different segmentation threshold γ, as shown in the first column of Table 1, to segment the same MSIS sonar image. Each floating scan was transformed into 100 different reference scans with transformation parameters generated following the above process. The registration accuracy was measured by the average and standard deviation of errors between the estimated transformation parameters and the predefined values, as given in Table 1.

As shown in the first column and sixth column of Table 1, the lower parameter γ is set to be, the more points will be kept. As the number of points increase, the computation speed of the registration algorithm is bound to be slowed down, as one can see in the fifth column of Table 1. In fact, the choice of the parameter γ is a tradeoff between the amount of environment information being kept and the computational burden of the registration algorithm.

In a word, the setting of parameter γ mainly affects the speed, not the precision of the SKLD-D2D algorithm.

4.2 *Known Scan Pairs*

It is unfair to evaluate the effectiveness of a registration method by comparing its estimated transformation parameters from a pair of images with the corresponding physical movement information of the underwater vehicle measured by the INS. There exists accumulated error in the INS. Consequently, the physical movement information of the underwater vehicle cannot work as the ground-truth transformation parameters.

In the second experiment, the way to test the performance of a registration algorithm was as the same as the Sec. 4.1. Here, 108 different scans in the dataset were randomly chosen as the floating scans. Each scan was transformed into 10 different reference scans with the generated transformation parameters. The resulted 1080 scan pairs were used to test the performance of different methods, including ICP, KC, NDT, D2D 2D-NDT, KLD-D2D, and SKLD-D2D. The performances of all the methods were provided in Table 2. An example of the estimated transformation parameters along with the ground truth is presented in Table 3.

As one can see in Table 2, the registration error of our method is comparable to that of ICP, NDT, and D2D 2D-NDT, but the running time of our method is much shorter than these three methods, indicating that our method is more efficient. Our method outperformed KC in terms of both the registration precision and the computational time. Even though our method is slightly slower than KLD-D2D, the registration error of our method is much smaller than that of KLD-D2D method. The standard deviation of our method is consistently smaller than that of the KLD-D2D method, indicating that our method is more robust than KLD-D2D method. The inclusion of the symmetric constraint not only improves the estimation precision but also increases the practicability. However, it also slightly increases the computational time, as two KL divergences need to be calculated in the SKLD-D2D method every time, only one KLD in the KLD-D2D method needs to be computed. Therefore, taking both registration precision and computational time into consideration, our method is superior to other scan registration methods.

4.3 *Register a Scan Pair*

In this experiment, we evaluate the effectiveness of our method on registration of two scans in the real-world dataset to form a local map. The crispness measure from Ref. 29 was adopted to evaluate the registration precision of different methods. Explicitly, we divided the registered scans into equal-sized voxels and counted the number of occupied units, with a lower value indicating a better alignment. Formally, it can be defined as

Table 1 The registration performance of SKLD-D2D influenced by the parameter γ settings.

γ	Δt_x	Δt_y	$\Delta\theta$	Time (s)	Number
100	0.0344 ± 0.2283	-0.0092 ± 0.2934	-0.1519 ± 2.8324	0.19 ± 0.02	1210
90	0.0339 ± 0.1284	0.0492 ± 0.1422	-0.0618 ± 1.5622	0.53 ± 0.06	2468
80	0.0322 ± 0.0975	0.0248 ± 0.0903	-0.0052 ± 0.8062	1.12 ± 0.18	3641
70	0.0188 ± 0.0588	0.0120 ± 0.0787	0.0468 ± 0.4593	1.88 ± 0.27	4834
60	0.0224 ± 0.0563	0.0055 ± 0.0694	0.1315 ± 0.4369	3.09 ± 0.44	6044
50	0.0176 ± 0.0430	0.0052 ± 0.0586	0.1091 ± 0.4439	5.22 ± 0.83	7717

Table 2 The average and standard deviation of errors and runtime of different registration methods.

Error&Runtime	Δt_x	Δt_y	$\Delta\theta$	Runtime (s)
ICP	0.0001 ± 0.5525	0.0166 ± 0.2820	0.0135 ± 1.2194	4.92 ± 4.79
KC	0.0234 ± 0.6986	-0.0407 ± 0.7849	-0.0885 ± 2.9451	28.16 ± 10.77
NDT	0.0311 ± 0.5922	-0.0021 ± 0.3993	-0.0379 ± 0.5323	383.94 ± 147.04
D2D 2-D NDT	-0.0019 ± 0.7760	0.0135 ± 0.3711	0.0132 ± 1.1207	8.13 ± 2.70
KLD-D2D	-0.0466 ± 1.9528	-0.0283 ± 1.2921	-0.1559 ± 4.8492	0.81 ± 0.51
SKLD-D2D	-0.0063 ± 0.6232	-0.0037 ± 0.4121	-0.0148 ± 1.0082	1.38 ± 0.94

Table 3 Estimate motion parameters for a known sonar image pair with different registration methods.

Estimation	Truth	ICP	KC	NDT	D2D 2-D NDT	KLD-D2D	SKLD-D2D
t_x	2.881	2.181	3.601	2.589	2.39	2.469	2.818
t_y	0.113	−0.154	−0.285	−0.064	−0.043	−0.008	−0.016
θ	3.396	2.377	3.396	3.231	3.139	3.161	3.426

$$\text{crispness} = \sum_{i=1}^{|v|} \mathcal{S}_{\text{occupied}}(\text{voxel}_i), \tag{13}$$

where $|v|$ is the total number of voxels. A voxel is considered to be occupied, i.e., $\mathcal{S}_{\text{occupied}}(\text{voxel}_i) = 1$, if at least one point falls into it. The voxel size was set to be 0.5 m in this paper.

An example for registering two scans with different methods are shown in Fig. 2. To ease the visualization, the floating scan is plotted in green, the reference scan is in red, while those points falling into the overlapped region are plotted in blue. The crispness indexes for different methods are shown in Table 4.

From Table 4, we can see that the crispness index of the proposed SKLD-D2D method is apparently smaller than other methods, which is consistent with the observation that the blue points in the synthesized image of Fig. 2(g) are obviously more than other methods. It not only demonstrates that the estimation performance of SKLD-D2D is more precise than other methods but also indicates that the crispness index is a feasible quantitative measure for evaluating the parameter estimation precision of different methods. Moreover, all methods are able to reduce the crispness index of the map generated by the original INS, which can be again validated by the fact that the blue points remained in the synthesized images generated by different methods are more than that generated by the INS [Fig. 2(a)]. It shows that all the registration methods listed here are capable of correcting the accumulated error from the dead-reckoning system.

4.4 *Register Consecutive Sonar Sequence*

The estimation errors among different registration algorithms are too insignificant to be used for performance comparison. However, the high-frequency information, such as edges and contours, becomes blurred if the consecutive scans are registered to form a panoramic view, due to the rapid accumulation of local errors.[30] To assess the abilities of different methods in reducing the accumulation error, we registered five consecutive sonar scans to form a panoramic map. A short sonar scan sequence, ranging from 21st to 25th, was selected for registration because the error from INS of this sequence was more severe than other segments.

The panoramic images generated by different methods are plotted in Fig. 3 and the corresponding crispness measure are displayed in Table 5. The KC method generates a panoramic image with the smallest crispness index. A closer observation to the panoramic image plotted in Fig. 3(c) shows that the point clouds belong to the same object, like the wall of the tank, have a higher overlapping ratio.

However, when facing outliers, the parameter involving outlier dealing of ICP and KC methods has to be carefully tuned. The KC method is very sensitive to the bandwidth. The crispness index of the panoramic map generated by the KC method presented in Table 5 was obtained by optimizing the bandwidth to be 3. A larger crispness index, i.e., 3452 was obtained when the bandwidth was kept as 2, the same as the previous two experiments. Much endeavor needs to be made to select a proper kernel bandwidth when using KC methods. The ICP method includes the maximum distance of point correspondence to filter outliers. This method is very sensitive to the setting of maximum distance parameter. With the optimal maximum distance, the crispness index of the corresponding panoramic map, i.e., 3183 was obtained. However, with the default value, the crispness index grew to 3549.

From Table 5, we can see that our method is comparable to NDT and D2D 2-D NDT in terms of crispness index. But many efforts have been taken to obtain the proper grid sizes for NDT and D2D 2-D NDT methods. For this experiment,

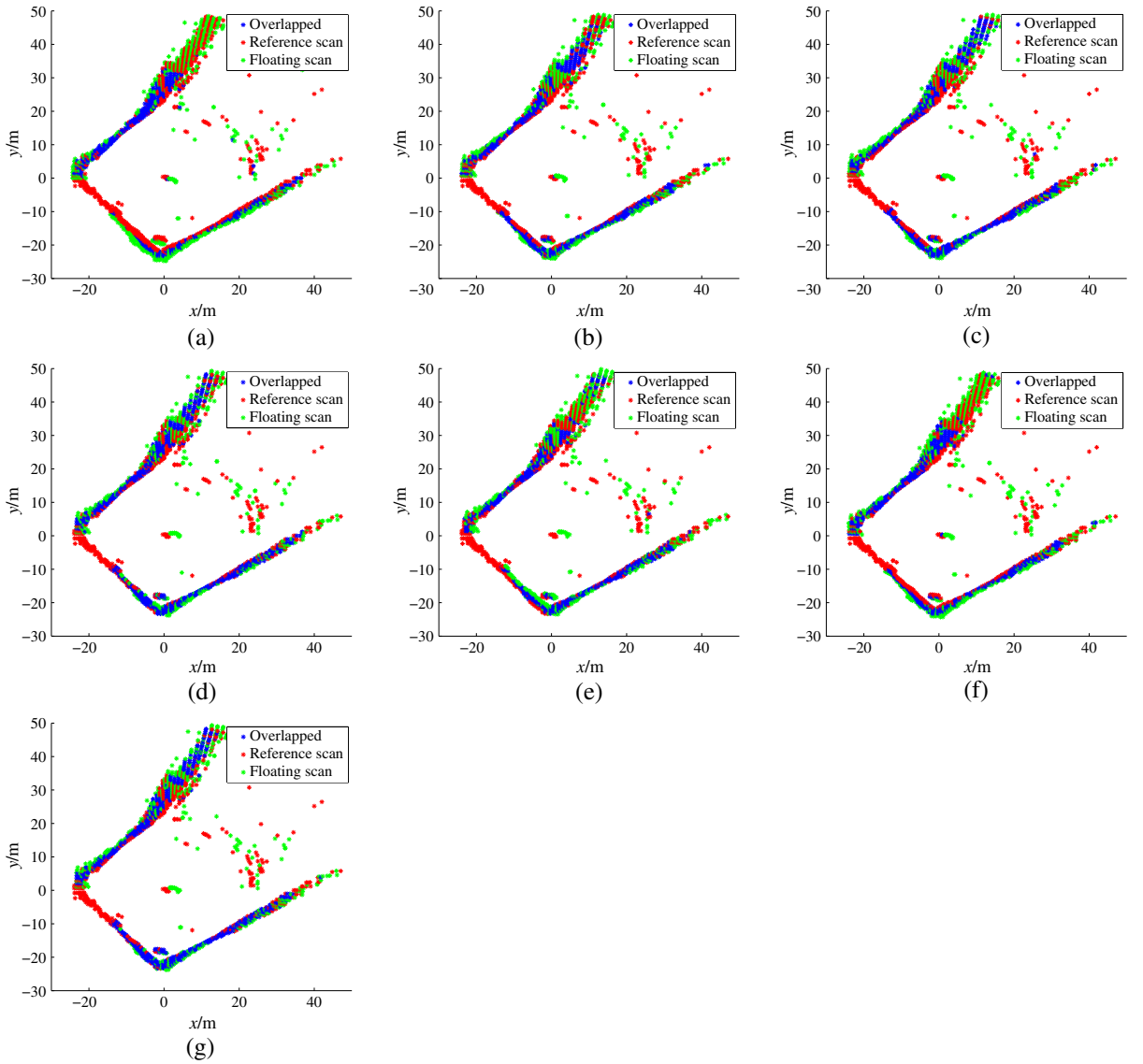

Fig. 2 Register two consecutive scans into a local map by (a) INS, (b) ICP, (c) KC, (d) NDT, (e) D2D 2-D NDT, (f) KLD-D2D, and (g) SKLD-D2D. The reference scan and floating scan are shown in red and green, respectively; the correct alignment is shown in blue.

Table 4 Crispness measurement for scan registration in Fig. 2.

Methods	INS	ICP	KC	NDT	D2D 2D-NDT	KLD-D2D	SKLD-D2D
Crispness	1570	1455	1431	1416	1488	1511	1413

the optimal grid sizes for NDT and D2D 2-D NDT changed to be {25, 20, 10} and {25,20,10,5,2,1}, respectively. It is mainly for that a larger initial deviation is more likely to trap the registration method in a local minimum. For example, when registering the 24th to the 23rd sonar scan with NDT, where the initial rotation parameter was far away from the ground truth, i.e., 27 deg, the estimation error was 14.62 degrees with the grid size setting the same as the previous two experiments, while it reduced to 0.17 deg with the optimized grid sizes. Smaller grid size leads to more Gaussian components, generating more local minima and preventing the gradient descend procedure converging to the global minimum. The registration precision is sensitive to the grid size setting, indicating that the performance of NDT depends on the prior knowledge.

The crispness index of our proposed SKLD-D2D is smaller than that of KLD-D2D in this experiment, demonstrating again that the symmetrical constraint to the KLD is beneficial to find a better parameter estimation. The average estimation time of SKLD-D2D is about 1.47 s. The scanning period of the MSIS sonar used in this paper is about 14 s. The former is far smaller than the latter, demonstrating that the proposed registration method is feasible for online data association with high estimation precision.

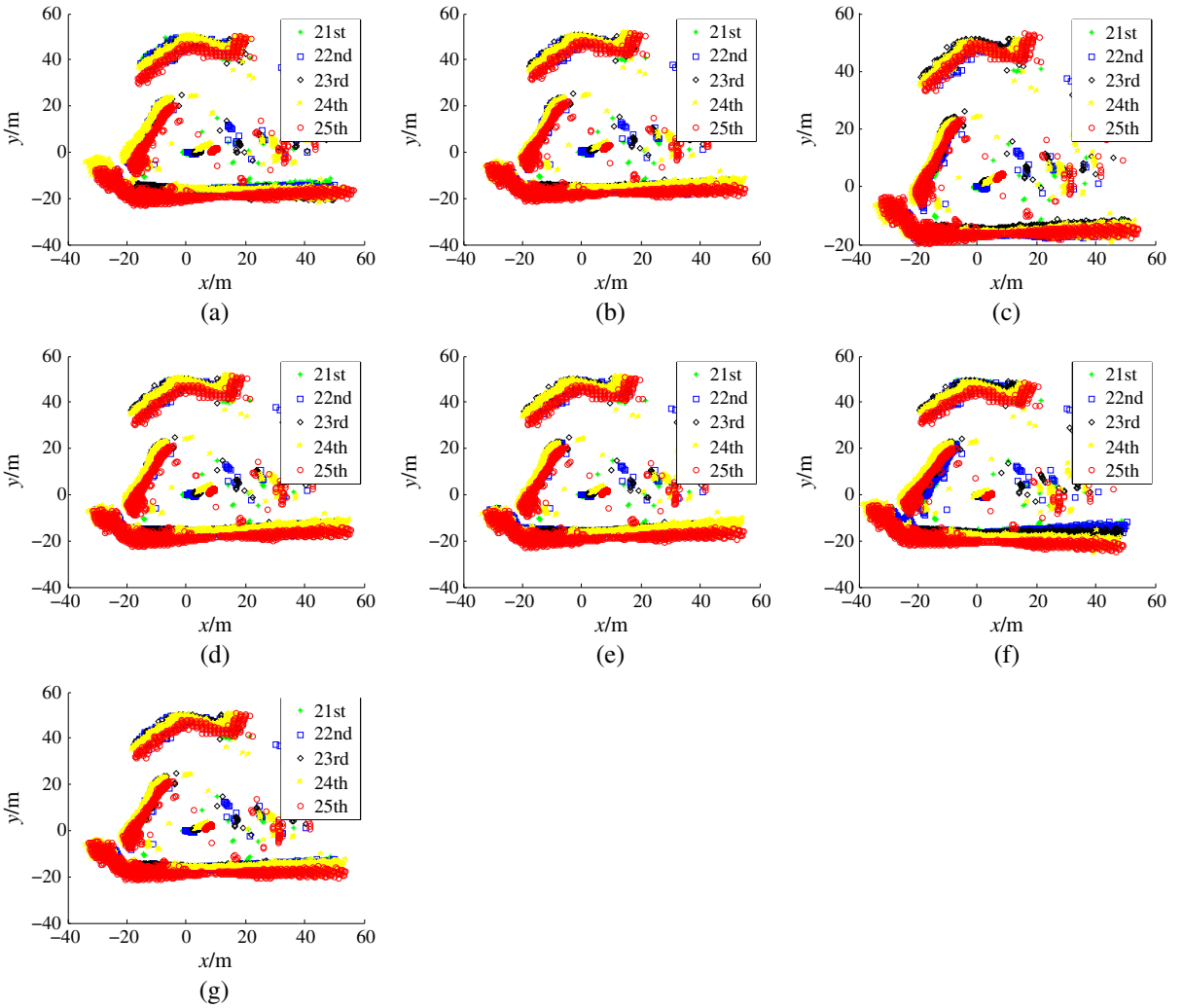

Fig. 3 Register five consecutive scans into a local map by (a) INS, (b) ICP, (c) KC, (d) NDT, (e) D2D 2D-NDT, (f) KLD-D2D, and (g) SKLD-D2D.

Table 5 Crispness measurement for scan registration in Fig. 3.

Methods	INS	ICP	KC	NDT	D2D 2D-NDT	KLD-D2D	SKLD-D2D
Crispness	3475	3183	3099	3258	3229	3904	3225

5 Conclusion

In this article, we propose a scan registration method termed as SKLD-D2D to register the underwater MSIS scans that have an ultra-low resolution both in spatial range and temporal dimension. Contrary to the point-to-point methods such as ICP or point-to-distribution methods such as NDT, SKLD-D2D works in the distribution-to-distribution framework, significantly reducing the computational cost in scan registration. Both the floating and reference scans are grouped into compact blobs using K-means clustering algorithm and mathematically described as Gaussian mixture models (GMMs). To better measure the distance between the probability distributions and increase the robustness, the symmetrical KL-divergence is adopted to construct the global cost function for scan registration. An approximation scheme is designed to solve the problem of deriving a tractable solution for KLD between two GMMs. The registration parameters are obtained by minimizing the cost function through the gradient descent algorithm. Through the sonar scan registration experiments, we can safely claim that the proposed SKLD-D2D greatly reduces the computational cost, however with comparable performance with the well-known point-to-point or point-to-distribution methods in terms of the accuracy and robustness.

Our future work will be devoted to improving the registration performance of the SKLD-D2D method further. More suitable clustering method will be developed to improve the registration accuracy of the SKLD-D2D method. Better parameter initialization procedure will be designed to avoid the divergence of the method, thus increasing the robustness

of SKLD-D2D. Moreover, we intend to incorporate the proposed method into an SLAM system to construct a global map for an AUV.

6 Appendix A: Gradient Vector and Hessian Matrix of the Objective Function

The increment to the parameter vector Ψ in the Newtonian gradient descend method can be written as

$$\Delta\Psi = -\eta \frac{\nabla\mathcal{F}(\Psi)}{\nabla\mathcal{F}^2(\Psi)}, \tag{14}$$

where $\nabla\mathcal{F}(\Psi)$ and $\nabla\mathcal{F}^2(\Psi)$ are the gradient vector and Hessian matrix of the objective function Eq. (10), respectively, η is the learning rate. Because symmetrical KL-divergence is composed of two KLDs, both the gradient vector $\nabla\mathcal{F}(\Psi)$ and Hessian matrix $\nabla\mathcal{F}^2(\Psi)$ consist of two terms:

$$\begin{aligned}\nabla\mathcal{F}(\Psi) &= \sum_{i=1}^{m} W_g(\Psi,\mathcal{N}_i,\mathcal{N}_{j_n}) + \sum_{j=1}^{m'} W_g(\Psi,\mathcal{N}_j,\mathcal{N}_{i_n}),\\ \nabla\mathcal{F}^2(\Psi) &= \sum_{i=1}^{m} W_H(\Psi,\mathcal{N}_i,\mathcal{N}_{j_n}) + \sum_{j=1}^{m'} W_H(\Psi,\mathcal{N}_j,\mathcal{N}_{i_n}).\end{aligned} \tag{15}$$

The detailed calculations of the second summands in the above two equations are given as follows. Note that the computations of the first summands are similar to that of the second summands. To simplify the presentation, the calculations of the first summands will not be given.

The gradient vector $W_g(\Psi,\mathcal{N}_j,\mathcal{N}_{i_n})$ is given as

$$W_g(\Psi,\mathcal{N}_j,\mathcal{N}_{i_n}) = \begin{pmatrix} \mathbf{j}_1^T B\boldsymbol{\mu}_{i_n j} \\ \mathbf{j}_2^T B\boldsymbol{\mu}_{i_n j} \\ \mathrm{trace}(BC_1 + C\Sigma_j) + (j_3^T B + \mu_{i_n j}^T C)\mu_{i_n j} \end{pmatrix}, \tag{16}$$

where

$$\begin{aligned} \boldsymbol{\mu}_{i_n j} &= R\boldsymbol{\mu}_{i_n} + \mathbf{t} - \boldsymbol{\mu}_j, & B &= R^T\Sigma_{i_n}^{-1}R,\\ \mathbf{j}_1 &= \frac{\partial\boldsymbol{\mu}_{i_n j}}{\partial t_x}, & C &= \frac{\partial R^T}{\partial\theta}\Sigma_{i_n}^{-1}R,\\ \mathbf{j}_2 &= \frac{\partial\boldsymbol{\mu}_{i_n j}}{\partial t_y}, & C_1 &= \frac{\partial R^T}{\partial\theta}\Sigma_{i_n}R\\ \mathbf{j}_3 &= \frac{\partial\boldsymbol{\mu}_{i_n j}}{\partial\theta} = \frac{\partial R}{\partial\theta}\boldsymbol{\mu}_{i_n}. \end{aligned} \tag{17}$$

The elements of Hessian matrix

$$W_H(\Psi,\mathcal{N}_j,\mathcal{N}_{i_n}) = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}, \tag{18}$$

are given by the following nine equations:

$$\begin{aligned} h_{11} &= \mathbf{j}_1^T B\mathbf{j}_1,\\ h_{12} &= \mathbf{j}_1^T B\mathbf{j}_2,\\ h_{13} &= \mathbf{j}_1^T(D\boldsymbol{\mu}_{i_n j} + B\mathbf{j}_3),\\ h_{21} &= \mathbf{j}_2^T B\mathbf{j}_1,\\ h_{22} &= \mathbf{j}_2^T B\mathbf{j}_2,\\ h_{23} &= \mathbf{j}_2^T(D\boldsymbol{\mu}_{i_n j} + B\mathbf{j}_3),\\ h_{31} &= \mathbf{j}_3^T B\mathbf{j}_1 + \mathbf{j}_1^T(C + C^T)\boldsymbol{\mu}_{i_n j},\\ h_{32} &= \mathbf{j}_3^T B\mathbf{j}_2 + \mathbf{j}_2^T(C + C^T)\boldsymbol{\mu}_{i_n j},\\ h_{33} &= \mathrm{trace}[DC_1 + B(E_1 + G_1) + (E + G)\Sigma_j] + \mathbf{j}_{33}^T B\boldsymbol{\mu}_{i_n j}\\ &\quad + 2\mathbf{j}_3^T(C + C^T)\boldsymbol{\mu}_{i_n\mathbf{j}} + \mathbf{j}_3^T B\mathbf{j}_3 + \boldsymbol{\mu}_{i_n\mathbf{j}}^T(E + G)\boldsymbol{\mu}_{i_n\mathbf{j}}, \end{aligned} \tag{19}$$

where

$$\begin{aligned} E &= \frac{\partial^2 R^T}{\partial\theta^2}\Sigma_{i_n}^{-1}R, \quad G = \frac{\partial R^T}{\partial\theta}\Sigma_{i_n}^{-1}\frac{\partial R}{\partial\theta},\\ E_1 &= \frac{\partial^2 R^T}{\partial\theta^2}\Sigma_{i_n}R, \quad G_1 = \frac{\partial R^T}{\partial\theta}\Sigma_{i_n}\frac{\partial R}{\partial\theta},\\ D &= \frac{\partial B}{\partial\theta} \qquad \mathbf{j}_{33} = \frac{\partial^2\boldsymbol{\mu}_{i_n j}}{\partial\theta^2} = \frac{\partial^2 R}{\partial\theta^2}\boldsymbol{\mu}_{i_n},\\ &= \frac{\partial R^T}{\partial\theta}\Sigma_{i_n}^{-1}R + R^T\Sigma_{i_n}\frac{\partial R}{\partial\theta}. \end{aligned} \tag{20}$$

The details of the involved partial derivatives are given in Eq. (21):

$$\begin{aligned} \frac{\partial\boldsymbol{\mu}_{i_n j}}{\partial t_x} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \frac{\partial R}{\partial\theta} = \begin{pmatrix} -\sin(\theta) & -\cos(\theta) \\ \cos(\theta) & -\sin(\theta) \end{pmatrix},\\ \frac{\partial\boldsymbol{\mu}_{i_n j}}{\partial t_y} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \frac{\partial^2 R}{\partial\theta^2} = \begin{pmatrix} -\cos(\theta) & \sin(\theta) \\ -\sin(\theta) & -\cos(\theta) \end{pmatrix}. \end{aligned} \tag{21}$$

Acknowledgments

Research supported by the National Natural Science Foundation of China (No. 41506121), the National Key Research and Development Program of China (Nos. 2016YFC0300801, 2017YFC030 5901, and 2016YFC0301601), the Strategic Priority Program of the Chinese Academy of Sciences (Nos. XDA13030205 and XDA11040103), the State Key Laboratory of Robotics of China (No. 2017-Z010), the Natural Science Foundation of Jiangsu Province, China (No. BK20170558), the project of "R&D Center for Underwater Construction Robotics," funded by the Ministry of Ocean and Fisheries (MOF) and Korea Institute of Marine Science & Technology Promotion (KIMST), Korea (No. PJT200539), the Public Science and Technology Research Funds projects of ocean (No. 201505017). The authors report grants from the National Natural Science Foundation of China, grants from National Key Research and Development Program of China, grants from Strategic Priority Program of the Chinese Academy of Sciences, grants from State Key Laboratory of Robotics of China, grants from Natural Science Foundation of Jiangsu Province, China, grants from Ministry of Ocean and Fisheries (MOF) and Korea

Institute of Marine Science & Technology Promotion (KIMST), Korea, the Public science and technology research funds projects of ocean, during the conduct of the study.

References

1. P. Ozog et al., "Long-term mapping techniques for ship hull inspection and surveillance using an autonomous underwater vehicle," *J. Field Rob.* **33**, 265–289 (2016).
2. B. J. Boom et al., "A research tool for long-term and continuous analysis of fish assemblage in coral-reefs using underwater camera footage," *Ecol. Inf.* **23**, 83–97 (2014).
3. J. McCarthy and J. Benjamin, "Multi-image photogrammetry for underwater archaeological site recording: an accessible, diver-based approach," *J. Maritime Archaeol.* **9**, 95–114 (2014).
4. J. Solari Franco et al., "Artificial potential fields for the obstacles avoidance system of an auv using a mechanical scanning sonar," in *3rd IEEE/OES S. Am. Int. Symp. on Oceanic Eng. (SAISOE), Buenos Aires, Argentina*, IEEE, pp. 1–6 (2016).
5. L. Chen et al., "Improving localization accuracy for an underwater robot with a slow-sampling sonar through graph optimization," *IEEE Sens. J.* **15**, 5024–5035 (2015).
6. M. Dong, W. Chou, and B. Fang, "Underwater matching correction navigation based on geometric features using sonar point cloud data," *Sci. Program.* **2017**, 1–10 (2017).
7. W. Cui, "Development of the jiaolong deep manned submersible," *Marine Technol. Soc. J.* **47**, 37–54 (2013).
8. B. He et al., "Autonomous navigation for autonomous underwater vehicles based on information filters and active sensing," *Sensors* **11**, 10958–10980 (2011).
9. D. Ribas et al., "Underwater slam in man-made structured environments," *J. Field Rob.* **25**, 898–921 (2008).
10. U. Castellani et al., "Efficient on-line mosaicing from 3D acoustical images," in *OCEANS '04. MTTS/IEEE Techno-Ocean '04*, Kobe, Japan, IEEE, Vol. 2, pp. 670–677 (2004).
11. E. Hernández et al., "Probabilistic sonar scan matching for an AUV," in *IEEE/RSJ Int. Conf. on Intell. Rob. Syst.*, St. Louis, MO, USA, IEEE, pp. 255–260 (2009).
12. P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *Proc. 2003 IEEE/RSJ Int. Conf. on Intell. Rob. and Syst. (IROS 2003) (Cat. No. 03CH37453)*, Las Vegas, NV, USA, IEEE, Vol. 3, pp. 2743–2748 (2003).
13. M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3D-NDT," *J. Field Rob.* **24**, 803–827 (2007).
14. Y. Tsin and T. Kanade, "A correlation-based approach to robust point set registration," in *Comput. Vision—ECCV 2004*, T. Pajdla and J. Matas, Eds., Springer-Verlag Berlin Heidelberg, Prague, Czech Republic, Vol. 3023, pp. 558–569 (2004).
15. J. Straub et al., "Efficient global point cloud alignment using Bayesian nonparametric mixtures," in *IEEE Conf. on Comput. Vision and Pattern Recognit. (CVPR)*, Honolulu, HI, USA, IEEE, pp. 2403–2412 (2017).
16. B. Jian and B. C. Vemuri, "Robust point set registration using Gaussian mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**, 1633–1645 (2011).
17. T. Stoyanov et al., "Fast and accurate scan registration through minimization of the distance between compact 3D-NDT representations," *Int. J. Rob. Res.* **31**, 1377–1393 (2012).
18. F. Briggs, R. Raich, and X. Z. Fern, "Audio classification of bird species: A statistical manifold approach," in *Ninth IEEE Int. Conf. on Data Mining*, Miami, FL, USA, IEEE, pp. 51–60 (2009).
19. A. Mallios et al., "Scan matching slam in underwater environments," *Auton. Rob.* **36**, 181–198 (2014).
20. A. Burguera, Y. González, and G. Oliver, "The UspIC: performing scan matching localization using an imaging sonar," *Sensors* **12**(6), 7855–7885 (2012).
21. A. Das and S. L. Waslander, "Scan registration with multi-scale k-means normal distributions transform," in *IEEE/RSJ Int. Conf. on Intell. Rob. and Syst.*, Vilamoura, Portugal, IEEE, pp. 2705–2710 (2012).
22. S. Kullback, *Information Theory and Statistics*, Dover Publications Inc., Mineola, New York (1968).
23. A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 2262–2275 (2010).
24. J. Klemelä, *Smoothing of Multivariate Data: Density Estimation and Visualization*, Wiley-Blackwell, Hoboken, New Jersey (2009).
25. A. B. Said, R. Hadjidj, and S. Foufou, "Cluster validity index based on Jeffrey divergence," *Pattern Anal. Appl.* **20**, 21–31 (2017).
26. J. Goldberger, S. Gordon, and H. Greenspan, "An efficient image similarity measure based on approximations of KL-divergence between two Gaussian mixtures," in *Proc. Ninth IEEE Int. Conf. on Comput. Vision*, Nice, France, France, IEEE, Vol. 1, pp. 487–493 (2003).
27. D. Ribas, "Underwater SLAM for structured environments using an imaging sonar," PhD Thesis, University de Girona, Girona, Spain (2008).
28. D. Ribas, "Towards simultaneous localization & mapping for an AUV using an imaging sonar," *Report on the Research Project*, University de Girona, Girona, Spain (2005).
29. B. Douillard et al., "Scan segments matching for pairwise 3D alignment," in *IEEE Int. Conf. on Rob. and Autom.*, Saint Paul, MN, USA, IEEE, pp. 3033–3040 (2012).
30. S. Song et al., "Two-dimensional forward-looking sonar image registration by maximization of peripheral mutual information," *Int. J. Adv. Rob. Syst.* **14**, 1–17 (2017).

Min Jiang is a doctoral candidate at Shenyang Institute of Automation, Chinese Academy of Science. He received his BS degree from the School of Electronics and Information Engineering, Southwest Petroleum University, China, in 2012. Currently, his research interests include sonar image processing, underwater SLAM, and the development of AUVs.

Sanming Song received his MS degree in computer science and technology from the Guangdong University of Technology, China, in 2008, and his PhD in computer science and technology from Harbin Institute of Technology, China, in 2013. During 2014 to 2015, he held a postdoctoral research position with the Shenyang Institute of Automation, Chinese Academy of Sciences, where he is currently an associate researcher. His interests include robotic intelligence, sonar image processing, and computational neuroscience.

Fengzhen Tang received her BEng degree in software engineering and her MSc degree in computer science and technology from Northeastern University, Shenyang, China, in 2009 and 2011, respectively, and her PhD in computer science from the University of Birmingham, Birmingham, UK, in 2015. She is an associate researcher with the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, China. Her research interests include machine learning, computational neuroscience, robotics, and signal processing.

Yiping Li received her MS degree from Shenyang Institute of Automation, Chinese Academy of Science, China, in 1988. Currently, she is a researcher with the Marine Robotics Department. She participated in the development of China's first 1000-meter AUV, Explorer in 1994, and China's first 6000-meter AUV, CR-01 in 1995. Her current research interests include marine robotic control, concept underwater robot development, and multi underwater robot system research.

Jian Liu received his MS degree from Dalian University of Technology, China. He is a researcher of Shenyang Institute of Automation, Chinese Academy of Science. His research interests include AUV general technology and AUV control. He is the chief designer of the Qianlong I, II AUV and CR02 AUV. He has 17 years of ocean experience with deep-sea AUV research and operations.

Xisheng Feng received his degree from the Department of Industrial Enterprises Electrification and Automation, Harbin Institute of Technology, China, in 1965. He was the deputy general designer (technical director) of China's first 6000-meter AUV, CR-01 in 1995. In 1999, he was elected as an academician of the Chinese Academy of Engineering. His current research interests include marine robotics, ROV and AUV, acoustic image processing, and intelligent control.