Observation Contribution Theory for Pose Estimation Accuracy

Zeyu Wan, Yu Zhang, Bin He, Zhuofan Cui, Weichen Dai, Lipu Zhou, Guoquan Huang

Abstract—The improvement of pose estimation accuracy is currently the fundamental problem in mobile robots. This study aims to improve the use of observations to enhance accuracy. The selection of feature points affects the accuracy of pose estimation, leading to the question of how the contribution of observation influences the system. Accordingly, the contribution of information to the pose estimation process is analyzed. Moreover, the uncertainty model, sensitivity model, and contribution theory are formulated, providing a method for calculating the contribution of every residual term. The proposed selection method has been theoretically proven capable of achieving a global statistical optimum. The proposed method is tested on artificial data simulations and compared with the KITTI benchmark. The experiments revealed superior results in contrast to ALOAM and MLOAM. The proposed algorithm is implemented in LiDAR odometry and LiDAR Inertial odometry both indoors and outdoors using diverse LiDAR sensors with different scan modes, demonstrating its effectiveness in improving pose estimation accuracy. A new configuration of two laser scan sensors is subsequently inferred. The configuration is valid for three-dimensional pose localization in a prior map and yields results at the centimeter level.

Index Terms—Robotics, Pose Estimation Accuracy, LiDAR.

1 INTRODUCTION

The basic ability of mobile robots is navigation which includes pose estimation, motion control, and path planning [1]. Generally, pose estimation involves the use of a map that is initially unavailable before the robot enters an unfamiliar environment. To solve localization and map building problems concurrently, simultaneous localization and mapping (SLAM) methods have been applied. Different types of SLAM sensors, such as inertial measurement units (IMUs), LiDAR, and cameras, are available. Because they are incremental sensors, the integration process is introduced in pose estimation, unavoidably causing the accumulation of errors in a SLAM system. This drift causes map distortion and even estimation failure. Although SLAM has a loop-closing function, it disperses the current error into the history trace instead of eliminating every pose error compared with the ground truth [2]. Currently, to overcome these deficiencies, LiDAR odometry (LO), visual odometry (VO), and multisensor fusion methods are widely utilized [3]; however, improving long-term autonomy accuracy remains a challenge. Compared with existing methods, we presume that the core to boosting the pose estimation accuracy is the improved use of observation information. Based on the reports in [4]–[7] and our experiences [8], as shown in Figs. 1 and 2(a)–2(c) utilizing various measurements in estimation can enhance optimization.

Fig. 1. Velodyne HDL64 laser scan angles are distributed from 2° to −24.8°; 64 scans are defined (i.e. 0-63).

Fortunately, recent studies have confirmed the foregoing and applied entropy based information theory to determine the better use of observation information [6], [7], [10], [11]. These studies concluded that higher entropy levels imply stronger informative points [12], which can improve pose estimation [6], [7]. The studies refer to entropy at the top level using SLAM, model estimation problems by identifying the best subset of a linear matrix. With this metric, satisfactory feature selection is achieved by solving an NP-hard problem. A lazy greedy algorithm is also designed to find the maximum sub-matrix. However, although the algorithm’s search time is considerable, achieving the global optimal is not ensured.

To avoid the selection of a non-maximum sub-matrix, observation uncertainty and sensitivity models are derived and combined with the observation contribution theory (from coarse to fine). In summary, estimation accuracy has two parts: (1) LiDAR measurement and geometry pattern uncertainty models (these indicate the characteristics of points, planes, and lines); (2) residual term sensitivity models (these describe the effects on calculation transformation). Their creation is defined as an observation contribution to the accuracy of estimation. In this implementation, the

• Zeyu Wan, Yu Zhang, Bin He and Zhuofan Cui are with State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China. E-mail: {zeyuwan,zhangyu80,binhe,zhuofancui}@zju.edu.cn
• Weichen Dai is with the College of Computer Science, Hangzhou Dianzi University, Hangzhou, China. E-mail: weichen.dai@hotmail.com
• Lipu Zhou is with Meituan, 7 Rongda Road, Chaoyang District, Beijing, 100012, China. E-mail: zhoulipu@meituan.com
• Guoquan Huang is with the Robot Perception and Navigation Group (RPNG), University of Delaware, Newark, DE 19716, USA. E-mail: ghuang@udel.edu.

(Corresponding author: Yu Zhang. Zeyu Wan and Yu Zhang contributed equally to this work.)
ALOAM uses 25 to 63 scans
\[ \varepsilon_R = 0.0043 \text{ deg/cm} \]
\[ \varepsilon_t = 0.4201\% \text{ m} \]

Consequently, this avoids the selection of a sub-matrix or the optimization of mutual information \[6, 7, 11]\). Another advantage is that the proposed decoupling approach leads to the creation of finer laser points and a map geometry model. This enables a fusion process supporting the above mentioned studies. Nevertheless, these studies do not clarify why the selection of informative points receives higher estimation accuracy. Although this conclusion is evident in information theory \[12\], it is used as a consequence. To the best of our knowledge, our study is the first to provide a global statistical optimum demonstration of a singular value decomposition (SVD)-based registration method \[13\].

After analyzing related studies and comparing them with the present work, the main contributions of our study are as follows; details are provided in succeeding sections.

1. To the best of our knowledge, this is the first study to theoretically prove the global statistical optimal selection for pose estimation.

2. This study devised a LiDAR point measurement method, a map geometry pattern uncertainty model and point-plane, point-line residual term sensitivity models with a calculation method. The decoupling and fusion mechanisms are also defined.

3. Experiments on the KITTI benchmark were conducted, The ALOAM translation error decreased from 1.7318\% to 1.5781\% in virtually half the number of used planes and lines. Different types of LiDAR scan modes were evaluated in indoor and outdoor environments with LO and LIO (LiDAR inertial odometry), increasing the average accuracy by approximately 20\%. An orthogonal laser configuration valid for locating a three-dimensional (3D) pose in a prior map was inferred; the root mean square error (RMSE) is in the centimeter level.

The remained of this article is structured as follows. Relevant basic works are introduced in section 2. Notations and preliminaries are presented in section 3. Demonstrations of the global optimum are presented in section 4. The details of contribution theory in LiDAR pose estimation are presented in section 5. For the experiments, the simulation of contribution selection mechanism using artificial data is presented in section 6.1. The proposed algorithm on the KITTI benchmark \[9\] is discussed in section 6.2, and tests on different types of LiDAR sensors in real indoor and outdoor environments are elucidated in section 6.3. The inference regarding multi-laser assembly based on our theory and its validation when applied to a real building are also elaborated. Finally, conclusions are summarized in section 7.

\section{2 Related Work}

In this work, related studies on improving accuracy are classified into four: traditional methods, uncertainty models, sensitivity models and contribution theory.

\subsection{2.1 Traditional Method}

To reduce drift, traditional LO, VO, and SLAM use theories to ameliorate performances. These can be summarized into three categories. The first includes filter-based methods: GMapping \[14\], LOAM \[15\], IMLS \[5\], MonoSLAM \[16\], PTAM \[17\]. With these methods, the current state is assumed
to only pertain to the former state and current observations. The Markov process [18] is a well-known technique that simplifies computations and yields small errors when the state approaches the assumed state point [2]. The second category is the nonlinear optimization-based methods, such as SuMa [19], SuMa++ [20], LiOmapping [21], LiTAMIN [22], LiTAMIN2 [23], MULS [24], BALM [25], LSD SLAM [26], SVO [27], ORB SLAM [28], DSO [29]. These techniques store all old states and use current observations to be updated under the maximum posterior estimation principle. All valid information is incorporated into the optimization structure with significant computations [2]. The third category pertains to methods based on incremental smoothing and mapping. These techniques include iSAM [30], iSAM2 [31], LeGOLOAM [32], AprilSAM [33], LIOSAM [34]. These methods model states as a factor graph; moreover, current observations only update relevant factors. As a result, accuracy and speed are well balanced [2]. Although filter format and optimization differ, both agree mathematically and share the same target. These theories focus on establishing the geometry or texture residuals of information obtained at different dates. Because of the Gaussian or linear assumption, a filter can directly drive the optimal result, obtaining the posterior by adding the likelihood to the prior. Nonlinear optimization solves the entire cost function at a single instance. For example, iSAM compromises the computing complexity and solves the problem in multiple steps. The minimization and ensured convergence of this function have been wildly and thoroughly studied [35, 36]. In general, traditional methods consider all valid observation information based on distinct weights.

2.2 Uncertainty Model

The uncertainty model consists of two parts. The first part independently models the uncertainty of every laser point measurement in 3D. The second part models the uncertainty of the geometry pattern in a map. For laser points, [37] proposed a rigorous first-order error analysis. It measures the horizontal and vertical errors of a laser pulse and finds the nonlinear error growth as recently reported in [38]. In comparing the various LiDAR sensors available in the market, measurement errors are found to be relevant to the target range. In [39], a laser point is modeled as projected footprint, and used to represent an uncertainty matrix. A 3D Gaussian distribution is proposed [40] to model LiDAR uncertainty points and clarify their propagation. For the geometry pattern, we believe that point cloud data (PCD) are direct and easy to use for localization. Accordingly, this investigation is focused on the map geometry pattern implemented using PCD. The LOAM algorithm [15] generates five points to simulate a plane and a line by decoupling eigenvalues. It fundamentally calculates them as a 3D Gaussian distribution, but discards irrelevant directions. The Gaussian mixture model [41] (GMM) is a continuous distribution function method; however it adopts multiple Gaussians and regroups them with different weights. A multi-layers tree structure [42] can fuse flat areas into one Gaussian or decompose a complicated area into several Gaussians. In addition to these implicit function representations, implicit methods are also relevant to PCD application.

A Moving Least Square (MLS) surface is defined in [43], it is a $C^\infty$ smooth surface generated from raw PCD. An implicit version is defined in [44]; it represents the distance of a location on a surface composed of neighbor points. Based on the studies mentioned above, real LiDAR emitting and receiving structures are considered to build a laser point uncertainty model. Then, the measured points are added to the map and fused. The concepts of GMM are also adopted and considered as goals in the SLAM system. Our study aims to improve the real-time estimation accuracy of mobile robots rather than real world precision. Compared with 3D reconstruction, sacrificing some of the complicated area details and focusing on the main direction constraints is advantageous for SLAM. Because nearby LiDAR points are dense, remote points are sparse. Leaves, trees and other irregular objects are not suitable for pose estimation; this means that their fused uncertainties are greater. Therefore, based on the LOAM algorithm, the Gaussian method is preferred, and the uncertainties of map points are considered. Consequently, the complete use of these points in a plane or line is applied.

2.3 Sensitivity Model

The report in [4] provided the current study with considerable motivation to employ the residual term sensitivity model. The paper proposes a technique for determining whether a pair of meshes are unstable in the iterative closest point (ICP) algorithm by estimating a covariance matrix from the sparse uniform sampling of input. Then, it develops a sample strategy that attempts to minimize this instability by drawing a new set of sample points primarily from the stable areas of input meshes. However, this study is confronted with the registration problem; it does not consider measurement uncertainties and only analyses mesh plane errors. Nevertheless, the technique remains fundamental to our theory. In the formulated theory, LODegeneracy [45] aims to avoid a degenerate environment, which could be regarded as a condition where sensitivity is zero. It determines and separates degenerate directions in state space and partially solves the problem in well-conditioned directions. It linearizes the cost function and uses the dot product of coefficient matrix and its transpose to form a covariance matrix containing the geometric structures of problem constraints. The IMLS [5] technique is a complete LO and uses the covariance method of [4] to select points. It has an experiment and a conclusion with numerous points that can alter constraints to determine the final pose. However, it does not solve the problem theoretically. It only considers a point-plane residual type such that variations in the point-line type are observed. The LeGOLOAM algorithm [32] uses normal vector clustering to detect true line points and obtain better matching. Optimization is achieved through two stages using the ground vehicles hypothesis. The LeGOLOAM algorithm can be regarded as improving accuracy from the pattern recognition perspective, which is limited by the strong ground vehicle hypothesis. Two-stage optimization enables all observations to calculate rotation and translation separately. Hence, this method can improve accuracy to some extent. LION [46] can self-assess its performance using an observability metric that evaluates
whether pose estimation is geometrically ill-constrained. It is similar to LODegeneracy \cite{45} and is applied to a real tunnel scene LO. Details regarding the derivative of residuals are considered by SGLO \cite{47}, however, it is not discussed in-depth. It does not consider constraint information in every transformation dimension, which is the core content. Point-point, point-line, point-plane types are constructed by MULLS \cite{24}, distinctly clarifying linearization details. It uses all observations in estimations with diverse weights, implying that estimation in different directions can be balanced; however, MULLS simply ensures observability. This study aims to fully clarify this field and theoretically improve estimation accuracy. In \cite{48}, inline set cardinality maximization was used to select good feature for a 2D-3D pose estimation. Bearing vectors play the role to select sensitive observations avoid degeneration.

2.4 Contribution Theory

Since 2020, some researches have applied entropy from information theory to the SLAM system, endeavoring to improve its robustness and accuracy. The author of \cite{6} proposed sub-matrix selection by choosing a scoring metric in VO. They modeled estimation as a linear matrix to find the best subset. Its simulation yields the best metric Max-LogDet because of its lower computation cost and considerable accuracy. With this metric, satisfactory features selection becomes an NP-hard problem. They designed a lazy greedy algorithm to find the maximum sub-matrix. Although they did not explicitly express that they applied the information theory as that employed in their concurrent work \cite{7}, the same author focused on selecting a satisfactory pose for graph optimization using the same concept and method. In section 3 and 4, entropy was used to model the estimation problem, and the most numerous mutual information points were selected from RGBD Odometry. The authors of \cite{11} imitated and applied entropy to a multi-LiDAR field to increase robustness and accuracy. In the aforementioned studies, two LiDAR sensors were set in diverse angles to cover wide areas. Moreover, a greedy method was designed to solve the NP-hard problem. After the points were selected, the two LiDARs ran in real time with satisfactory accuracy.

3 Notations and Preliminaries

Before introducing our contribution theory, interpreting the ICP registration can aid in understanding the principles involved. Assume that the current set of measurement points is \( P = \{ p_i \}, i = 0, \ldots, N \) and its corresponding set of referent model points is \( Q = \{ q_i \}, i = 0, \ldots, N \). Define a standard \( L_2 \) norm point-point ICP problem as

\[
T^* = \arg \min_{T \in SE(3)} \sum_{i=0}^{N} ||Rt_e(T)p_i - t - q_i||^2 \tag{1}
\]

where \( T = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \) is a transformation matrix. \( e_i \) is the residual vector. By linearizing the transformation parameters, \( T \in SE(3) \), from the Lie group manifold to the corresponding location in tangent vector space (i.e. \( \xi^\wedge \in se(3), \xi \in \mathbb{R}^6 \) Lie algebra), the problem becomes a linear least square. Accordingly, it can be solved by the Gauss-Newton \cite{49} or Levenberg-Marquardt \cite{50} technique.

To full demonstrate the foregoing, a standard ICP solution SVD is employed. Zero-mean normalization is applied to the sets of points, decoupling the problem into rotation and translation. In other words, it solves \( R \) in \( SO(3) \) space and then back to \( SE(3) \) space to calculate \( t \). For convenience, \( p_i \) and \( q_i \) are used to represent points in \( SO(3) \). This is known as the Wahba problem \cite{51} or rotation search \cite{52}, given as follows:

\[
R^* = \arg \min_{R \in SO(3)} \sum_{i=0}^{N} ||Rp_i - q_i||^2
\]

\[
= \arg \min_{R \in SO(3)} \sum_{i=0}^{N} (q_i^Tq_i + p_i^TR^TRp_i - 2q_i^TRp_i)
\]

\[
= \arg \max_{R \in SO(3)} \sum_{i=0}^{N} (q_i^T R p_i)
\]

\[
= \arg \max_{R \in SO(3)} tr(R \sum_{i=0}^{N} p_iq_i^T)
\]

Then by defining

\[
H = \sum_{i=0}^{N} p_iq_i^T = U\Sigma V^T
\]

the optimal rotation is found as

\[
R^* = VU^T
\]

The rotation matrix, \( R \), and corresponding axis angle (rotation vector) representation, \( \phi_R = \theta_R \omega_R \), shown in Fig. 3 are connected by the exponential map from Lie group to Lie algebra, \( R = \exp(\phi_R^\wedge) = \exp(\theta_R \omega_R^\wedge) \). In section 3 and 4 \( \theta_R \) is presented as the angle, and \( \omega_R \) is the axis of \( R \); \( \phi^\wedge \) denotes the symmetric skew matrix of vector \( \phi \).

\[
\text{LiDAR}
\]

\[
\omega'_L
\]

\[
\theta_L
\]

\[
L'
\]

\[
\text{LiDAR}
\]

\[
\omega'_L
\]

\[
\theta_L
\]

\[
L'
\]

Fig. 3. Axis angle (rotation vector) representation is \( \phi_L' = \theta_L' \omega_L' \) and \( L' = \exp(\phi_L'^\wedge) = \exp(\theta_L' \omega_L'^\wedge) \).

To measure the difference between the two rotation matrix, the Riemannian metric distance under the Frobenius
4 Estimation Contribution

This section presents proof that the proposed contribution theory-based selection method achieves the global statistical optimal. Although uncertainty models are involved, the method is statistical because these models are still probability distributions. In deriving the least number of uncertainty points in one test, a slight possibility that the remaining points are close to the accurate location exists. In one test, a point with a higher uncertainty may be more accurate than other points. Therefore multi-sampling and calculation expectations can reveal superiority; statistical analysis results are discussed in a later section. In ICP-SVD, the zero-mean normalization causes the registration to attain optimal rotation in SO(3) and then solves its translation. For brevity, the demonstration of SE(3) can be naturally inferred from SO(3).

As shown in Fig. 4, there exists an accurate point $p^*$ in the current LiDAR coordinate. It is measured as $p'$ because of noises. In SO(3), a dynamic rotation matrix $L'$ describes this error, $p' = L'p^*$ and they lie on a 3D manifold surface. In real LiDAR measurements, errors are within several centimeters and point distances are dozens of meters. Assuming the measured point is uniform (gaussian is both ok) distributed around the accurate location. Then applying these relationships to $so(3)$, $p'$ is uniformed within a $h \in (0, \epsilon)$ radius plane circle in Fig. 5. $h$ is variable and $\epsilon$ is a given distance to the maximal far location.

**Lemma 1 (Point Uncertainty Model).** The point (i.e. $p^*$) uncertainty model is given by

$$\Phi_{p^*} = \frac{\epsilon}{2}$$

Fig. 5. Effect of noise, $p' = L'p^*$, on measurement of accurate points. Dynamic rotation matrix, $L'$, describes this error in SO(3). After application to $so(3)$, it lies in radius, $h \in (0, \epsilon)$, of plane circle, where $h$ (i.e. given distance to far maximal location $\epsilon$) is variable.

Because $||h||^2$ indicates the error amplitude, its expected integration is

$$\Phi_{p^*} = E(h^2) = \frac{1}{\epsilon} \int_{0}^{\epsilon} \frac{1}{2\pi h} \int_{0}^{2\pi} h^2 dh dh = \frac{\epsilon}{2}$$

where $\alpha$ indicates the round integration of $p'$.

**Lemma 2 (Point Sensitivity Model).** Under the point-point residual, the point (i.e. $p^*$) sensitivity model is

$$J_{p^*} = ||p^*||^2$$

Points that are distant from the center of a LiDAR sensor undergo more changes when the same rotation is applied to them. Thus, in point-point registration, sensitivity is given by the norm.

Consider the given rotation disturbance, $L'$ (simulated noise), in the current LiDAR points, $p' = L'p^*$, and $K'$ in the reference map points, $q' = K'q^*$. The target is to derive an optimal rotation result, $R'$, satisfying $q' = R'p'$ shown in Fig. 6. The original accurate connection is $q' = K'q^*$. The Riemannian distance, $B' = R'^{-T}R'$, indicates the accuracy of this estimation. In other words, if the residual constructed by a pair of these points (whose contribution is large), the Riemannian distance must be small.

Fig. 6. Wahba problem involves two sets of points $P^* = \{p^*_i\}, i = 1, \ldots, N$ and $Q^* = \{q^*_i\}, i = 1, \ldots, N$, without disturbances and rotation, $R^*$; these two sets are calculated. Disturbances and amplitudes can be bound within two rotations, $L'$ and $K'$; and optimal rotation result is $R'$. This work proves that with our selection, $R'$ is closer to $R^*$.
Theorem 1 (Optimal Rotation Under Small Disturbance).

\[ R' = K'R^*L'^T \]  

(10)

Proof 1 (Optimal Rotation Under Small Disturbance).

Substitute disturbances into Eqs. (3) and (4):

\[ H' = \sum_{i=0}^{N} p_i'q_i'^T = \sum_{i=0}^{N} L'p_i'q_i'^TK'^T \]

(11)

\[ = L'H^*K'^T = (L'U) \Sigma (K'V)^T \]

Thus, the optimal rotation is

\[ R' = K'V(L'U)^T = K'VU^T L'^T = K'R^*L'^T \]

(12)

Q.E.D.

Theorem 2 (Riemannian Distance Under Small Disturbance).

\[ \text{Riem}(B') = \| \log (R'^*R') \|_{F}^2 \]

\[ = \frac{1}{2}[2(R^*\phi_{K'})^T(R^*\phi_{K'}) + 2\phi_L^T\phi_L'] \]

(13)

\[ + \frac{1}{2}[\phi_{K'}^T\phi_L + (R^*\phi_{K'})^T\phi_L'] \]

\[ + 4(R^*\phi_{K'})^T\phi_L'] \]

Proof 2 (Riemannian Distance Under small disturbance).

Applying the Lie group adjoint property and the exponential map to Eq. (10) yields

\[ R'^*R' = (R^T K^* R'^* L'^T)^T \]

(14)

\[ = (R^T \exp(\phi_K^*) R'^*) L'^T \]

\[ = \exp(- (R^*\phi_K)^\wedge) \exp(-\phi_L^\wedge) \]

Because the two rotations are small, the left Jacobin or right Jacobin model is not suitable. Directly apply the Baker-Campbell-Hausdorff (BCH) formula and keep the first-order terms as follows:

\[ \log[\exp(A) \exp(B)] = A + B + \frac{1}{2} \{A, B\} + \ldots \]

(15)

where \{A, B\} is the Lie bracket; it satisfies

\[ \{A, B\} = AB - BA = \phi^A \phi^B - \phi^B \phi^A \]

\[ = (\phi^A \phi^B)^\wedge \]

(16)

Substituting Eqs. (15) and (16) into Eq. (14) yields

\[ \text{Riem}(B') = \| \log (R'^*R') \|_{F}^2 \]

\[ = \| \log[\exp((-R^*\phi_K)^\wedge) \exp(-\phi_L^\wedge)] \|_{F}^2 \]

\[ = \| - (R^*\phi_K)^\wedge - \phi_L^\wedge + \frac{1}{2}[(R^*\phi_K)^\wedge \phi_L^\wedge] \|_{F}^2 \]

(17)

Using the Forobenius norm and trace properties yields

\[ \|A + B\|_{F}^2 = \frac{1}{2} \text{tr}[(A + B)^T(A + B)] \]

\[ = \frac{1}{2} [\text{tr}(A^T A) + \text{tr}(B^T B) + 2\text{tr}(A^T B)] \]

\[ = \frac{1}{2} [\text{tr}(a^\wedge^T b^\wedge)] \]

\[ = 2a^T b \]

(18)

(19)

Substituting Eqs. (18) and (19) into Eq. (17) results in

\[ \text{Riem}(B') = \frac{1}{2}[2(R^*\phi_{K'})^T(R^*\phi_{K'}) + 2\phi_L^T\phi_L'] \]

(20)

\[ + \frac{1}{2}[\phi_{K'}^T\phi_L + (R^*\phi_{K'})^T\phi_L' + 2(R^*\phi_{K'})^T\phi_L'] \]

\[ - 2(R^*\phi_{K'})^T(R^*\phi_{K'})^T\phi_L'] \]

\[ = 0 \]

Eq. (20) is composed of 6 terms. Based on Eqs. (22) and (23), the last two terms in Eq. (20) are consistently zero. A new rotation vector is defined in Fig. 7

\[ \phi_{G'} = R^* \phi_{K'} \]

(21)

Substitute Eqs. (18) and (19) into Eq. (17) results in

\[ \text{Riem}(B') = \frac{1}{2}[2(R^*\phi_{K'})^T(R^*\phi_{K'}) + 2\phi_L^T\phi_L'] \]

\[ + \frac{1}{2}[\phi_{K'}^T\phi_L + (R^*\phi_{K'})^T\phi_L' + 2(R^*\phi_{K'})^T\phi_L'] \]

(22)

\[ - 2(R^*\phi_{K'})^T(R^*\phi_{K'})^T\phi_L'] \]

\[ = 0 \]

Thus, Theorem 2 is confirmed;

Q.E.D.

Remark 1 (Disturbance Terms). Based on Theorem 2 and Fig. 3, the first and second terms are squares of disturbance angle values:

\[ (R^*\phi_{K'})^T(R^*\phi_{K'}) = \theta_{K'}^2 \]

\[ \phi_L^T\phi_L' = \theta_{L'}^2 \]

(25)

The third and fourth terms are dynamic depending on the included angle of these two disturbance rotation vectors. The expression \( R^*\phi_{K'} \) can be regarded as a rotating \( q^* \) disturbance in \( p^* \). The cross product can be used to calculate the angle difference and reflect the effect on the
Theorem 3 (Point Contribution Model). The contribution model of point \( p^* \) and corresponding \( q^* \) is

\[
\Psi_{pq} = \Phi_{pq}/\Phi_{pq} \tag{26}
\]

We assume that \( \Psi_{pq} \) is the residual contribution, \( \Phi_{pq} \) is its sensitivity, \( \Phi_{pq} \) is its uncertainty:

\[
\Psi_{pq} = J_{pq}/[\Phi_{p}^2 + \Phi_{q}^2] \tag{27}
\]

Given two residual pairs (i.e. \( p', q' \) and \( p'', q'' \)), the larger contribution residual is better for estimation; it has a short Riemannian distance to the optimal result, as follows:

\[
\Psi_{p'q'} < \Psi_{p''q''} \Leftrightarrow Riem(B'') > Riem(B') \tag{28}
\]

Proof 3. The Riemannian distance expectation by double integration is

\[
E(Riem(B')) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} \| \log(R^T R') \|^2_2 d\alpha_p d\alpha_q \tag{29}
\]

where \( \alpha_p \) and \( \alpha_q \) are the possible angles of \( p' \) and \( q' \), respectively, as shown in Fig. 8.

Considering the third term, a new vector is formulated as cross product, as shown in Fig. 9:

\[
\phi_{New} = \phi_{G'} \times \phi_{L'} \tag{31}
\]

The integration of the third term becomes \( \phi_{New}^T \phi_{New} \). Moreover, it contains one angle, \( \alpha_p' \), when LiDAR is implemented. The inner integration and exchange of integration variables from \( \alpha_q' \) to \( \beta \) shown in Fig. 9 indicate that \( \phi_{G'} \times \phi_{L'} \) as a fixed value with inner integration:

\[
= \int_0^{2\pi} \int_0^{2\pi} \phi_{New}^T \phi_{New}' d\alpha_p d\alpha_q' \tag{32}
\]

The substitution Eqs. (25), (30), and (32) into Eq. (29) yields

\[
E(Riem(B')) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} (\theta^2_{L'} + \theta^2_{G'}) d\alpha_p d\alpha_q' + \frac{\pi^2}{2} \tag{33}
\]

Eq. (33) demonstrates that the expected Riemannian distance between \( R' \) and \( R^* \) is only influenced by disturbances. Based on Lemmas 1 and 2, the disturbance in \( so(3) \) consists of uncertainty and sensitivity, as shown in Fig. 10.
θ_{K'} = \arctan\left( \frac{\epsilon q'}{||q'||^2} \right) \quad (35)

This can be demonstrated using proof by contradiction if a residual pair (i.e. $p'', q''$) exists and the contribution of this pair exceeds $p', q'$, and the Riemannian distance, $\text{Riem}(B'')$, is closer to zero than $\text{Riem}(B')$:

$$\Psi_{p''q''} < \Psi_{p'q'}$$

$$\iff \frac{||p'||^2}{\epsilon_p^2 + \epsilon_q^2} < \frac{||p'^{*}||^2}{\epsilon_p^2 + \epsilon_q^2}$$

$$\iff \theta_{K'p''}^2 + \theta_{K'q''}^2 > \theta_{K'p'}^2 + \theta_{K'q'}^2$$

$$\iff \text{Riem}(B'') > \text{Riem}(B')$$

(36)

A conflict occurs in the foregoing, demonstrating Theory 3.

Q.E.D.

Fig. 11. Outline of demonstration process.

Traditional SVD method

Theorem 1

Optimal Rotation Under Small Disturbance

Theorem 2

Riemannian Distance Under Small Distance

Remark 1

Disturbance Term

Lemma 1

Point Uncertainty Model

Lemma 2

Point Sensitivity Model

Theorem 3

Point Contribution Model

Fig. 12. Observation contribution selection process. Upper part details observation contribution selection, and lower part indicates location of selection added to LO. Red parts indicate data of laser points and poses; blue parts are operations.

5 LIDAR POSE ESTIMATION

Section 4 describes some of the major contents of this work. The previous section discusses that residual contributions can be decoupled into the uncertainty and sensitivity. However, it does not explain how uncertainty and sensitivity can be calculated using a real LiDAR sensor. This section elaborates on the approach implemented in practice.

The complete procedure of the proposed method is outlined in Fig. 12. The upper part of the outline details the selection of observation contributions. The inputs include map referent models, LiDAR current points, and an initial pose available from a uniform motion model or IMU. Next, the algorithm uses an octree to find neighbors that establish the closest matches. Because current points are classified into surf and corner (plane and line), the algorithm computes uncertainties and sensitivities to form contributions. Finally, it sorts residual terms by scores and sends them to the nonlinear solver to derive the optimal pose.

The remainder of this paper first presents the LiDAR measurement error distribution, uncertainties and fusion methods depending on the type of point-plane and point-line residuals. This part clarifies the primary reasons for the different estimation accuracies of various points. The second part presents the use of Taylor expansion and eigenvalue projection tool to decouple residuals into every dimension.

5.1 Uncertainty Model

As shown in Fig. 13, the uncertainty region of the camera’s triangulation depth considerably increases with distance. Due to laser time-fly ranging and careful rotary mechanism calibration, the LiDAR’s uncertainty region does not grow as larger as that of cameras. Therefore, a LiDAR has more distinct effects on our theory; accordingly, it is first illustrated. Here, (1) single-laser scan beam uncertainty, (2) geometry map patterns uncertainty, and (3) uncertainty fusion.

Fig. 13. Different measurement errors between LiDAR and camera. Camera triangulation depth considerably relies on baseline length. For distant points, the recovery of errors considerably increases; however, for time-fly ranging LiDAR, these effects are not distinct.
5.1.1 Laser Scan Beam

The standard uncertainty associated with measured variable, $\theta$ and $\Delta \theta$, is the accuracy provided by the manufacturer. If the error associated with $\theta$ is a uniform distribution, then $\sigma_\theta = \frac{\Delta \theta}{\sqrt{3}}$. If it is a Gaussian distribution, then $\sigma_\theta = \Delta \theta$, and if $\theta$ is not measured, then $\sigma_\theta = 0$.

Based on a multi-beam laser scanner system [37], [39], [40], point $P_s$ in the scanner coordinate frame, $F_s$, is given by the following.

$$P_s = \begin{bmatrix} x^s \\ y^s \\ z^s \end{bmatrix} = R_l^T(\alpha, \omega) \begin{bmatrix} x^l \\ y^l \\ z^l \end{bmatrix} F_l$$

$$R_l(\alpha, \omega) = \begin{bmatrix} \cos \omega & 0 & \sin \omega \\ 0 & 1 & 0 \\ -\sin \omega & 0 & \cos \omega \end{bmatrix}$$

In the foregoing, $F_l$ is the laser coordinate frame defined by the $z_l$ pointing laser beam direction; $R_l(\alpha, \omega)$ is the rotation matrix from $F_l$ to $F_s$; $\alpha$ is the azimuth angle; $\omega$ is the elevation angle of laser beams channel; $z^l$ refers to depth, and $x^l$ and $y^l$ must be zero.

Following self-rotation, a laser scan line is formed, as shown in Fig. [14(c)]. Every elevation angle, $\omega$, is calibrated and rectified; thus, $\sigma_\omega = 0$. For the azimuth, $\alpha$, the manual states rotation angular resolution is 0.01 °. All traditional studies, as reported in [37], [39], [40], assume that

$$0 < \sigma_\alpha < \frac{\pi \Delta \alpha}{180 \sqrt{3}}$$

where $\Delta \alpha = 0.005$ ° is half the resolution.

The LiDAR parameters and coordinates are shown in Fig. [14(d)]. Because self-rotation is nonlinear, $R_l(\alpha, \omega)$ must be linearized during uncertainty propagation. Then, the uncertainty of point $P_s$ is simulated as $\Sigma_s$, which is a $3 \times 3$ matrix:

$$\Sigma_s = J R_l^T \Sigma_s J R_l$$

Although [38] compares 10 different LiDAR sensors available in the market and states that their measurement uncertainties are not independent of depth, their uncertainties are found to be diverse. Hence, they require calibration at different distance because of varied LiDAR hardware instructions (sending or receiving using different optical axes). Independence is assumed in the proposed method and those in [37], [39], [40].

5.1.2 Map Patterns

As illustrated in Fig. [14(a)], a laser diode stack emits three light beams on the surface. These are received by the LiDAR observing window shown in Fig. [14(b)]. The LiDAR records the emission time and the most intense three light beams on the surface. These are received by the LiDAR observing window shown in Fig. [14(b)]. The LiDAR records the emission time and the most intense beam footprint. The true location may lie anywhere within the beam footprint. Based on the manual, the horizontal and vertical divergence angles of the rectangular window are $\delta_h = 3 \times 10^{-3}$ rad, $\delta_v = 1.5 \times 10^{-3}$ rad. Therefore, assuming that point $P_s$ is uniform in the region, it is modeled as a Gaussian.

$$\sigma_{x^l} = \frac{\Delta x^l}{\sqrt{3}} = \frac{z^l \tan(\frac{\pi}{3})}{\sqrt{3}}, \sigma_{y^l} = \frac{\Delta y^l}{\sqrt{3}} = \frac{z^l \tan(\frac{\pi}{3})}{\sqrt{3}}$$

Because of inhomogeneous noise, sparse density and missing data in LiDAR sampling [54], the estimation in LO typically employs plane and line patterns by minimizing the alignment distance [15], [21], [24], [32], [34]. The LO baselines, such as LOAM [15], use five neighboring points to model the plane or line shown in Fig. [15(a)] and [15(b)].

Eigenvalues $\lambda_0 < \lambda_1 < \lambda_2$ and eigenvectors $\nu_0, \nu_1$ and $\nu_2$ corresponding to $x^l, y^l$ and $z$ dimensions are calculated by LOAM, respectively. This involves modeling a surface as a 3D Gaussian ellipsoid. The Gaussian is treated as a prior, and their uncertainty influence is considered to obtain.
a posterior. However, in Bayes inference, the likelihood function and marginalized distribution are nonlinear and unknown. Because our purpose is to model uncertainties in registration, the main error direction uncertainties are modeled using the sigma point transform \[55\].

To infer the posterior Gaussian, one sigma point, \( P_{sj} \) is resampled around the ellipsoid, \( P_s \sim \mathcal{N}(\mu_s, \Sigma_s) \) as follows:
\[
P_{sj} = \mu_s \pm \lambda_j \nu_j, \quad j = x, y, z
\]
where point \( P_s \) along its eigenvector location has an eigenvalue of 1. Therefore, \( 2 \times 3 \times 5 = 30 \) points for recalculation and approximation are regressed as the posterior.

5.1.3 Currence and Reference Fusion

Referent posterior Gaussians are defined as \( G^{ref} \sim \mathcal{N}(\mu^{ref}, \Sigma^{ref}) \); current measurement points are defined as \( G^{curr} \sim \mathcal{N}(\mu^{curr}, \Sigma^{curr}) \). Data association uses the closest principle to match point-plane and point-line terms. As for referent and current uncertainty fusion, the use of a new principle to match point-plane and point-line terms. As for referent and current uncertainty fusion, the use of a new distribution to fit them by minimizing the Kullback-Leibler referent and current uncertainty fusion, the use of a new principle to match point-plane and point-line terms. As for referent and current uncertainty fusion, the use of a new principle to match point-plane and point-line terms.

\[
\Phi_{e_i} = \begin{cases} 
(\lambda^{ref}_i + \lambda^{curr}_i)/2 & \text{(plane)} \\
(\lambda^{ref}_i + \lambda^{ref}_i + \lambda^{curr}_i + \lambda^{curr}_i)/4 & \text{(line)} 
\end{cases}
\]

where \( \Phi_{e_i} \) is a scalar that evaluates the uncertainty.

5.2 Sensitivity Model

\[\begin{align*}
\text{(a) LO uses red measurements to estimate pitch} \\
\text{(b) LO uses green measurements to estimate pitch}
\end{align*}\]

Based on \( \Delta e = J \Delta T \) and Theorem 3, the estimation error is found to be proportional to the residual term’s sensitivity and uncertainty. As shown in Fig. 17 if a point deviates along the normal of its modeling plane, a farther point has a greater influence on rotation estimation. In Fig. 17 the orange line is the current point in the LiDAR coordinate, and the dark hole is central to the map referent surf points whose uncertainty region is a blue ellipsoid.

In Fig. 17(a), the current point is close to the LiDAR and must be aligned with the referent surf point (i.e. red line). In Fig. 17(b), the current point must be aligned with the green line. The sensitivity of residual terms shown in (a) to pitch angle estimation is lower than of terms shown in (b). This means that with the same measurement accuracy, the use of distant points in (b) yields higher accuracy. The measurement of (a) forms a residual whose addition to an optimization problem corrupts the convex qualities of the problem. This is the reason for defining the residual term’s sensitivity. The construction of point-plane and point-line sensitivities are discussed next, and some detailed realistic problems are solved.

To satisfy the assumption of infinitesimal rotation and translation, the linearization error approach zero:
\[
R \approx \begin{bmatrix}
1 & -r_z & r_y \\
-r_z & 1 & -r_x \\
r_y & r_x & 1
\end{bmatrix} = I_3 + \begin{bmatrix}
r_x \\
r_y \\
r_z
\end{bmatrix}^\wedge
\]

The sensitivity model is introduced from two residual types.

5.2.1 Point to Plane

\[\begin{align*}
\text{Fig. 18. Plane residual.}
\end{align*}\]

A current point, \( p_i \), is defined as a referent point, \( q_i \), on the plane. Its normal is \( n_i \), as shown in Fig. 18 \( \lambda^{ref}_i \) is the eigenvalue along \( n_i \). The three current eigenvector directions are calculated, and the angle closest to \( n_i \) corresponding to the eigenvector is used as \( \lambda^{curr}_i \). The error of the \( i \)-index residual term point-plane distance is
\[
e_i = (R p_i + t - q_i)^T n_i
\]

and \( e_i \) is a scalar. The Jacobin and linearized rotation are calculated:
\[
e_i \approx [(I + r \wedge) p_i + t - q_i] n_i
\]
\[
J_{e_i} = \frac{\partial e_i}{\partial (r, t)} = [(p_i \times n_i)^T n_i^T]
\]

5.2.2 Point to Line

\[\begin{align*}
\text{Fig. 19. Line residual.}
\end{align*}\]

A current point, \( p_i \), is defined as a referent point, \( q_i \), on the line with the pointing direction unit vector, \( n_i \),
as shown in Fig. [19]. Because the line model collapses in two dimensions, the uncertainty must consider two orthogonal directions with their eigenvalues. In this case, 
\[ \lambda^\text{ref}_i = (\lambda^\text{ref}_0 + \lambda^\text{ref}_2)/2, \]
and \( \lambda^\text{curr}_i \) is used as the maximum eigenvalue. The error of the i-index residual term point-line distance is
\[ d_i = (q_i - Rp_i - t) \times (q_i + n_i - Rp_i - t) \] (49)
where \( d_i \) is a 3 × 1 vector, and \( e_i \) is a scalar, as follows:
\[ e_i = d_i^T d_i \] (50)
It differs from the point-plane distance. First, \( d_i \) must be derived:
\[ d_i \approx (q_i - p_i - t - r \times p_i) \times (q_i - p_i - t - r \times p_i + n_i) \]
\[ = (q_i - p_i - t - r \times p_i) \times n_i \] (51)
\[ J_{d_i} = \left[ \frac{\partial d_i^T}{\partial (r, t)} \right]^T = [(n_i^T p_i) I_3 - p_i n_i^T n^\wedge] \] (52)
where \( J_{d_i} \) is a 6 × 3 matrix. Moreover,
\[ \Delta e_i = \Delta \begin{bmatrix} r \ t \end{bmatrix} J_{d_i}^T J_{d_i} \Delta \begin{bmatrix} r \ t \end{bmatrix} \] (53)
Thus, the Hessian matrix, \( H_{d_i} = J_{d_i}^T J_{d_i} \) is defined. This means that point-line \( e_i \) is a quadratic form of optimization parameters. Different from the linear form in point-plane distance, directly decoupling the contribution into six dimensions is impossible because the partial derivatives of the quadratic function approximating \( \Delta r = 0, \Delta t = 0 \) are consistently zero. Fortunately, this study only focuses on the growing gradient in a small region. The hessian matrix is projected onto six axes. It is regrouped into a linear form,
\[ H_{6 \times 6} = V \text{diag} [\lambda_0 \ldots \lambda_5] V^T, V = [\nu_0 \ldots \nu_5] \] (54)
then every eigenvalue with vectors is projected on the j-index axis:
\[ \text{Axis}_j = \begin{bmatrix} 0 & 1_j \end{bmatrix}^T, j = 0, \ldots, 5 \] (55)
\[ \text{Proj}_j = \sum_{k=0}^{m} \lambda_k \nu_k^T \text{Axis}_j, k = 0, \ldots, 5 \] (56)
the j index is formed as a part of a 1 × 6 vector, \( \text{Proj}_j \):
\[ \text{Proj}_j = \begin{bmatrix} \text{Proj}_0 \ldots \text{Proj}_5 \end{bmatrix} \] (57)

### 5.3 Contribution Model

By combining the uncertainty and sensitivity models from Theorem 5 and Eq. (27), the observation contribution to estimation is derived as:
\[ \Psi_{e_i} = J_{e_i}/\Phi_{e_i} \] (58)
where score \( \Psi_{e_i} \) is a 6 × 1 vector corresponding to three rotations and three translations. By substituting Eqs. (44) and (48) into Eq. (58), the contribution of the point-plane residual term is derived. The point-line residual term’s contribution is obtained by substituting Eqs. (44) and (57) into Eq. (58).

### 6 Experiments

The complete procedure of the contribution analysis is presented in the previous sections. The algorithm sorts all contribution scores of terms from high to low and gathers them in fixed amounts, exhibiting its real-time performance. This contribution theory-based residual term selection method can be easily applied. It achieves significant improvements in accuracy but with fewer residual terms and lower computational cost in nonlinear optimization.

Our codes are implemented in C++. The program is executed on six-core CPU AMD 2600X, 48-GB RAM, and Nvidia RTX 2070 GPU. The proposed method is tested on the simulation (section 6.1) and the well-known autonomous driving benchmark KITTI [9] (section 6.2). The multi-scan mode LiDAR sensors in a real environment are also tested (section 6.3). For the localization of the 3D pose in a prior map, two orthogonal laser scan sensors are inferred to be sufficient. This configuration is evaluated and found valid for MAV; its RMSE is within the centimeter level. The IMU and LiDAR on the device are calibrated offline. The IMU, which is only used in LIOMapping [21] for comparison purposes, is unnecessary for our proposed algorithm.

#### 6.1 Simulation

To evaluate the proposed contribution theory, a two-frame registration simulation is implemented. Transformation and PCD are randomly generated, and matching pairs are fixed to avoid incorrect data association. The steps involved are as follows.

1. LiDAR points are randomly generated from 1 m to 100 m in a circle with 64 scans. The PCD density satisfies the Velodyne HDL64 angle resolution.
2. Every point is randomly allocated to a pattern model, a plane (three points), or a line (two point). The measured points remain associated with their model, and matches are not changed during optimization.
3. Transformation \( T_{GT} \) is randomly generated as the ground truth, and the PCD is moved to a new location.
4. Matching models and observations are derived. The nonlinear solver uses the Gauss-Newton method and sets
the initial pose as an identity matrix. The proposed algorithm is then applied to derive results $T_{ours}$ and record the residual term size.

(5) Finally, the same size terms are randomly selected and inputted into a nonlinear solver to obtain result $T_{random}$. Translation error between $||T_{GT} - T_{ours}||_{trans}$ and $||T_{GT} - T_{random}||_{trans}$ are compared.

The foregoing registration is repeated 100 times and then the average error is derived. The resulting variation curves after the separate addition of random disturbances to measurement points and models are shown in Fig. 20.

The disturbance amplitude grows along the horizontal axis. When the disturbance is zero, data associations are accurate and do not change during the optimization. The proposed and random methods both converge to zero. The proposed method gradually increases with the disturbance amplitude. The random method probably selects some large error residual terms that considerably deviates from optimization. The proposed method can calculate and sort all terms based on contributions; thus, selected terms are optimal. This simulation demonstrates the proposed theory of contribution influences on estimation.

### 6.2 KITTI Benchmark

The KITTI benchmark is a well-known benchmark in autonomous driving [2]. It was created by the Karlsruhe Institute of Technology and Toyota Technological Institute (Chicago). It has Velodyne LiDAR (64 scans), two gray cameras, and two color camera. GPS and IMU are used for ground truth. It provides 11 sequences with ground truths in urban, city, natural, and highway environments and has been wildly used for VO and LO evaluation.

The proposed algorithm is applied to ALOAM and run on KITTI. In the first test, it is directly introduced into its mapping thread; this means that our method’s potential residual terms are a subset of the original code used. The result are summarized in Table 1. On average, the proposed method employs fewer planes and lines for mapping registration compared with the original. Although the information it uses is a subset of the original ALOAM, improvements in seven sequences are achieved. In the other four sequences, the proposed method is not superior. Nevertheless, it only fell behind by approximately 0.05%, including 0.2 cm drift over 100 m. The reasons for this are explained after next test.

This subset is presumed to limit the effects of the proposed method. Accordingly, in the second test, we changed the threshold in surf corner points detection and then used twice amounts of potential of plane and line for contribution selection. The resulting paths are illustrated in Fig. 21. As summarized in Table 2, accuracy improves by approximately 0.2 m, with advancements in ten sequences. Moreover, only approximately half of the number of planes and lines are used to obtain this result. These two test substantially demonstrate the validity of our theory. For sequence 10 in the second test, the checking of LiDAR frame’s PCD is implemented to identify shortages, as shown in Fig. 22. The car in this sequence is traversing a wild field road and with bushes on both sides. This environment means that the LiDAR’s current observations are restricted to a local region. Ours calculation yields similar scores. The difference depends on the residuals selected by the algorithm; particularly considering how small they are. The proposed method trades off robustness to achieve accuracy. Error matches have a stronger influence; thus, falling behind by 7 cm is possible under this extreme environment.

MLOAM [11] is a relevant work in this field; it is designed for a multi-LiDAR system. Its results are compared on KITTI, as summarized in Table 3. Only sequence 02 falls behind, retarding the proposed method; the other sequences have better results. The contribution is decoupled into uncertainty and sensitivity, enabling the consideration of a more accurate map model to improve accuracy. As for sequence 02, the proposed method is implemented to a road corner because of fast turning. From this point onward, it is considerably misled, resulting a large path error. The MLOAM is originally intended for multi-LiDAR sensors. Considering the limitation on the length of this article, following comparison with real data focuses on single LiDAR LO baseline ALOAM.

### 6.3 Real Data

The contribution theory is applied to ALOAM and LiDAR mapping for the captured real data. The collection device...
consists of two sets. The first is a Velodyne LiDAR (Puck VLP16) with IMU (Xsens MTI-100), as shown in Fig. 23(a). The second is a Robosense LiDAR (Blind Spot) in Fig. 23(b). These two LiDARs have completely different scan modes. The horizon of VLP16 is $360^\circ$, and the BS LiDAR is a half-sphere window with 32 scans (designed for blind spots on the front or back of a car). The environments involve walking inside building with corridors. The capturing tour starts and ends at the same location. The same path is captured five times to be more credible.

**TABLE 1**
KITTI subset points

| Sequence | FrameNum | Average |
|----------|----------|---------|
| 00       | 4541     | 1101    |
| 01       | 1461     | 1010    |
| 02       | 461      | 271     |
| 03       | 271      | 1101    |
| 04       | 1101     | 1101    |
| 05       | 4071     | 1591    |
| 06       | 1201     | 1201    |

ALOAM/m 0.7556% 1.9629% 4.5316% 0.9507% 0.7201% 0.5421% 0.6053% 0.4203% 1.0482% 0.7235% 1.7318%

| FrameNum | LineNum | PlaneNum |
|----------|---------|----------|
| 4541     | 1051    | 4177     |
| 1101     | 1477    | 1070     |
| 461      | 271     | 1101     |
| 271      | 1304    | 1166     |
| 1101     | 1414    | 1050     |
| 1101     | 1214    | 1201     |
| 4071     | 700     | 674      |
| 1591     | 690     | 670      |

**TABLE 2**
KITTI twice potential points

| Sequence | FrameNum | Average |
|----------|----------|---------|
| 00       | 4541     | 1101    |
| 01       | 1461     | 1010    |
| 02       | 461      | 271     |
| 03       | 271      | 1101    |
| 04       | 1101     | 1101    |
| 05       | 4071     | 1591    |
| 06       | 1201     | 1201    |

ALOAM/m 0.7244% 1.9339% 4.4820% 0.8636% 0.7295% 0.4828% 0.5798% 0.4493% 1.0662% 0.6609% 1.7041%

| FrameNum | LineNum | PlaneNum |
|----------|---------|----------|
| 4541     | 1051    | 700      |
| 1101     | 1070    | 674      |
| 461      | 271     | 690      |
| 271      | 1304    | 670      |
| 1101     | 700     | 670      |

**TABLE 3**
KITTI MLOAM

| Sequence | FrameNum | Average |
|----------|----------|---------|
| 00       | 4541     | 1101    |
| 01       | 1461     | 1010    |
| 02       | 461      | 271     |
| 03       | 271      | 1101    |
| 04       | 1101     | 1101    |
| 05       | 4071     | 1591    |
| 06       | 1201     | 1201    |

MLOAM/m 1.7015% 2.3043% 2.3271% 1.0544% 1.1347% 0.8285% 1.4445% 1.4053% 1.0679% 1.5106% 1.9189% 1.6152%

| FrameNum | LineNum | PlaneNum |
|----------|---------|----------|
| 4541     | 1051    | 1101     |
| 1101     | 1070    | 1101     |
| 461      | 271     | 1101     |
| 271      | 1304    | 1101     |
| 1101     | 1414    | 1101     |
| 1101     | 1214    | 1201     |
| 4071     | 700     | 674      |
| 1591     | 690     | 670      |

**TABLE 4**
VLP16 LO indoor

| Environment | indoor (loop closure error/m) |
|-------------|------------------------------|
| Sequence    | 00   | 01   | 02   | 03   | 04   |
| ALOAM       | 0.0334 | 0.0163 | 0.1131 | 0.0338 | 6.7250 |
| ALOAM(Ours) | 0.0274 | 0.0184 | 0.0654 | 0.0162 | 0.0400 |

**6.3.1 VLP16 LO Indoor**

Indoor loop closure errors are summarized in TABLE 4. Because the LiDAR range is sufficiently long to measure the far wall, the LO drift is within the centimeter level. Comparing with ALOAM, the proposed method achieves better results in four sequences. In sequence 01, because ALOAM’s translation is virtually $1$ cm, we believe that the accuracy of the start and end locations of this sequence is not sufficient for evaluation. Sequence 04 involves walking into a rest room. More surrounding points are used by ALOAM, but the proposed method uses more points outside door and window. The ALOAM drift is distinct, whereas that of the proposed method is low. The path and map details are presented in Fig. 24.

**6.3.2 VLP16 LO Outdoor**

Outdoor loop closure errors are summarized in TABLE 5. The long outdoor path demonstrates the superiority of the proposed method. This path is approximately 1.1 km long. The proposed method achieves virtually twice the accuracy of all five sequences. These results further sup-
Drift

Room

(a) ALOAM  
(b) Ours

Fig. 24. VLP16 LO indoor drift of ALOAM is high in rest room because ALOAM use more surrounding points, whereas proposed method uses more points outside door and window.

| Environment       | outdoor (loop closure error / m) |
|-------------------|----------------------------------|
| Sequence          | 00  | 01  | 02  | 03  | 04  |
| ALOAM             | 7.8131 | 10.0851 | 8.0538 | 5.0375 | 6.9043 |
| ALOAM(Ours)       | 5.0001  | 5.8372  | 4.4030  | 3.3837  | 5.4447 |

port the analysis results of the proposed method in the KITTI benchmark. Under large-scale conditions, our method achieves more significant results. Some details regarding the wall mapping quality are shown in Fig. 25 and 26. In Fig. 25, blue represents the ALOAM building wall point cloud in the map; green is that of the proposed method; and orange is the LIOmapping result. In Fig. 26, the LiDAR’s moving path is green, and the scanned wall is red. The ALOAM building wall is thick, which means that the poses have drifts. The wall generated by the proposed method is as thin as that of LIOmapping, indicating higher accuracy; and only LO is used to reach this level.

6.3.3 VLP16 LIO Indoor

| Environment       | indoor (loop closure error / m) |
|-------------------|----------------------------------|
| Sequence          | 00  | 01  | 02  | 03  | 04  |
| LIOmapping        | 0.0352 | 0.0399 | 0.0297 | 0.0363 | 0.0163 |
| LIOmapping(Ours)  | 0.0335  | 0.0158  | 0.0198  | 0.0380  | 0.0359 |

The loop closure error is small for LIO, because the indoor path is not long as outdoor; results are summarized in TABLE 6. The path starts from a hall, and the front side of the corridor is observed; the LiDAR’s data are mainly repeated on surrounding walls. The results of LIOmapping and those of the proposed method are close. We presume that the small region restricts improvement.

6.3.4 VLP16 LIO Outdoor

| Environment       | outdoor (loop closure error / m) |
|-------------------|----------------------------------|
| Sequence          | 00  | 01  | 02  | 03  | 04  |
| LIOmapping        | 3.6/19 | 4.5632 | 5.0961 | 4.2081 | 3.6909 |
| LIOmapping(Ours)  | 3.2095  | 3.6629  | 4.2257  | 4.0576  | 3.1955 |

Outdoor results are summarized in TABLE 7. The path is virtually 1.1 km long. The average improvement accuracy of the proposed method is approximately 0.5 m. This means that the application of our theory to an LIO system is also valuable. The path details are presented in Fig. 27. Although the start and end are at the same location, they do not close. Moreover, the same car becomes two cars in the built map. Because LIOmapping does not have a loop-closing method, this error cannot be eliminated.

Fig. 26. Real wall scene: Green is LiDAR moving path; red is a scanned wall; blue indicates LiDAR scanning surface, far or near.

Fig. 25. The building map quality. Green is that of the proposed method; blue is that of ALOAM; orange is that of LIOmapping. ALOAM building wall is thick, indicating that poses have drifts. Wall thickness generated by proposed method is as thin as that of LIOmapping, indicating higher accuracy; and only LO is used to reach this level.

Fig. 27. VLP16 LIO outdoor details on drift.
6.3.5 Blind Spot LO Indoor

Fig. 28. BS LiDAR point distribution considerably differs from that of Velodyne VLP16. Many points lie in small region within 5 m in front of LiDAR. Traditional LO uses more points in this region, but distant points on wall are more suitable for estimation.

Fig. 29. BS LiDAR is directed toward ceiling, which is less likely to be scanned by VLP16; chandelier is well reconstructed.

Fig. 30. BS LiDAR indoor built map. Map of ground floor and ceiling generated by blue ALOAM are distorted because poses drift are at meter level. Green obtained by the proposed method can be maintained with considerably lower drift.

TABLE 8
BS LO indoor

| Environment | indoor (loop closure error/m) |
|-------------|-------------------------------|
| Sequence    | 00   | 01   | 02   | 03   | 04   |
| ALOAM       | 2.9735 | 3.2668 | 3.2795 | 2.1737 | 5.0697 |
| ALOAM(Ours) | 0.1301 | 0.1012 | 0.0886 | 0.7348 | 0.1299 |

6.3.6 Blind Spot LO Outdoor

Tests were also conducted outdoors. Because of BS LiDAR scan characteristics, distant points are extremely sparse, and near points are considerably dense. According to our contribution theory, the estimation accuracy is significantly lower than that of VLP16. The results summarized in TABLE 9 explain this phenomenon. The translation error exceeds that of VLP16. Although certain improvements are achieved, the resulting errors are extremely considerable to ignore. Thus, this BS LiDAR is unsuitable for outdoor SLAM. Meanwhile, SLAM requires a LiDAR sensor capable of capturing distant points, which are more favorable for estimation.

TABLE 9
BS LO outdoor

| Environment | outdoor (loop closure error/m) |
|-------------|-------------------------------|
| Sequence    | 00   | 01   | 02   | 03   | 04   |
| ALOAM       | 4.4345 | 19.3979 | 17.8245 | 20.5416 | 43.1282 |
| ALOAM(Ours) | 3.1535 | 13.5462 | 15.1132 | 12.7554 | 52.5632 |

6.3.7 Orthogonal Laser Scan

This is an inference from our contribution theory. A single-scan laser is insufficient to localize the 3D pose. If another laser is placed orthogonally to the first one, the three rotations and three translations may be limited; this configuration is illustrated in Fig. 31(a). Its weight is approximately 400 g, and power is 12 V. Compared with VLP16, it is considerably cheaper and lighter; it also has low power cost. Compared with cameras, their measurements are more surface is well reconstructed. However, it can not scan the floor up to the ceiling; this explains the absence of ground points in Fig. 29. Upon entering the corridor, the BS LiDAR is placed toward the front to measure more points. In Fig. 28 the BS LiDAR point distribution show extremely differs from that of VLP16. Virtually 99% of the laser points lie in a small region within 5 m in front of the LiDAR. Traditional LOs use more points, but points distant from the wall are more useful for estimation. In our contribution theory selection, more suitable points are selected. The results are summarized in TABLE 8; those of the proposed method are distinct. The accuracy of BS LO improves from the meter to the decimeter level. The building map details are shown in Fig. 30. When we return to the hall, the floor and ceiling of the blue ALOAM map are distorted because poses drift are at the meter level. Green obtained by the proposed method can be maintained using the starting PCD with considerably lower drift.
TABLE 10
orthogonal laser scan

| Environment     | Sequence | Parameter pd | 00 | 01 | 02 | 03 | 04 | 05 |
|-----------------|----------|--------------|----|----|----|----|----|----|
| indoor with Motive Track |          | pd = 30      | 0.3218 | 144 | 0.0991 | 161 | Failed | 163 | 0.5779 | 155 | 0.0994 | 155 | 0.1602 | 156 |
|                 |          | pd = 60      | 0.1425 | 265 | 0.2086 | 296 | 0.0899 | 293 | 0.3222 | 277 | 0.1151 | 291 | 0.1291 | 279 |
|                 |          | pd = 100     | 0.0662 | 416 | 0.0797 | 450 | 0.0986 | 436 | 0.0696 | 414 | 0.1134 | 473 | 0.0712 | 419 |
|                 |          | pd = 200     | 0.0626 | 738 | 0.0546 | 759 | 0.0932 | 698 | 0.0395 | 699 | 0.1074 | 782 | 0.0630 | 704 |
|                 |          | pd = 300     | 0.0629 | 939 | 0.0500 | 992 | 0.0833 | 927 | 0.0915 | 900 | 0.0958 | 1058 | 0.0762 | 905 |
|                 |          | pd = 400     | 0.0572 | 1076 | 0.0480 | 1191 | 0.0807 | 1137 | 0.1768 | 1060 | 0.0889 | 1213 | 0.0583 | 1082 |
|                 |          | pd = 500     | 0.0546 | 1196 | 0.0479 | 1336 | 0.0798 | 1326 | 0.1745 | 1205 | **0.0865** | 1379 | **0.0579** | 1218 |
|                 |          | pd = 600     | 0.0515 | 1267 | 0.0472 | 1457 | 0.0795 | 1405 | 0.1610 | 1281 | 0.1142 | 1580 | 0.0587 | 1321 |
|                 |          | pd = 700     | 0.0492 | 1359 | 0.0474 | 1521 | 0.0825 | 1467 | 0.0361 | 1348 | 0.0988 | 1683 | 0.0583 | 1409 |
|                 |          | pd = 800     | 0.0467 | 1395 | 0.0463 | 1618 | 0.0841 | 1559 | 0.0374 | 1410 | 0.0971 | 1719 | 0.0580 | 1465 |
|                 |          | pd = 900     | 0.0326 | 1438 | 0.0459 | 1653 | 0.0842 | 1616 | 0.1403 | 1466 | 0.1464 | 1774 | 0.0578 | 1513 |
|                 |          | pd = 1000    | 0.0351 | 1446 | 0.0458 | 1712 | 0.0838 | 1644 | 0.1414 | 1501 | 0.1457 | 1864 | 0.0581 | 1554 |
|                 |          | pd = 1100    | 0.0405 | 1490 | **0.0457** | 1722 | 0.0837 | 1685 | 0.1306 | 1518 | 0.1393 | 1883 | 0.0583 | 1578 |
|                 |          | pd = 1200    | 0.0494 | 1504 | 0.0462 | 1730 | 0.0826 | 1700 | 0.1342 | 1538 | 0.1304 | 1898 | 0.0583 | 1588 |
|                 |          | pd = 1300    | 0.1555 | 1541 | 0.0464 | 1771 | 0.0830 | 1724 | 0.1367 | 1557 | 0.1359 | 1909 | 0.0580 | 1607 |
|                 |          | pd = 1400    | 0.0525 | 1587 | 0.0464 | 1780 | 0.0835 | 1749 | 0.1429 | 1582 | 0.1358 | 1917 | 0.0581 | 1620 |

Fig. 31. (a) Third collecting device: RPLiDAR A3 and Hokuyo UTM-30LX; both are single-scan laser. (b) Orthogonal laser scan points localize in a prior map. Orthogonal laser scans are blue points. Their external parameters are manually calibrated offline.

Fig. 32. Orthogonal scan laser motion paths and ground truths captured by OptiTrack.

7 CONCLUSION

In this paper, we propose a theory of observation contribution for pose estimation. We demonstrate that the contribution model can be decoupled into observation uncertainty and sensitivity models. Our selection method is the global statistical optimal method. To explain the method in details, LiDAR measurement uncertainties and fusing mechanisms are calculated, and residual contributions are analyzed and decoupled into six dimensions. Then, the algorithm sorts and selects residuals for optimization. The experiments reveal that superior pose estimation accuracy is achieved.

Because our theory is a general model for observation contributions, it is designed to solve nonlinear optimization problems. Sensors with multiple measurements in
single-shot sampling are suitable for capturing considerable contribution measurements. The selected observations can achieve high optimization accuracy, and the small quantity of residuals leads to a low computational cost.

Thus far, the problem of data association has not been encountered. This work has adopted traditional methods in LiDAR, relying on a uniform motion model or IMU, which is the neighborhood principle in ICP for matching current points and referent points. Because this work concentrates on improving pose estimation accuracy, a uniform motion model in walking or low-speed driving is found sufficient. The proposed observation contribution theory attempts to select residual terms with small uncertainties and higher sensitivities. This method fundamentally decreases the robustness of the pose estimation system to increase its accuracy; this is the reason for trading off robustness for accuracy. To improve the accuracy of pose estimation from another perspective, our next objective is to investigate against disturbances in pose estimation.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant No.62088101, 61673341), the Project of State Key Laboratory of Industrial Control Technology, Zhejiang University, China (No.ICT2021A10), the Fundamental Research Funds for the Central Universities (No.2021FZZX003-01-06), National Key R&D Program of China (2016YFD020701-3), Double First Class University Plan (CN). Zeyu Wan and Yu Zhang contributed equally to this work.

REFERENCES

[1] D. Wooden, M. Malchano, K. Blankepoo, A. Howard, A. Rizzi, and M. Raibert, “Autonomous navigation for bigdog,” 2010 IEEE International Conference on Robotics and Automation, pp. 4736–4741, 2010.

[2] T. Barfoot, “State estimation for robotics,” 2017.

[3] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, J. Reid, and J. Leonard, “Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age,” IEEE Transactions on Robotics, vol. 32, pp. 1309–1332, 2016.

[4] N. Gelfand, S. Rusinkiewicz, L. Ikemoto, and M. Levoy, “Geometrically stable sampling for the icp algorithm,” Fourth International Conference on 3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings., pp. 260–267, 2003.

[5] J.-E. Deschaud, “Imls-slam: Scan-to-model matching based on 3d data,” 2018 IEEE International Conference on Robotics and Automation (ICRA), pp. 2480–2485, 2018.

[6] Y. Zhao and P. Wan, “Good feature matching: Toward accurate, robust vo/vslam with low latency,” IEEE Transactions on Robotics, vol. 36, pp. 657–675, 2020.

[7] A. Fontán, J. Civera, and R. Triebel, “Information-driven direct rgb-d odometry,” 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4928–4936, 2020.

[8] W. Dai, Y. Zhang, P. Li, Z. Fang, and S. Scherer, “Rgb-d slam in dynamic environments using point correlations,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1–1, 2020.

[9] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? the kitti vision benchmark suite,” in Conference on Computer Vision and Pattern Recognition (CVPR), 2012.

[10] Y. Zhao, J. Smith, and P. Vela, “Good graph to optimize: Cost-effective, budget-aware bundle adjustment in visual slam,” ArXiv, vol. abs/2008.10123, 2020.

[11] J. Jiao, Y. Zhu, H. Ye, H. Huang, P. Yun, L. Jiang, L. Wang, and M. Liu, “Greedy-based feature selection for efficient lidar slam,” ArXiv, vol. abs/2103.13090, 2021.

[12] C. Shannon, “A mathematical theory of communication,” Bell Syst. Tech. J., vol. 27, pp. 379–423, 1948.

[13] P. Besl and N. D. McKay, “A method for registration of 3-d shapes,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 14, pp. 239–256, 1992.

[14] G. Grisetti, C. Stachniss, and W. Burgard, “Improved techniques for grid mapping with rao-blackwellized particle filters,” IEEE Transactions on Robotics, vol. 23, pp. 3–14, 2007.

[15] J. Zhang and S. Singh, “Loam: Lidar odometry and mapping in real-time,” in Robotics: Science and Systems, 2014.

[16] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, “Monoslam: real-time single camera slam,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 1052–1067, 2007.

[17] G. Klein and D. Murray, “Parallel tracking and mapping for small ar workspaces,” 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, pp. 225–234, 2007.

[18] M. Puterman, “Markov decision processes: Discrete stochastic dynamic programming,” in Wiley Series in Probability and Statistics, 1994.

[19] J. Behley and C. Stachniss, “Efficient surfel-based slam using 3d laser range data in urban environments,” in Robotics: Science and Systems, 2018.

[20] X. Chen, A. M. E. Palazzolo, P. Giguère, J. Behley, and C. Stachniss, “Suma++: Efficient lidar-based semantic slam,” 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4530–4537, 2019.

[21] H. Ye, Y. Chen, and M. Liu, “Tightly coupled 3d lidar inertial odometry and mapping,” 2019 International Conference on Robotics and Automation (ICRA), pp. 3144–3150, 2019.

[22] M. Yokozuka, K. Koide, S. Oishi, and A. Banno, “Litamin: Lidar-based tracking and mapping by stabilized icp for geometry approximation with normal distributions,” 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5143–5150, 2020.

[23] M. Yokozuka, K. Koide, S. Oishi, and A. Banno, “Litamin2: Ultra light lidar-based slam using geometric approximation applied with kl-divergence,” ArXiv, vol. abs/2103.00784, 2021.

[24] Y. Pan, P. Xiao, Y. He, Z. Shao, and Z. Li, “Mulls: Versatile lidar slam via multi-metric linear least square,” ArXiv, vol. abs/2102.03771, 2021.

[25] Z. Liu and F. Zhang, “Balm: Bundle adjustment for lidar mapping,” IEEE Robotics and Automation Letters, vol. 6, pp. 3184–3191, 2021.

[26] J. Engel, T. Schöps, and D. Cremers, “Lsd-slam: Large-scale direct monocular slam,” in ECCV, 2014.

[27] C. Forster, M. Pizzoli, and D. Scaramuzza, “Svo: Fast semi-direct monocular visual odometry,” 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 15–22, 2014.

[28] R. Mur-Artal and J. D. Tardós, “Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras,” IEEE Transactions on Robotics, vol. 33, pp. 1255–1262, 2017.

[29] J. Engel, V. Koltun, and D. Cremers, “Direct sparse odometry,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, pp. 611–625, 2018.

[30] M. Kaess, A. Ranganathan, and F. Dellaert, “isAM: Incremental smoothing and mapping,” IEEE Trans. on Robotics (TRO), vol. 24, pp. 1363–1378, Dec. 2008.

[31] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, “isAM2: Incremental smoothing and mapping with fluence linearization and incremental variable reordering,” in IEEE Intl. Conf. on Robotics and Automation (ICRA), (Shanghai, China), pp. 3281–3288, May 2011.

[32] T. Shan and B. Englot, “Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain,” 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4758–4765, 2018.

[33] X. Wang, R. J. Marcotte, G. Ferrer, and E. Olson, “Aprism: Real-time smoothing and mapping,” 2018 IEEE International Conference on Robotics and Automation (ICRA), pp. 2486–2493, 2018.

[34] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, “Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping,” 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5135–5142, 2020.

[35] A. Ben-Tal and A. Nemirovski, “Robust convex optimization,” Math. Oper. Res., vol. 23, pp. 769–805, 1998.

[36] S. P. Boyd and L. Vandenberghe, “Convex optimization,” IEEE Transactions on Automatic Control, vol. 51, pp. 1899–1895, 2006.
