Balanced ICP for precise lidar odometry from non bilateral correspondences

Matteo Azzini, Ezio Malis, Philippe Martinet

Balanced ICP for precise lidar odometry from non bilateral correspondences

Matteo Azzini, Ezio Malis and Philippe Martinet

***Abstract*— In the field of lidar odometry for autonomous navigation, the Iterative Closest Point (ICP) algorithm is a prevalent choice for estimating robot motion by comparing point clouds. However, ICP accuracy is strictly dependent on the nature of the features involved, but also on the directional choice of the extraction and matching, either from the current to the reference point cloud or vice-versa. Point-to-line or point-to-plane correspondences have been proven to provide the more accurate odometry results. The matching is generally done in a mono-directional framework: extract the features (lines or planes) in the current point cloud and match them to points in the reference point cloud. This paper introduces a novel formulation, named Balanced ICP, that performs feature extraction (lines or planes) in both point clouds and consequent matching in both the directions. Therefore, the cost function is designed to perform a simultaneous optimization of all available data balancing the noise and extraction errors. The experiments, conducted both on simulated and real data from the KITTI dataset, reveal that our method outperform the classical mono-directional formulations, in terms of robustness, accuracy and stability.**

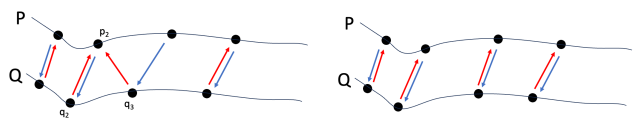
I. INTRODUCTION

Nowadays, lidar sensors gained huge popularity thanks to their ability to generate high-resolution 3D point clouds. Its application in various domains, such as autonomous navigation, robot manipulation or 3D reconstruction, and their rapid development constantly require increasingly accurate localization. To meet this need, in the context of lidar odometry, the Iterative Closest Point (ICP) algorithm [1] emerges as the most popular solution to estimate the motion of the robot by comparing the current point cloud with the previous one, or with a saved reference map. Even if it is widely exploited, the accuracy of this method is strictly dependent on the nature of the features involved, but also on the directional choice of the extraction and matching, either from the current to the reference point cloud or vice-versa.

To overcome these limitations different techniques were attempted, either keeping the same algorithm formulation while focusing on another block of the odometry pipeline [2], or trying to work on the feature extraction. The most popular example of this second approach is LOAM [3], which is based on the extraction of lines and planes and on the matching of those features. Inspired by the same idea, [4] insisted on the same geometric features, but they tried to improve accuracy by enriching the matching with ground constraints and splitting into two odometry pipelines the pose estimation, for ground and not ground points. A similar philosophy is followed by [5], but complexifying the features extraction in order to classify the points in different categories (ground, facade, pillar, beam, etc.) and use this information to weight the matching in the cost function. Even more elaborated is the approach of [6] where the authors proposed the couples of points and normals as features, enriched also by semantic classification obtained by RangeNet [7], used then to weight the matching like [5].

Other works focused more on the cost function in order to obtain a more precise estimation. For example, [8] proposed a distribution-based feature extraction and a modified cost function that takes into account the KL-divergence between the two distributions. A different approach is the work proposed by [9] where the authors presented a point-to-plane formulation of the ICP, enriched by terms that constrain the smoothness of the motion and the constant velocity assumption.

All the previous works are based on the extraction of features in the current point cloud and then matching in the reference point cloud, therefore using a mono-directional approach. Differently, the authors of [10] guessed the possibility of exploiting a double matching information and they proposed a method based on the bilateral correspondences, namely when a point in the current point cloud is matched with a point in the reference point cloud and that point is matched with the other one uniquely in the first point cloud (see Fig.1).



(a) Non bilateral matching (b) Bilateral matching

Fig. 1: Bilateral correspondences

In their case, they exploit the bidirectionality, in the sense that the matching is performed in both directions, from the current to the reference and vice-versa, and the bilateral correspondences are used to weight the cost function. This approach provides a more precise estimation of the motion, but their formulation first applies just to point-to-point correspondences and secondly, the minimization involves just the transformation in one direction, from the current to the reference point cloud.

Leveraging a similar intuition, but with a different proposal, the authors of [11] presented another version, called

This work has been supported by the AISENSE project.

Authors are with ACENTAURI team at Centre Inria d'Université Côte d'Azur, Sophia-Antipolis, France matteo.azzini@inria.fr ezio.malis@inria.fr philippe.martinet@inria.fr

Symmetric ICP. In this case, keeping a mono-directional matching from the current to the reference, the symmetry is given by the fact that the cost function is defined in order to take into account the point-to-plane distances with the normals extracted for both the matched points in the current and in the reference point cloud. This formulation allows the exploitation of higher precision of a point-to-plane ICP with respect to the classic point-to-point formulation, but also includes points that are consistent with a locally-second-order surface centered between them. On the other hand, this method doesn't exploit the information that a bidirectional matching can provide, consequently, the minimization still involves just the transformation from the first to the second point cloud.

In this work we present a novel formulation of the ICP algorithm, named Balanced ICP, that, in contrast with the other approaches, exploits the features extraction and matching in both the point clouds to define a cost function, which includes the motion estimation in both directions, reference to current and vice-versa. By performing the simultaneous optimization of all available data, our formulation is able to balance the additional error inherited from the inevitable dependency of the feature extraction and matching processes on the noisy point cloud. To the best of our knowledge, this is the first time that a method fully exploits the bi-laterality of the scan registration.

II. Theoretical background

In this section, we recall the ICP algorithm for different features that can be extracted from a 3D point cloud: points, lines and planes.

A. Iterative Closest Point algortihms

We separate the algorithms into two categories. The first category (see II-A.1) is obtained when considering bilateral features in the reference and current point cloud: point-to-point, line-to-line and plane-to-plane. The second category (see II-A.2) is obtained by considering features of different nature in the reference and current point cloud: point-to-line and point-to-plane

1) Bilateral Features:

a) Point-to-point correspondences: When we have n current noisy points $\underline{\mathbf{m}}_{ck}, k \in \{1, 2, ..., n\}$ in correspondence to n reference noisy points $\underline{\mathbf{m}}_{rk}, k \in \{1, 2, ..., n\}$ we look for the transformation matrix ${}^{r}\mathbf{T}_{c}$ that minimizes the distances between the reference points and the transformed current points:

$$ {}^{r}\mathbf{T}_{c} = \operatorname{argmin} \frac{1}{2} \sum_{k=1}^{n} w_k^2 \left\| {}^{r}\mathbf{T}_{c} \underline{\mathbf{m}}_{ck} - \underline{\mathbf{m}}_{rk} \right\|^2 \qquad (1) $$

Note that one can also define the following problem that minimizes the distances between the current points and the transformed reference points:

$$ {}^{c}\mathbf{T}_{r} = \operatorname{argmin} \frac{1}{2} \sum_{k=1}^{n} w_k^2 \left\| \underline{\mathbf{m}}_{ck} - {}^{c}\mathbf{T}_{r} \underline{\mathbf{m}}_{rk} \right\|^2 \qquad (2) $$

These problems are naturally bilateral and the solutions of the two separated minimizations are such that:

$$ {}^{r}\mathbf{T}_{c} = {}^{c}\mathbf{T}_{r}^{-1} $$

even in the presence of noisy data.

b) Line-to-line correspondences: When we have n_r current lines with origin the point $\underline{\mathbf{o}}_{ck}$ and direction the vector $\mathbf{d}_{ck}$, in correspondence to n reference noisy lines with origin the point $\underline{\mathbf{o}}_{rk}$ and direction the vector $\mathbf{d}_{rk}$, we look for the transformation matrix ${}^{r}\mathbf{T}_{c}$ that minimizes the distances between the reference points and the transformed current points:

$$ \begin{aligned} {}^{r}\mathbf{T}_{c} &= \operatorname{argmin} \frac{1}{2} \sum_{k=1}^{n} w_k^2 \left\| {}^{r}\mathbf{R}_{c} \boldsymbol{d}_{ck} - \boldsymbol{d}_{rk} \right\|^2 + \\ &+ \frac{1}{2} \sum_{k=1}^{n} w_k^2 \left\| {}^{r}\mathbf{R}_{c} \boldsymbol{o}_{ck} + {}^{r}\mathbf{t}_{c} - \boldsymbol{o}_{rk} \right\|^2 \end{aligned} \qquad (3) $$

Note that one can also define the following problem that minimizes the distances between the current lines and the transformed reference lines:

$$ \begin{aligned} {}^{c}\mathbf{T}_{r} &= \operatorname{argmin} \frac{1}{2} \sum_{k=1}^{n} w_k^2 \left\| \boldsymbol{d}_{ck} - {}^{c}\mathbf{R}_{r} \boldsymbol{d}_{rk} \right\|^2 + \\ &+ \frac{1}{2} \sum_{k=1}^{n} w_k^2 \left\| \boldsymbol{o}_{ck} - {}^{c}\mathbf{R}_{r} \boldsymbol{o}_{rk} - {}^{c}\mathbf{t}_{r} \right\|^2 \end{aligned} \qquad (4) $$

These problems are naturally bilateral and the solutions of the two separated minimizations are such that:

$$ {}^{r}\mathbf{T}_{c} = {}^{c}\mathbf{T}_{r}^{-1} $$

even in the presence of noisy data.

c) Plane-to-plane correspondences: When we have n current noisy planes $\boldsymbol{\pi}_{ck}, k \in \{1, 2, ..., n\}$ in correspondence to n reference noisy planes $\boldsymbol{\pi}_{rk}$, $k \in \{1, 2, ..., n\}$, we look for the transformation matrix ${}^{r}\mathbf{T}_{c}$ that minimizes the distances between the reference planes and the transformed current planes:

$$ {}^{r}\mathbf{T}_{c} = \operatorname{argmin} \frac{1}{2} \sum_{k=1}^{n} w_k^2 \left\| {}^{r}\mathbf{T}_{c}^{-\top} \boldsymbol{\pi}_{ck} - \boldsymbol{\pi}_{rk} \right\|^2 \qquad (5) $$

Note that one can also define the following problem that minimizes the distances between the current points and the transformed reference points:

$$ {}^{c}\mathbf{T}_{r} = \operatorname{argmin} \frac{1}{2} \sum_{k=1}^{n} w_k^2 \left\| \boldsymbol{\pi}_{ck} - {}^{c}\mathbf{T}_{r}^{-\top} \boldsymbol{\pi}_{rk} \right\|^2 \qquad (6) $$

These problems are naturally bilateral and the solutions of the two separated minimizations are such that:

$$ {}^{r}\mathbf{T}_{c} = {}^{c}\mathbf{T}_{r}{}^{-1} $$

even in the presence of noisy data.

2) Non Bilateral Features: When the features extracted from the reference and current point clouds are not bilateral, the result of the optimization depends on the cost function.

a) Point-to-line correspondences: Consider n_r lines in the reference frame and for each line, with origin the point $\underline{\mathbf{o}}_{ri}$ and direction the vector $\mathbf{d}_{ri}$, a corresponding point in the current frame $\underline{\mathbf{m}}_{ci}$, where $i \in \{1, 2, ..., n\}$. We look for the optimal transformation matrix ${}^{r}\mathbf{T}_c$ that minimizes the following weighted least squares problem:

$$ {}^{r}\mathbf{T}_c = \text{argmin} \frac{1}{2} \sum_{i=1}^{n_r} w_{ri}^2 \|\mathbf{D}_{ri}({}^{r}\mathbf{T}_c \underline{\mathbf{m}}_{ci} - \underline{\mathbf{o}}_{ri})\|^2 \quad (7) $$

where w_{ri} are positive weights, and

$$ \mathbf{D}_r = \begin{bmatrix} [\mathbf{d}_r]_\times^2 & 0 \\ 0 & 0 \end{bmatrix} $$

If we also have n_c lines in the current frame and for each line, with origin the point $\underline{\mathbf{o}}_{cj}$ and direction the vector $\mathbf{d}_{cj}$, a point in the reference frame $\underline{\mathbf{m}}_{rj}$, where $j \in \{1, 2, ..., n_c\}$ we can also solve the following problem:

$$ {}^{c}\mathbf{T}_r = \text{argmin} \frac{1}{2} \sum_{j=1}^{n_c} w_{cj}^2 \|\mathbf{D}_{cj}({}^{c}\mathbf{T}_r \underline{\mathbf{m}}_{rj} - \underline{\mathbf{o}}_{cj})\|^2 \quad (8) $$

where w_{cj} are positive weights, and

$$ \mathbf{D}_c = \begin{bmatrix} [\mathbf{d}_c]_\times^2 & 0 \\ 0 & 0 \end{bmatrix} $$

In this case, the features are not bilateral and in general with noisy data solving the two problems separately will give two solutions such that:

$$ {}^{r}\mathbf{T}_c \neq {}^{c}\mathbf{T}_r{}^{-1} $$

b) Point-to-plane correspondences: Consider n_r planes in the reference frame and for each plane $\boldsymbol{\pi}_{ri}$ a corresponding point in the current frame $\underline{\mathbf{m}}_{ci}$, where $i \in \{1, 2, ..., n\}$. We look for the optimal transformation matrix ${}^{r}\mathbf{T}_c$ that minimizes the following weighted least squares problem:

$$ {}^{r}\mathbf{T}_c = \text{argmin} \frac{1}{2} \sum_{i=1}^{n_r} w_{ri}^2 \|\boldsymbol{\pi}_{ri}^\top {}^{r}\mathbf{T}_c \underline{\mathbf{m}}_{ci}\|^2 \quad (9) $$

where w_{ri} are positive weights.

If we also have n_c planes in the current frame and for each plane $\boldsymbol{\pi}_{cj}$ a point in the reference frame $\underline{\mathbf{m}}_{rj}$, where $j \in \{1, 2, ..., n_c\}$ we can also solve the following problem:

$$ {}^{c}\mathbf{T}_r = \text{argmin} \frac{1}{2} \sum_{j=1}^{n_c} w_{ci}^2 \|\boldsymbol{\pi}_{cj}^\top {}^{c}\mathbf{T}_r \underline{\mathbf{m}}_{rj}\|^2 \quad (10) $$

In this case, the features are not bilateral and in general with noisy data solving the two problems separately will give two solutions such that:

$$ {}^{r}\mathbf{T}_c \neq {}^{c}\mathbf{T}_r{}^{-1} $$

III. Balanced ICP for non bilateral correspondences

In this section, an approach to balance non-bilateral correspondences is proposed by performing the simultaneous optimization of all available data, both from reference and current point clouds.

A. Cost function for non-bilateral correspondences

Consider n_r features extracted in the reference frame, along with their non bilateral matching in the current frame, and n_c features of the same nature extracted in the current frame, along with their non bilateral matching in the current frame. We look for the optimal transformation matrix $\mathbf{T}$ that minimizes the following balanced weighted least squares problem:

$$ \min_{\mathbf{T}} \frac{n_r}{n_r + n_c} \|\mathbf{e}_r(\mathbf{T})\|^2 + \frac{n_c}{n_r + n_c} \|\mathbf{e}_c(\mathbf{T})\|^2 \quad (11) $$

where $\mathbf{e}_r(\mathbf{T})$ and $\mathbf{e}_c(\mathbf{T})$ are vectors containing the errors of the features, respectively extracted in the reference and the current frame, relative to their non-bilateral matching in the other point cloud. The expression of $\mathbf{e}_r(\mathbf{T})$ and $\mathbf{e}_c(\mathbf{T})$ depends on the features and it is given in the following paragraphs.

a) point-to-line correspondences: Consider n_r lines with origin the point $\underline{\mathbf{o}}_{ri}$ and direction the vector $\mathbf{d}_{ri}$ in the reference frame and for each line a point in the current frame $\underline{\mathbf{m}}_{ci}$, where $i \in \{1, 2, ..., n_r\}$. Consider n_c lines with origin the point $\underline{\mathbf{o}}_{cj}$ and direction the vector $\mathbf{d}_{cj}$ in the current frame and for each line a point in the reference frame $\underline{\mathbf{m}}_{rj}$, where $j \in \{1, 2, ..., n_c\}$. Therefore the errors can be defined as:

$$ \begin{aligned} \mathbf{e}_r &= \begin{bmatrix} w_{r1}\mathbf{D}_{r1}(\mathbf{T}\underline{\mathbf{m}}_{c1} - \underline{\mathbf{o}}_{r1}) \\ w_{r2}\mathbf{D}_{r2}(\mathbf{T}\underline{\mathbf{m}}_{c2} - \underline{\mathbf{o}}_{r2}) \\ \vdots \\ w_{rn_r}\mathbf{D}_{rn_r}(\mathbf{T}\underline{\mathbf{m}}_{cn_r} - \underline{\mathbf{o}}_{rn_r}) \end{bmatrix} = \\ &= \begin{bmatrix} w_{r1}\,[\mathbf{d}_{r1}]_\times^2\,(\mathbf{R}\mathbf{m}_{c1} + \mathbf{t} - \mathbf{o}_{r1}) \\ w_{r2}\,[\mathbf{d}_{r2}]_\times^2\,(\mathbf{R}\mathbf{m}_{c1} + \mathbf{t} - \mathbf{o}_{r2}) \\ \vdots \\ w_{rn_r}\,[\mathbf{d}_{rn_r}]_\times^2\,(\mathbf{R}\mathbf{m}_{cn_r} + \mathbf{t} - \mathbf{o}_{rn_r}) \end{bmatrix} \end{aligned} \quad (12) $$

and

$$ \begin{aligned} \mathbf{e}_c &= \begin{bmatrix} w_{c1}\mathbf{D}_{c1}(\mathbf{T}^{-1}\underline{\mathbf{m}}_{r1} - \underline{\mathbf{o}}_{c1}) \\ w_{c2}\mathbf{D}_{c2}(\mathbf{T}^{-1}\underline{\mathbf{m}}_{r2} - \underline{\mathbf{o}}_{c2}) \\ \vdots \\ w_{cn_c}\mathbf{D}_{cn_c}(\mathbf{T}^{-1}\underline{\mathbf{m}}_{rn_c} - \underline{\mathbf{o}}_{cn_c}) \end{bmatrix} = \\ &= \begin{bmatrix} w_{c1}\mathbf{D}_{c1}(\mathbf{R}^\top\mathbf{m}_{r1} - \mathbf{R}^\top\mathbf{t} - \underline{\mathbf{o}}_{c1}) \\ w_{c2}\mathbf{D}_{c2}(\mathbf{R}^\top\mathbf{m}_{r2} - \mathbf{R}^\top\mathbf{t} - \underline{\mathbf{o}}_{c2}) \\ \vdots \\ w_{cn_c}\mathbf{D}_{cn_c}(\mathbf{R}^\top\mathbf{m}_{rn_c} - \mathbf{R}^\top\mathbf{t} - \underline{\mathbf{o}}_{cn_c}) \end{bmatrix} \end{aligned} \quad (13) $$

where w_{ri} and w_{cj} are positive weights.

b) Point to plane correspondences: Consider n_r planes in the reference frame and for each plane $\boldsymbol{\pi}_{ri}$ a point in the current frame $\underline{\mathbf{m}}_{ci}$, where $i \in \{1, 2, ..., n_r\}$. Consider n_c planes in the current frame and for each plane $\boldsymbol{\pi}_{cj}$ a point in the reference frame $\underline{\mathbf{m}}_{rj}$, where $j \in \{1, 2, ..., n_c\}$. Therefore

the errors can be defined as:

$$\mathbf{e}_r = \begin{bmatrix} w_{r1}(\boldsymbol{\pi}_{r1}^\top \mathbf{T}\underline{\mathbf{m}}_{c1}) \\ w_{r2}(\boldsymbol{\pi}_{r2}^\top \mathbf{T}\underline{\mathbf{m}}_{c2}) \\ \vdots \\ w_{r_n r}(\boldsymbol{\pi}_{rn_r}^\top \mathbf{T}\underline{\mathbf{m}}_{cn_r}) \end{bmatrix} = \tag{14}$$

$$= \begin{bmatrix} w_{r1}(\mathbf{n}_{r1}^\top \mathbf{R}\mathbf{m}_{c1} + \mathbf{n}_{r1}^\top \mathbf{t} + d_{r1}) \\ w_{r2}(\mathbf{n}_{r2}^\top \mathbf{R}\mathbf{m}_{c2} + \mathbf{n}_{r2}^\top \mathbf{t} + d_{r2}) \\ \vdots \\ w_{r_n r}(\mathbf{n}_{rn_r}^\top \mathbf{R}\mathbf{m}_{cn_r} + \mathbf{n}_{rn_r}^\top \mathbf{t} + d_{rn_r}) \end{bmatrix}$$

and:

$$\mathbf{e}_c = \begin{bmatrix} w_{c1}(\boldsymbol{\pi}_{c1}^\top \mathbf{T}^{-1}\underline{\mathbf{m}}_{r1}) \\ w_{c2}(\boldsymbol{\pi}_{c2}^\top \mathbf{T}^{-1}\underline{\mathbf{m}}_{r2}) \\ \vdots \\ w_{cn_c}(\pi_{cn_c}^\top \mathbf{T}^{-1}\underline{\mathbf{m}}_{rn_c}) \end{bmatrix} = \tag{15}$$

$$= \begin{bmatrix} w_{c1}(\mathbf{n}_{c1}^\top \mathbf{R}^\top \mathbf{m}_{r1} - \mathbf{n}_{c1}^\top \mathbf{R}^\top \mathbf{t} + d_{c1}) \\ w_{c2}(\mathbf{n}_{c2}^\top \mathbf{R}^\top \mathbf{m}_{r2} - \mathbf{n}_{c2}^\top \mathbf{R}^\top \mathbf{t} + d_{c2}) \\ \vdots \\ w_{cn_c}(\mathbf{n}_{cn_c}^\top \mathbf{R}^\top \mathbf{m}_{rn_c} - \mathbf{n}_{cn_c}^\top \mathbf{R}^\top \mathbf{t} + d_{cn_c}) \end{bmatrix}$$

where w_{ri} and w_{cj} are positive weights.

B. Iterative optimization

We optimise the cost function iteratively starting from an initial guess of the pose. We consider the optimization problem as a virtual sensor-based control problem by considering the error vectors $\mathbf{e}_r$ and $\mathbf{e}_c$. We have that:

$$\dot{\mathbf{e}}_r = \mathbf{L}_r \mathbf{v}$$

$$\dot{\mathbf{e}}_c = \mathbf{L}_c \mathbf{v}$$

where $\mathbf{v} \in \mathbb{R}^6$ is the velocity of a virtual lidar and $\mathbf{L}_r$ and $\mathbf{L}_c$ are called the interaction matrices. Stacking together both errors and interaction matrices we get:

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_r \\ \mathbf{e}_c \end{bmatrix}$$

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_r \\ \mathbf{L}_c \end{bmatrix}$$

and

$$\dot{\mathbf{e}} = \mathbf{L}\mathbf{v}$$

We choose as control law:

$$\mathbf{v} = -\lambda(\mathbf{L}^\top \mathbf{L})^{-1}\mathbf{L}^\top \mathbf{e}$$

where λ is a constant positive scalar factor that tunes the amplitude of the incremental step. Therefore, we get a Gauss-Newton like minimisation approach. The new pose is then computed as:

$$\mathbf{T}_{k+1} = \mathbf{T}_k \exp([\mathbf{v}])$$

and the process is iterated until convergence.

a) point-to-line correspondences: In the specific case of point-to-line correspondences, recalling the previously defined notation, the interaction matrices $\mathbf{L}_r$ and $\mathbf{L}_c$ are defined as:

$$\mathbf{L}_r = \begin{bmatrix} w_{r1}\begin{bmatrix} [\mathbf{d}_{r1}]_\times^2 \mathbf{R} & -[\mathbf{d}_{r1}]_\times^2 \mathbf{R}\,[\mathbf{m}_{c1}]_\times \end{bmatrix} \\ w_{r2}\begin{bmatrix} [\mathbf{d}_{r2}]_\times^2 \mathbf{R} & -[\mathbf{d}_{r2}]_\times^2 \mathbf{R}\,[\mathbf{m}_{c2}]_\times \end{bmatrix} \\ \vdots \\ w_{rn_r}\begin{bmatrix} [\mathbf{d}_{rn_r}]_\times^2 \mathbf{R} & -[\mathbf{d}_{rn_r}]_\times^2 \mathbf{R}\,[\mathbf{m}_{cn_r}]_\times \end{bmatrix} \end{bmatrix}$$

$$\mathbf{L}_c = -\begin{bmatrix} w_{c1}\begin{bmatrix} [\mathbf{d}_{c1}]_\times^2 & [\mathbf{d}_{c1}]_\times^2 \left[\mathbf{R}^\top(\mathbf{t}-\mathbf{m}_{r1})\right]_\times \end{bmatrix} \\ w_{c2}\begin{bmatrix} [\mathbf{d}_{c2}]_\times^2 & [\mathbf{d}_{c2}]_\times^2 \left[\mathbf{R}^\top(\mathbf{t}-\mathbf{m}_{r2})\right]_\times \end{bmatrix} \\ \vdots \\ w_{cn_c}\begin{bmatrix} [\mathbf{d}_{cn_c}]_\times^2 & [\mathbf{d}_{cn_c}]_\times^2 \left[\mathbf{R}^\top(\mathbf{t}-\mathbf{m}_{rn_c})\right]_\times \end{bmatrix} \end{bmatrix}$$

b) Point to plane correspondences: In the specific case of point-to-plane correspondences, recalling the previously defined notation, the interaction matrices $\mathbf{L}_r$ and $\mathbf{L}_c$ are defined as:

$$\mathbf{L}_r = \begin{bmatrix} w_{r1}\begin{bmatrix} \mathbf{n}_{r1}^\top \mathbf{R} & -\mathbf{n}_{r1}^\top \mathbf{R}\,[\mathbf{m}_{c1}]_\times \end{bmatrix} \\ w_{r2}\begin{bmatrix} \mathbf{n}_{r2}^\top \mathbf{R} & -\mathbf{n}_{r2}^\top \mathbf{R}\,[\mathbf{m}_{c2}]_\times \end{bmatrix} \\ \vdots \\ w_{rn_r}\begin{bmatrix} \mathbf{n}_{rn_r}^\top \mathbf{R} & -\mathbf{n}_{rn_r}^\top \mathbf{R}\,[\mathbf{m}_{cn_r}]_\times \end{bmatrix} \end{bmatrix}$$

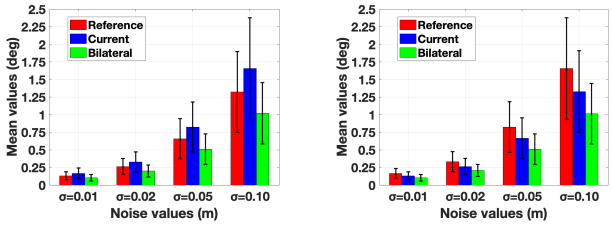
$$\mathbf{L}_c = \begin{bmatrix} w_{c1}\begin{bmatrix} -\mathbf{n}_{c1}^\top & \mathbf{n}_{c1}^\top \left[\mathbf{R}^\top(\mathbf{m}_{r1}-\mathbf{t})\right]_\times \end{bmatrix} \\ w_{c2}\begin{bmatrix} -\mathbf{n}_{c2}^\top & \mathbf{n}_{c2}^\top \left[\mathbf{R}^\top(\mathbf{m}_{r2}-\mathbf{t})\right]_\times \end{bmatrix} \\ \vdots \\ w_{cn_c}\begin{bmatrix} -\mathbf{n}_{cn_c}^\top & \mathbf{n}_{cn_c}^\top \left[\mathbf{R}^\top \mathbf{m}_{rn_c}\right]_\times \end{bmatrix} \end{bmatrix}$$

IV. Experiments

The proposed Balanced ICP pose estimation is evaluated qualitatively and quantitatively on both simulated, regardless of environmental or sensor artifacts, and real data from KITTI dataset [12]. For this purpose, our approach is compared with the result of the iterative optimization considering just the features extracted in the reference point cloud and matched in the current one and with the result of the iterative optimization considering just the features extracted in the current point cloud and matched in the reference one. In this section, we will refer to these methods as mono-directional.

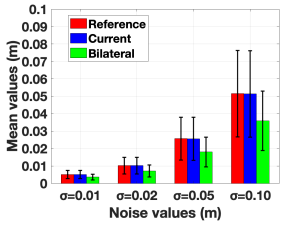
A. Simulated data

In the experiments with simulated data, 10000 experiments are performed, each one randomly generating a pair of point clouds. For each point cloud, 100 lines and 100 planes are extracted coherently with the points distribution and matched. Then, both the features and the points are perturbed with a Gaussian noise with mean zero and standard deviation σ. After the data generation, the tests are carried out with different values of σ and with different ratios of matching features reference-to-current/current-to-reference.

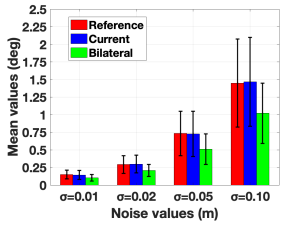


(a) ϵ_t with ratio 40/60

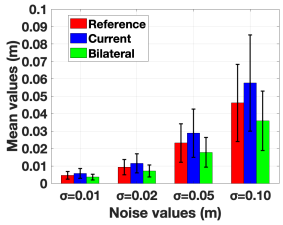
(b) ϵ_r with ratio 40/60



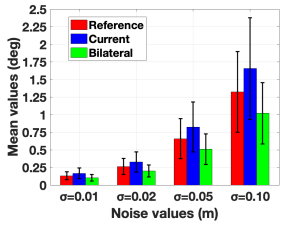
(c) ϵ_t with ratio 50/50



(d) ϵ_r with ratio 50/50

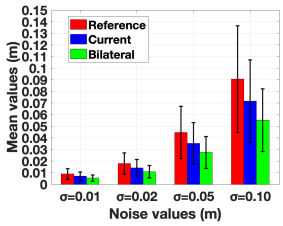


(e) ϵ_t with ratio 60/40

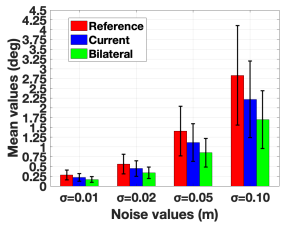


(f) ϵ_r with ratio 60/40

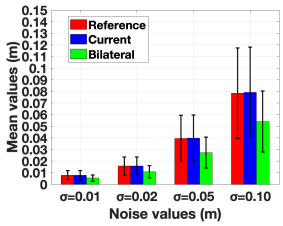
Fig. 2: Mean and std deviation of translation (a), (c), (e) and rotation (b), (d), (f) errors in point-to-line correspondences.



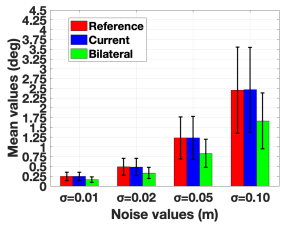
(a) ϵ_t with ratio 40/60



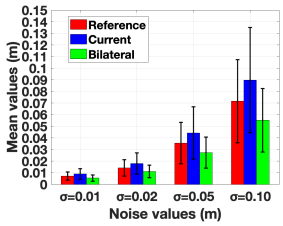
(b) ϵ_r with ratio 40/60



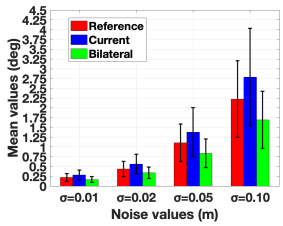
(c) ϵ_t with ratio 50/50



(d) ϵ_r with ratio 50/50



(e) ϵ_t with ratio 60/40



(f) ϵ_r with ratio 60/40

Fig. 3: Mean and std deviation of translation (a), (c), (e) and rotation (b), (d), (f) errors in point-to-plane correspondences.

In particular, the noise values considered are $\sigma \in \{0.01m; 0.02m; 0.05m; 0.10m\}$ and the ratios are $40/60$, $50/50$ and $60/40$ in percentage.

The results, for both for point-to-line and point-to-plane correspondences, are shown in Fig. 2 and Fig. 3. The plots, organized by distinguishing the matching ratio, illustrate the mean error ϵ_t and ϵ_r, respectively in translation and rotation, along with the standard deviation. Both sets of experiments demonstrate that the proposed algorithm is more precise than the mono-direction approaches, with a natural decrease of the precision with the increase of the noise. In particular, the algorithm brings a bigger improvement when the number of points is balanced, while obviously tends to the result of the best direction when the number of points is unbalanced, but still with higher accuracy. Additionally, the proposed balanced approach exhibits higher stability, consistently showing lower values of standard deviation.

B. Real data

These experiments were conducted on real data from the KITTI benchmark [12]. The dataset provides a set of 11 sequences in different driving scenarios recorded with a Velodyne HDL-64E laser scanner mounted on a car and the corresponding ground truth poses from GPS/IMU measurements. In order to evaluate the contribution of our method for a lidar odometry pipeline, we consider a simple scan-to-scan implementation, downsampling the point clouds with a voxel grid at 0.50 meters. Furthermore, we rely on point-to-plane correspondences since the dataset offers peri-urban areas where there is a lack of line features.

For each dataset scene, three methods are tested. The first method uses the classic point-to-plane ICP formulation extracting the planes in the reference scan. The second, uses the classic point-to-plane ICP formulation extracting the planes in the current scan. Finally, using the proposed method, the Bilateral ICP, extracting the planes both in the current and reference scan.

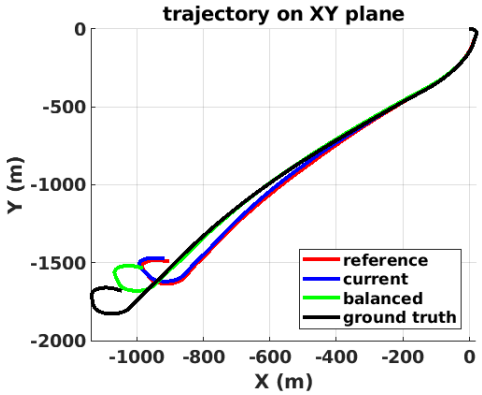


Fig. 4: Trajectory in the XY plane

A qualitative result is shown in Fig. 4, by plotting the estimated trajectory on the XY plane. The higher precision in the scan-to-scan registration provided by the Balanced ICP

produces a trajectory that is much closer to the ground truth than trajectories produced by the other approaches.

A quantitative evaluation of our method is shown, in tables I and II, in terms of mean μ and standard deviation σ of scan-to-scan motion estimation error in translation and rotation, respectively noted as δt and δr. The results are presented for all the 11 KITTI scenes with ground truth and lowest, thus the best, values are highlighted in bold text.

	Reference	Current	Balanced
Scene	$\mu(\delta t_r) \pm \sigma(\delta t_r)$	$\mu(\delta t_c) \pm \sigma(\delta t_c)$	$\mu(\delta t_b) \pm \sigma(\delta t_b)$
0	0.024 ± 0.022	0.024 ± 0.030	**0.023 ± 0.021**
1	0.299 ± 0.694	0.306 ± 0.724	**0.203 ± 0.562**
2	0.058 ± 0.179	0.066 ± 0.351	**0.047 ± 0.147**
3	0.032 ± 0.044	0.031 ± 0.022	**0.027 ± 0.017**
4	0.081 ± 0.257	0.053 ± 0.178	**0.036 ± 0.118**
5	0.016 ± 0.012	0.016 ± 0.012	**0.015 ± 0.010**
6	0.016 ± 0.013	0.016 ± 0.012	**0.015 ± 0.012**
7	0.018 ± 0.012	0.017 ± 0.012	**0.016 ± 0.011**
8	0.033 ± 0.039	0.033 ± 0.053	**0.030 ± 0.038**
9	0.034 ± 0.081	0.037 ± 0.101	**0.026 ± 0.019**
10	0.027 ± 0.061	0.026 ± **0.033**	**0.024** ± 0.041

TABLE I: Mean and standard deviation of translation error

	Reference	Current	Balanced
Scene	$\mu(\delta r_r) \pm \sigma(\delta r_r)$	$\mu(\delta r_c) \pm \sigma(\delta r_c)$	$\mu(\delta r_b) \pm \sigma(\delta r_b)$
0	0.082 ± 0.109	0.082 ± 0.113	**0.078 ± 0.107**
1	0.075 ± 0.143	0.076 ± 0.213	**0.058 ± 0.104**
2	0.082 ± 0.086	0.081 ± 0.091	**0.077 ± 0.080**
3	0.065 ± 0.045	0.065 ± 0.042	**0.062 ± 0.041**
4	0.038 ± 0.023	0.037 ± 0.023	**0.035 ± 0.022**
5	0.052 ± 0.044	0.051 ± 0.043	**0.049 ± 0.042**
6	0.037 ± 0.021	0.037 ± 0.021	**0.036 ± 0.020**
7	0.056 ± 0.051	0.054 ± 0.048	**0.050 ± 0.044**
8	0.068 ± 0.059	0.069 ± 0.060	**0.063 ± 0.055**
9	0.062 ± 0.049	0.064 ± 0.050	**0.057 ± 0.036**
10	0.079 ± 0.083	0.076 ± 0.067	**0.071 ± 0.065**

TABLE II: Mean and standard deviation of rotation error

It is noteworthy that upon close examination, the proposed method consistently attains lower mean errors and exhibits reduced standard deviations across all evaluated scenes. This commendable performance prompts a closer investigation into the efficacy of the method in alleviating potential errors arising from both the feature extraction and matching processes or some sensor artifact. Thus, the method demonstrates a notable capability in mitigating the impact of those additional sources of noise, which could potentially introduce cumulative errors in either direction during the pose estimation process.

Furthermore, observing the results of the first two approaches, we can conclude that a favorite point cloud to perform the features extraction cannot be identified, since is not always the same to reach a better result with respect to the other. This suggests that the proposed Balanced ICP is a better solution considering that in the applications there is no a priori knowledge available that could suggest considering the features in the reference scan and not in the current one, or vice-versa.

Finally, a timing study was conducted to evaluate the computational cost of the proposed method. The better accuracy of our method is paid in higher computational time. In fact, computing the average execution time per outer loop iteration over all the scenes, the Balanced ICP takes 2.34 times more than the monodirectional formulations of the ICP. This can be addressed by the fact that the number of matching considered is roughly doubled, but also by the fact that, having a more complex minimization function, the optimization process takes on average 2.69 iterations to converge against 1,82 of the monodirectional formulations.

V. Conclusions

This work presents a new approach for the ICP, named Balanced ICP, which extended the classic formulation leveraging the bidirectionality in the point cloud registration problem. Exploiting the feature extraction and matching those in both the point clouds, the Balanced ICP is able to balance the additional errors inevitably inherited. Simulated and experimental results have shown that the proposed method outperforms the classical monodirectional approaches for robustness, accuracy and stability. However, better performances are paid in higher computational complexity in our current Matlab implementation. For this reason, future works will focus on the matching selection, to reduce the number while still considering the features from both point clouds.

Acknowledgment

This work has been supported by EDSTIC d'Université Côte d'Azur. This work has been carried out within the context of the AISENSE project.

References

[1] P. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992.

[2] I. Vizzo, T. Guadagnino, B. Mersch, L. Wiesmann, J. Behley, and C. Stachniss, "Kiss-icp: In defense of point-to-point icp - simple, accurate, and robust registration if done the right way," *IEEE Robotics and Automation Letters*, 2023.

[3] J. Zhang and S. Singh, "Loam : Lidar odometry and mapping in real-time," *Robotics: Science and Systems Conference*, 2014.

[4] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018.

[5] Y. Pan, P. Xiao, Y. He, Z. Shao, and Z. Li, "Mulls: Versatile lidar slam via multi-metric linear least square," in *IEEE International Conference on Robotics and Automation*, 2021.

[6] X. Chen, A. Milioto, E. Palazzolo, P. Giguère, J. Behley, and C. Stachniss, "Suma++: Efficient lidar-based semantic slam," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2019.

[7] A. Milioto, I. Vizzo, J. Behley, and C. Stachniss, "Rangenet ++: Fast and accurate lidar semantic segmentation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2019.

[8] M. Yokozuka, K. Koide, S. Oishi, and A. Banno, "Litamin2: Ultra light lidar-based slam using geometric approximation applied with kl-divergence," in *IEEE International Conference on Robotics and Automation*, 2021.

[9] P. Dellenbach, J.-E. Deschaud, B. Jacquet, and F. Goulette, "Ct-icp: Real-time elastic lidar odometry with loop closure," in *International Conference on Robotics and Automation*, 2022.

[10] J. Zhu, D. Wang, X. Bai, H. Lu, C. Jin, and Z. Li, "Registration of point clouds based on the ratio of bidirectional distances," in *Fourth International Conference on 3D Vision*, 2016.

[11] S. Rusinkiewicz, "A symmetric objective function for ICP," *ACM Transactions on Graphics*, vol. 38, no. 4, 2019.

[12] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.