2010 IEEE International Conference on Robotics and Automation
Anchorage Convention District
May 3-8, 2010, Anchorage, Alaska, USA

A Generalization of the Metric-Based Iterative Closest Point Technique for 3D Scan Matching

Leopoldo Armesto Javier Minguez Luis Montesano

Abstract— **Scan matching techniques have been widely used to compute the displacement of robots. This estimate is part of many algorithms addressing navigation and mapping. This paper addresses the scan matching problem in three dimensional workspaces. We propose an generalization of the Metric based *Iterative Closest Point* (MbICP) to the 3D case. The main contribution is the development of all the mathematical tools required to formulate the ICP with this new metric, including the derivation of point to plane distances based on the new metric. We also provide experimental results to evaluate the algorithms and different combinations of ICP and MbICP to illustrate the advantages of the metric based approach.**

I. Introduction

Scan matching techniques are widely used to track the robot position using range data in many applications such as navigation and mapping. The principle is to compute the sensor displacement between two consecutive configurations by maximizing the overlap between the range measurements obtained at each configuration (see Figure 1). Although they are local in nature, in Robotics they have widely used as an improved odometry in navigation systems [1] or to solve the initial problem in SLAM [2], to ameliorate the loop-closing, etc. The most popular scan matching methods usually follow the Iterative Closest Point (ICP) algorithm (principle borrowed from the computer vision community [3]). The ICP algorithm addresses this problem with an iterative process in two steps. At each iteration:

1) **matching:** establishment of correspondent points between scans with a *closest point* criterion,
2) **minimization:** computation of the sensor displacement by a least square minimization of the error of the correspondences.

In two dimensions, a common feature of most versions of ICP is the usage of the Euclidean distance to establish the correspondences and to estimate the displacement [4], [5], [6]. The limitation of this distance is the difficulty to capture the sensor rotation [7]. To overcome this limitation, some new metric distances have been proposed to compensate translation and rotation simultaneously [8] improving the performance of previous methods. Based on these principles other techniques have been proposed to address the same issue but within probabilistic frameworks [9], [10]. Although outdoor 3D laser scanners are nowadays common in indoor and outdoor environments, only few 2D improvements have been studied in 3D. For instance, [11] uses one of the scans as a random generator for the other one and was extended to 3D in [12]. Another example is [13] that using the Hough transform aligns the scans and estimates the posterior distribution of relative poses. An approximated 3D version of the algorithm is able to align the scans, but further processing is required to improve accuracy [14].

Fig. 1. The scan matching problem: estimate the relative displacement between two configurations given two scans taken at these configurations, shown in black and red (gray in B&W), respectively.

In addition to this 3D generalization, many geometric ICP variants have been proposed in the Vision community to deal with the registration problem in 3D (see [15] for a survey). The different approaches modify the correspondences computation between meshes based on intersecting rays, projection of points to the mesh, orthogonal vectors, compatibility tests or on metrics using color information. This information is also used to re-weigh the correspondences in the minimization step. In [16], authors propose a variation to ICP by using a Z-buffer to find correspondences. The method does not necessarily obtain the closest point but accelerates the computation time using the GPU (Graphic Processor Unit). A closed-form estimate for the ICP covariance was proposed in [17]. Recently, Generalized-ICP [18] extended the point to plane distance proposed in [19]. By incorporating a probabilistic interpretation, it uses planar approximations to implement a plane-to-plane minimization and to take advantage of the structure of the environment. Another approach is to consider a global consistent matching between partially overlapping point sets, known as N-Registration in the Vision comunity (see [20] for a comparison) and N-Scan Matching problem in Robotics comunity [21].

In both communities, the idea of creating a geometric metric to improve the method performance, such as the one

L. Armesto is with Control and Systems Engineering, Universidad Politcnica de Valencia, Valencia, Spain leoaran@isa.upv.es

J. Minguez and L. Montesano are with the I3A, Computer Science Department, Universidad de Zaragoza jminguez@unizar.es, montesano@unizar.es

This work has been partially supported by the Vicerectorado de Investigacion, Desarrollo e Innovacion from Universidad Politcnica de Valencia, the Spanish project DPI2009-14732-C02-01 and the Portuguese FCT project PTDC/EEA-ACR/70174/2006.

proposed in [8], has not been explored in the 3D case. This paper describes the generalization of the metric-based ICP (MbICP) [8] to three dimensional workspaces. The emphasis of the work is on the development of all the mathematical formulation required to address the scan matching problem in three dimensional workspaces based on this new metric. This includes the derivation of the distance expression between points, segments, planes and facets; and the corresponding minimization expressions for pose estimation. Furthermore, we show extensive evaluations to compare the ICP, the MbICP and a natural combination of the MbICP and the ICP. Results show that the MbICP has the best performance specially in the presence of large errors in rotation.

II. Mathematical Tools

This section presents the generalization to 3D workspaces of the metric introduced in [8], to implement the two main steps of an ICP-like algorithm: 1) compute the correspondences (point to point and point to facet) between two scans and 2) estimate the displacement of the sensor.

A. Expression of the Metric distance

A rigid body transformation in $\mathbf{R}^3$ can be decomposed into a rotation of angle θ ($-\pi < \theta < \pi$) about a unit vector $n = (n_x, n_y, n_z)$ and a translation of vector (x, y, z) , and thus uniquely defined by vector $q = (x, y, z, \theta\ n_x, \theta\ n_y, \theta\ n_z)$. We define the norm of $\mathbf{q}$ as:

$$\|\mathbf{q}\| = \sqrt{\mathbf{t}^T\mathbf{t} + L^2\theta^2} = \sqrt{t_x^2 + t_y^2 + t_z^2 + L^2\theta^2} \tag{1}$$

where $L \in \mathbf{R}^+$. Let $d_p : \mathbf{R}^3 \times \mathbf{R}^3 \to [0, \infty)$ be the function defined as

$$d_p(\mathbf{p}_1, \mathbf{p}_2) = \min\{\|\mathbf{q}\| \mid \mathbf{p}_2 = \mathbf{R}(\mathbf{n}, \theta)\mathbf{p}_1 + \mathbf{t}\} \tag{2}$$

where $\mathbf{R}(\mathbf{n}, \theta)$ is a rotation matrix

$$\mathbf{R}(\mathbf{n}, \theta) = \mathbf{n}\mathbf{n}^T(1 - \cos\theta) + \mathbf{U}(\mathbf{n})\sin\theta + \mathbf{I}\cos\theta \tag{3}$$

and $\mathbf{U}(\mathbf{n})$ is a skew-symmetric matrix

$$\mathbf{U}(\mathbf{n}) = \begin{bmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{bmatrix} \tag{4}$$

It is easy to see that d_p as defined in equation (2) is a distance since it verifies:

1) $d_p(\mathbf{p}_1, \mathbf{p}_2) = d_p(\mathbf{p}_2, \mathbf{p}_2)$
2) $d_p(\mathbf{p}_1, \mathbf{p}_2) = 0 \Leftrightarrow \mathbf{p}_1 = \mathbf{p}_2$
3) $d_p(\mathbf{p}_1, \mathbf{p}_3) \leq d_p(\mathbf{p}_1, \mathbf{p}_2) + d_p(\mathbf{p}_2, \mathbf{p}_3)$

Notice that d_p is the distance between two points $(\mathbf{p}_1, \mathbf{p}_2)$, defined as the minimum norm among the rigid body transformations that move $\mathbf{p}_1$ to $\mathbf{p}_2$.

Unfortunately there is no closed form expression of d_p with respect to the coordinates of the points. However, assuming small rotations, equation (2) can be linearized about $\theta \to 0$. Then, $\cos\theta \approx 1$, $\sin\theta \approx \theta$ and $\mathbf{R}(\mathbf{n}, \theta) \approx \mathbf{I} + \mathbf{U}(\mathbf{n})\theta$. Consequently, the rigid transformation can be approximated as follows:

$$\mathbf{p}_2 \approx \mathbf{p}_1 + \mathbf{U}(\mathbf{n})\theta\mathbf{p}_1 + \mathbf{t} = \mathbf{p}_1 - \mathbf{U}(\mathbf{p}_1)\theta\mathbf{n} + \mathbf{t} \tag{5}$$

which can be expressed as

$$\mathbf{t} = \mathbf{U}(\mathbf{p}_1)\mathbf{r} + \boldsymbol{\delta} \tag{6}$$

where $\mathbf{r} = \theta\mathbf{n}$, $\boldsymbol{\delta} = \mathbf{p}_1 - \mathbf{p}_2$ and $\theta^2 = \mathbf{r}^T\mathbf{r} = \|\mathbf{r}\|_2^2$. Notice that equation (6) a linear system with respect to $\mathbf{r}$.

In this case the norm is given by

$$\begin{aligned} \|\mathbf{q}\|^2 &= \mathbf{t}^T\mathbf{t} + L^2\theta^2 = (\boldsymbol{\delta} + \mathbf{U}(\mathbf{p}_1)\mathbf{r})^T(\boldsymbol{\delta} + \mathbf{U}(\mathbf{p}_1)\mathbf{r}) + L^2\theta^2 \\ &= \mathbf{r}^T\left(\mathbf{U}^T(\mathbf{p}_1)\mathbf{U}(\mathbf{p}_1) + L^2\mathbf{I}\right)\mathbf{r} - 2\boldsymbol{\delta}^T\mathbf{U}^T(\mathbf{p}_1)\mathbf{r} + \boldsymbol{\delta}^T\boldsymbol{\delta} \\ &= \mathbf{r}^T\mathbf{A}\mathbf{r} - 2\mathbf{b}^T\mathbf{r} + c \end{aligned} \tag{7}$$

with $\mathbf{A} = \mathbf{U}^T(\mathbf{p}_1)\mathbf{U}(\mathbf{p}_1) + L^2\mathbf{I}$ and $\mathbf{b} = \mathbf{U}(\mathbf{p}_1)\boldsymbol{\delta}$.

According to equation (2), the approximated distance d_{ap} is given by the minimization of the norm (7) given the parameter $\mathbf{r}$. The result is $\mathbf{r}^* = \mathbf{A}^{-1}\mathbf{b}$, and thus the approximated distance is:

$$d_{ap}(\mathbf{p}_1, \mathbf{p}_2)^2 = \boldsymbol{\delta}^T\left[\mathbf{I} - \mathbf{U}^T(\mathbf{p}_1)\mathbf{A}^{-1}\mathbf{U}(\mathbf{p}_1)\right]\boldsymbol{\delta} = \boldsymbol{\delta}^T\mathbf{M}(\mathbf{p}_1)\boldsymbol{\delta} \tag{8}$$

with $\mathbf{M}(\mathbf{p}_1) = \mathbf{I} - \frac{\mathbf{U}^T(\mathbf{p}_1)\mathbf{U}(\mathbf{p}_1)}{k}$ and $k = \|\mathbf{p}_1\|_2^2 + L^2$. Developing (8) the approximated distance is

$$d_{ap}(\mathbf{p}_1, \mathbf{p}_2)^2 = \sqrt{\boldsymbol{\delta}^T\boldsymbol{\delta} - \frac{\boldsymbol{\delta}\mathbf{U}^T(\mathbf{p}_1)\mathbf{U}(\mathbf{p}_1)\boldsymbol{\delta}}{k}} = \tag{9}$$

$$= \sqrt{\|\boldsymbol{\delta}\|_2^2 - \frac{\|\mathbf{p}_1 \times \boldsymbol{\delta}\|_2^2}{k}}. \tag{10}$$

The iso-distance surfaces for such as new distance are computed by

$$\{\mathbf{p}_2 \in \mathbf{R}^3 \mid d_{ap}(\mathbf{p}_1, \mathbf{p}_2) = c\}. \tag{11}$$

These curves are ellipsoids centered in $\mathbf{p}_1$ with principal axis given by the vector $\mathbf{v}_1 = c\frac{\mathbf{p}_1}{\|\mathbf{p}_1\|_2}$ and any other two vectors orthogonal one to each other of magnitude $c(1 + \frac{\|\mathbf{p}_1\|_2^2}{L^2})^{\frac{1}{2}}$. Thus, L balances the trade-off between translation and rotation. When $L \to \infty$, the new distance tends to the Euclidean distance, otherwise, the parameter L acts as weighting factor between translation and rotation.

B. Expressions for the Correspondence Step

In the previous section, we derived the point to point distance based on the metric of expression 1. In practice, range data is discrete and, therefore, points in each scan are different. The standard way to solve this problem is to assume a continuous local structure between points in the reference scan. In 2D, this implies building segments in between points of the reference scan [7], while in 3D facets or planes are the most common structures [19].

In order to obtain the expression of the distance from a point to a triangular facet (geometric triangle among three 3D points, the distances from a point to segment and from a point to plane and facet are described first.

1) Point to Segment Distance: Let $\mathbf{s}$ be a segment defined by two points $\mathbf{s}_1, \mathbf{s}_2 \in \mathbf{R}^3$. The goal is to compute the point $\mathbf{p}^* \in [\mathbf{s}_1\ \mathbf{s}_2]$ with the minimum distance to $\mathbf{p}_1$:

$$d_{ps}(\mathbf{p}_1, \mathbf{p}^*) = \min_{\lambda \in [0,1]} d_{ap}(\mathbf{p}_1, (1 - \lambda)\mathbf{s}_1 + \lambda\mathbf{s}_2). \tag{12}$$

Denoting $d(\lambda) = d_{ap}(\mathbf{p}_1, (1 - \lambda)\mathbf{s}_1 + \lambda\mathbf{s}_2)$, the approximated distance between $\mathbf{p}_1$ and the line that contains the segment $\mathbf{s}$ is

$$d^2(\lambda) = \|\boldsymbol{\delta}(\lambda)\|_2^2 - \frac{\mathbf{p}_1 \times \boldsymbol{\delta}(\lambda)}{k} = a\lambda^2 - 2b\lambda + c \tag{13}$$

with $a = \mathbf{u}_2^T\mathbf{M}(\mathbf{p}_1)\mathbf{u}_2 = \|\mathbf{u}_2\|_2^2 - \frac{\|\mathbf{p}_1 \times \mathbf{u}_2\|_2^2}{k}$, $b = -\boldsymbol{\delta}_1\mathbf{M}(\mathbf{p}_1)\mathbf{u}_2 = -\boldsymbol{\delta}_1^T\mathbf{u}_2 + \frac{(\mathbf{p}_1 \times \boldsymbol{\delta}_1)^T(\mathbf{p}_1 \times \mathbf{u}_2)}{k}$, $k = \|\mathbf{p}_1\|_2^2 + L^2$

and $c = \boldsymbol{\delta}_1^T \mathbf{M}(\mathbf{p}_1)\boldsymbol{\delta}_1 = \|\boldsymbol{\delta}_1\|_2^2 - \frac{\|\mathbf{p}_1 \times \boldsymbol{\delta}_1\|_2^2}{k}$ (where $\mathbf{u}_2 = \mathbf{s}_2 - \mathbf{s}_1$ and $\boldsymbol{\delta}_1 = \mathbf{s}_1 - \mathbf{p}_1$).

The minimum of equation (13) is $\lambda^* = \frac{b}{a}$. Therefore, the distance of point $\mathbf{p}_1$ to segment $\mathbf{s}$ is given by:

$$d_{ps}(\mathbf{p}_1, [\mathbf{s}_1\ \mathbf{s}_2]) = \begin{cases} d_{ap}(\mathbf{p}_1, \mathbf{s}_1) = \sqrt{c} & \text{if } \lambda^* < 0, \\ d_{ap}(\mathbf{p}_2, \mathbf{s}_2) = \sqrt{a - 2b + c} & \text{if } \lambda^* > 1, \\ d(\lambda^*) = \sqrt{-\lambda^* b + c} & \text{if } 0 \leq \lambda^* \leq 1. \end{cases} \tag{14}$$

2) Point to Plane Distance: Let $[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]$ be the plane defined by three points $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbf{R}^3$. The minimum distance to a point $\mathbf{p}_1$ is:

$$d_{pp}(\mathbf{p}_1, [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]) = \min_{\lambda_1, \lambda_2 \in [0,1]} d(\boldsymbol{\lambda}) \tag{15}$$

where $d(\boldsymbol{\lambda}) = d_{ap}(\mathbf{p}_1, \mathbf{v}_1 + \lambda_1(\mathbf{v}_3 - \mathbf{v}_1) + \lambda_2(\mathbf{v}_2 - \mathbf{v}_1))$. The approximated distance between $\mathbf{p}_1$ and the plane that contains the facet $\mathbf{v}$ is:

$$d^2(\boldsymbol{\lambda}) = \|\boldsymbol{\delta}(\boldsymbol{\lambda})\|_2^2 - \frac{\|\mathbf{p}_1 \times \boldsymbol{\delta}(\boldsymbol{\lambda})\|_2^2}{k} \tag{16}$$

where $\boldsymbol{\lambda} = [\lambda_1\ \lambda_2]^T$, $\boldsymbol{\delta}(\boldsymbol{\lambda}) = \boldsymbol{\delta}_1 + \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2$, $\mathbf{u}_1 = \mathbf{v}_3 - \mathbf{v}_1$, $\mathbf{u}_2 = \mathbf{v}_2 - \mathbf{v}_1$ and $\boldsymbol{\delta}_1 = \mathbf{v}_1 - \mathbf{p}_1$. Equation (16) can be written

$$d^2(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^T \mathbf{A}\boldsymbol{\lambda} - 2\mathbf{B}^T\boldsymbol{\lambda} + C \tag{17}$$

where $\mathbf{A} = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} d \\ e \end{bmatrix}$ and $C = f$, and with:

$$a = \mathbf{u}_1^T \mathbf{M}(\mathbf{p}_1)\mathbf{u}_1 = \|\mathbf{u}_1\|_2^2 - \frac{\|\mathbf{p}_1 \times \mathbf{u}_1\|_2^2}{k}, \tag{18}$$

$$b = \mathbf{u}_2^T \mathbf{M}(\mathbf{p}_1)\mathbf{u}_2 = \|\mathbf{u}_2\|_2^2 - \frac{\|\mathbf{p}_1 \times \mathbf{u}_2\|_2^2}{k}, \tag{19}$$

$$c = -\boldsymbol{\delta}_1^T \mathbf{M}(\mathbf{p}_1)\mathbf{u}_1 = -\boldsymbol{\delta}_1^T \mathbf{u}_1 + \frac{(\mathbf{p}_1 \times \boldsymbol{\delta}_1)^T(\mathbf{p}_1 \times \mathbf{u}_1)}{k}, \tag{20}$$

$$d = -\boldsymbol{\delta}_1^T \mathbf{M}(\mathbf{p}_1)\mathbf{u}_2 = -\boldsymbol{\delta}_1^T \mathbf{u}_2 + \frac{(\mathbf{p}_1 \times \boldsymbol{\delta}_1)^T(\mathbf{p}_1 \times \mathbf{u}_2)}{k}, \tag{21}$$

$$f = \boldsymbol{\delta}_1^T \mathbf{M}(\mathbf{p}_1)\boldsymbol{\delta}_1 = \|\boldsymbol{\delta}_1\|_2^2 - \frac{\|\mathbf{p}_1 \times \boldsymbol{\delta}_1\|_2^2}{k}. \tag{22}$$

where, again, $k = \|\mathbf{p}_1\|_2^2 + L^2$.

The minimum of equation (17) is $\boldsymbol{\lambda}^* = \mathbf{A}^{-1}\mathbf{B}$ and the closest point is $\mathbf{p_0} = \mathbf{p}(\boldsymbol{\lambda}^*) = \mathbf{v}_1 + \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2$.

3) Point to Facet Distance: Let $\mathbf{v}$ be a facet (triangle) defined by the three previous points $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbf{R}^3$. The minimum distance to a point $\mathbf{p}_1$ is a piecewise function that depends on the position of the point $\mathbf{p}(\boldsymbol{\lambda}^*)$ with respect to the facet given by (Figure 2):

$$d_{pf}(\mathbf{p}_1, \mathbf{v}) = \begin{cases} d_{ap}(\mathbf{p}_1, [\mathbf{v}_1\ \mathbf{v}_2]) & \text{if } \lambda_1^* < 0 \text{ and } \begin{cases} \lambda_2^* > 0 \text{ and } \lambda_1^* + \lambda_2^* < 1, \\ \text{or } \lambda_2^* < 0 \text{ and } d < 0, \\ \text{or } \lambda_1^* + \lambda_2^* > 1 \text{ and } e < b \end{cases} \\ d_{ap}(\mathbf{p}_1, [\mathbf{v}_1\ \mathbf{v}_3]) & \text{if } \lambda_2^* < 0 \text{ and } \begin{cases} \lambda_1^* > 0 \text{ and } \lambda_1^* + \lambda_2^* < 1, \\ \text{or } \lambda_1^* < 0 \text{ and } d \geq 0, \\ \text{or } \lambda_1^* + \lambda_2^* > 1 \text{ and } d < a \end{cases} \\ d_{ap}(\mathbf{p}_1, [\mathbf{v}_2\ \mathbf{v}_3]) & \text{if } \lambda_1^* + \lambda_2^* > 1 \text{ and } \begin{cases} \lambda_1^* > 0 \text{ and } \lambda_2^* > 0, \\ \text{or } \lambda_1^* < 0 \text{ and } e \geq b, \\ \text{or } \lambda_2^* < 0 \text{ and } d \geq a \end{cases} \\ d_{ap}(\mathbf{p}_1, \mathbf{p}(\boldsymbol{\lambda}^*)) & \text{otherwise} \end{cases} \tag{23}$$

The distances of equation (23) are

$$d_{ap}(\mathbf{p}_1, [\mathbf{v}_1\ \mathbf{v}_2]) = \begin{cases} \sqrt{f} & \text{if } \frac{e}{b} < 0 \\ \sqrt{b - 2e + f} & \text{if } \frac{e}{b} > 1 \\ \sqrt{-\frac{e^2}{b} + f} & \text{otherwise} \end{cases} \tag{24}$$

$$d_{ap}(\mathbf{p}_1, [\mathbf{v}_1\ \mathbf{v}_3]) = \begin{cases} \sqrt{f} & \text{if } \frac{d}{a} < 0 \\ \sqrt{a - 2d + f} & \text{if } \frac{d}{a} > 1 \\ \sqrt{-\frac{d^2}{a} + f} & \text{otherwise} \end{cases} \tag{25}$$

$$d_{ap}(\mathbf{p}_1, [\mathbf{v}_2\ \mathbf{v}_3]) = \begin{cases} \sqrt{i} & \text{if } \frac{h}{g} < 0 \\ \sqrt{g - 2h + i} & \text{if } \frac{h}{g} > 1 \\ \sqrt{-\frac{h^2}{g} + i} & \text{otherwise} \end{cases} \tag{26}$$

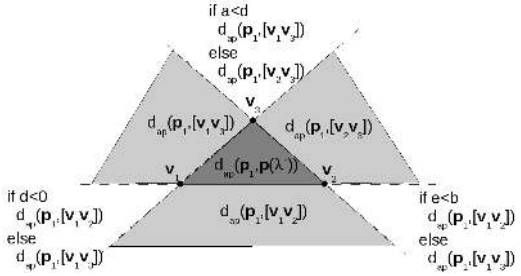
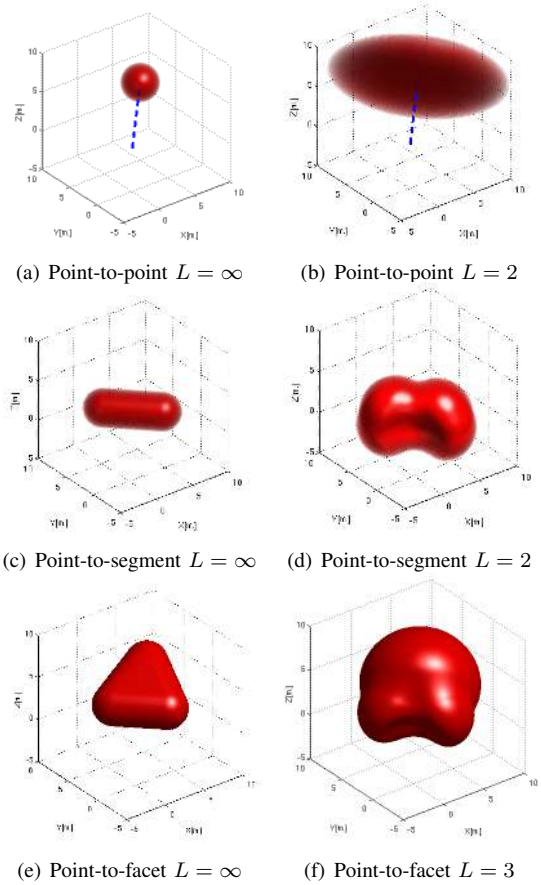


Fig. 2. This figure depicts the different zones where the closest point to p_1 in the plane defined by the segment or facet could lie.



(a) Point-to-point $L = \infty$ (b) Point-to-point $L = 2$

(c) Point-to-segment $L = \infty$ (d) Point-to-segment $L = 2$

(e) Point-to-facet $L = \infty$ (f) Point-to-facet $L = 3$

Fig. 3. Isodistance curves of the the new metric, including the particular case of the euclidean distance ($L = \infty$).

with $g = \mathbf{u}_3^T \mathbf{M}(\mathbf{p}_1)\mathbf{u}_3 = \|\mathbf{u}_3\|_2^2 - \frac{\|\mathbf{p}_1 \times \mathbf{u}_3\|_2^2}{k}$, $h = -\boldsymbol{\delta}_2 \mathbf{M}(\mathbf{p}_1)\mathbf{u}_3 = -\boldsymbol{\delta}_2^T \mathbf{u}_3 + \frac{(\mathbf{p}_1 \times \boldsymbol{\delta}_2)^T(\mathbf{p}_1 \times \mathbf{u}_3)}{k}$, $i = \boldsymbol{\delta}_2^T \mathbf{M}(\mathbf{p}_1)\boldsymbol{\delta}_2 = \|\boldsymbol{\delta}_2\|_2^2 - \frac{\|\mathbf{p}_2 \times \boldsymbol{\delta}_2\|_2^2}{k}$ and $\mathbf{u}_3 = \mathbf{v}_3 - \mathbf{v}_2$, $\boldsymbol{\delta}_2 = \mathbf{v}_2 - \mathbf{p}_1$.

Figures 3(a) to 3(f) illustrate the effect of the L parameter. Notice that the new metric captures the rotation by deforming the Euclidean distance to a segment or to a facet with a combination of the ellipsoids of their corresponding vertices.

C. Expression for the minimization

Based on the distance of Eq. 23, ICP algorithms establish a set of correspondences between two scans. The minimization step uses these correspondences to estimate the relative displacement between the two scans. This is done by minimizing the sum of squared distances of the correspondences i.e., using a least squares estimator. To derive the estimator, we use the following notation: $\mathbf{p}_i$ is a point in the new scan while $\mathbf{p}'_i$ is its corresponding point in the reference scan

(model). The goal is to compute the rotation and translation that minimizes the performance index

$$E_{dist}(\mathbf{q}) = \sum_{i=1}^{n} d_{pp}^2(\mathbf{p}_i, \mathbf{p}'_i) \tag{27}$$

with $\mathbf{p}'_i = \mathbf{R}(\mathbf{n}, \theta)\mathbf{p}_i + \mathbf{t}$. This expression does not have a closed form solution. By linearizing about $\theta = 0$, we obtain:

$$E_{dist}(\mathbf{q}) = \sum_{i=1}^{n} \boldsymbol{\delta}_i^T(\mathbf{q})\mathbf{M}(\mathbf{p}_i)\boldsymbol{\delta}_i(\mathbf{q}) \tag{28}$$

being $\boldsymbol{\delta}_i(\mathbf{q}) = \boldsymbol{\delta}_i + \mathbf{U}(\mathbf{p}_i)\mathbf{r} - \mathbf{t}$ and $\boldsymbol{\delta}_i = \mathbf{p}'_i - \mathbf{p}_i$. Expanding equation (28),

$$E_{dist}(\mathbf{q}) = \mathbf{q}^T\mathbf{A}\mathbf{q} - 2\mathbf{B}^T\mathbf{q} + C \tag{29}$$

$$\mathbf{A} = \sum_{i=1}^{n} \begin{bmatrix} \mathbf{M}(\mathbf{p}_i) & -\mathbf{M}(\mathbf{p}_i)\mathbf{U}(\mathbf{p}_i) \\ -\mathbf{U}^T(\mathbf{p}_i)\mathbf{M}(\mathbf{p}_i) & \mathbf{U}(\mathbf{p}_i)\mathbf{M}(\mathbf{p}_i)\mathbf{U}(\mathbf{p}_i) \end{bmatrix}, \tag{30}$$

$$\mathbf{B} = \sum_{i=1}^{n} \begin{bmatrix} \mathbf{M}(\mathbf{p}_i) \\ \mathbf{M}(\mathbf{p}_i)\mathbf{U}(\mathbf{p}_i) \end{bmatrix} \boldsymbol{\delta}_i, \; C = \sum_{i=1}^{n} \boldsymbol{\delta}_i^T\mathbf{M}(\mathbf{p}_i)\boldsymbol{\delta}_i \tag{31}$$

The minimum of equation (29) is $\mathbf{q}^* = \mathbf{A}^{-1}\mathbf{B}$, which is the displacement that minimizes the correspondence error.

III. Experimental results

This section analyses the performance of the new 3D metric. We considered three different algorithms that differ in the distances used to estimate the correspondences and to define the minimization criterion:

- **ICP**: Standard ICP based on the Euclidean metric,
- **MbICP**: Uses the new metric to compute the correspondences and to estimate the relative displacement,
- **MbICP-mixed**: Uses the new metric to establish the correspondences and the Euclidean one to estimate the relative displacement.

The MbICP-mixed algorithm is motivated by the fact that the metric-based minimization step does not have a closed form and solves a linearized version of the problem. Thus, a closed form solution may provide better solutions. In particular, for known correspondences the Euclidean minimization ICP behaves better than the metric based version, but this difference disappears in the presence of noise and wrong correspondences.

The analysis is done using a dataset of 72 3D laser scans [22] recorded in a bakery with a 2D laser scan mounted on a robotic platform. In order to obtain 3D scans, the 2D laser was mounted on a servo drive. Since the robot moved during data acquisition, the robot motion was compensated to create 3D scans prior to the scan matching process. The implementation of the three algorithms is essentially the same, but for the distance expression used to compute the correspondences (Eq. 23) and the estimation of the displacement (based on Eq. 29). In order to create the facets, we assume that scans form quads typically obtained by the pan and tilt sweeping of a laser ray. Thus, we do not perform any Delaunay triangulation to obtain the local structure, but simply form two facets per each quad (four consecutive points in both directions). All the algorithms use a sliding window search which restricts the correspondence search to points or facets that lay within a given angular window. We filter out boundary points by implementing a simple adaptive break-point detector similar to the one proposed in [23] so we can detect gaps between points and reject points which are too far or too close (to avoid considering points that correspond to the robot used to collect the measurements). We used the trimmed version [24] to filter spurious correspondences when obtaining the maps. For the stopping criteria, we followed the criterion proposed in [4] and considered the error ratio during several time steps.

A. Robustness, Precision and Convergence Analysis

The robustness and precision analysis follows the same methodology as the one presented in [8]. Each scan was matched against itself, but with random initial locations. We defined 8 different levels of error, starting form $\pm 0.025m$ in translation and 7.5 degrees in rotation (Experiment 1) up to $0.2m$ and 60 degrees (Experiment 8), with increments of $0.025m$ and 7.5 degrees on each experiment. Based on these values, we run a 50 samples of Monte Carlo simulation for each scan and each noise level of noise making a total of 28800 runs. Although we also carried out the same tests using point to point distances, we only report results for the point to facet case which provided better results. Since we match a scan against itself, the matching of both scans should in principle be perfect at convergence and we can calculate the error made by each method against the ground truth (zero rotation and zero translation).

We analyze first the robustness of the methods. For this analysis, we considered a successful run those solutions that converged in less than 150 iterations, whose error in translation was lower than 25 mm and whose error in rotation was lower than 0.25 degrees. Based on this thresholds we compute the percentage of true positives (the solution converged below the threshold); false positives (the algorithm converged with errors above the threshold); true negatives (the algorithm did not converged and was still above threshold); and false negatives (the algorithm did not converge but the solution was already below threshold).

Table I shows the results for each method. When the error is small all the methods achieved 100% of true positives. As the error increases from experiment #1 to #8, the percentage of true positives decreases for all the methods. The MbICP performed better than the other methods since it has the highest percentage of true positives and the lowest number of false positives.

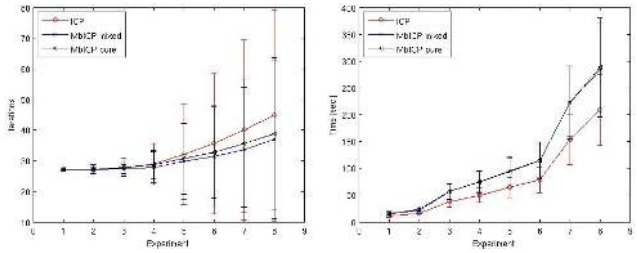
We now study the precision achieved by each algorithm. For this analysis, we considered only the true positives and define the error (25 mm and 0.25 degrees) in three equal intervals corresponding to high (H), medium (M) and low (L) precision. Table II shows the results for each method. The performance of all methods is quite similar, although the MbICP obtained slightly better results.

Finally, we consider the convergence rate and the execution time of each method. Figure 4(a) depicts the mean number of iterations used by each method for each experiment. There are no significant differences between the methods

TABLE I

ROBUSTNESS ANALYSIS OF SCAN MATCHERS.

Exp.	Method	ICP	MbICP mixed	MbICP pure
#1	TP	100	100	100
	FP	0	0	0
	TN	0	0	0
	FN	0	0	0
#2	TP	100	100	100
	FP	0	0	0
	TN	0	0	0
	FN	0	0	0
#3	TP	100	100	100
	FP	0	0	0
	TN	0	0	0
	FN	0	0	0
#4	TP	99.89	99.89	100
	FP	0	0	0
	TN	0.1111	0.1111	0
	FN	0	0	0
#5	TP	98.58	99.19	99.36
	FP	0.3056	0.25	0.1111
	TN	1.111	0.3889	0.4722
	FN	0	0.1667	0.05556
#6	TP	96.67	98.5	99
	FP	0.6944	0.1944	0.1944
	TN	2.583	1.306	0.8056
	FN	0.05556	0	0
#7	TP	94.11	97.28	97.67
	FP	1.528	0.7222	0.2778
	TN	4.333	1.944	1.889
	FN	0.02778	0.05556	0.1667
#8	TP	90.56	95.25	96.58
	FP	3.444	1.556	0.9167
	TN	5.944	3.167	2.361
	FN	0.05556	0.02778	0.1389



(a) Iterations (b) Computational Time

Fig. 4. Number of iterations and computational time required by the ICP, MbICP and the MbICP mixed for the different levels of error.

(smaller than five iterations in average). The MbICP-mixed required the least number of iterations followed by the MbICP and the ICP. Regarding execution times, MbICP iterations are more expensive than ICP. The ICP is the fastest algorithm (see Figure 4(b)). The figure shows how the time depends mainly on the computation of the correspondences, which is quadratic in the number of points. Since the angular window is bigger for bigger errors, the computational time increased in a quadratic way. The constant overhead required by the MbICP algorithm is the price to pay for the improvement in robustness and precision.

B. Maps

Figures 5(a) to 5(f) show robot trajectories estimation and super-imposed laser scans at each position for a whole experiment (one loop in an old bakery), where in this case the matching has been performed between consecutive scans. In the experiment, the odometry data is quite accurate and it is not hard to obtain a good estimation since the environment

TABLE II

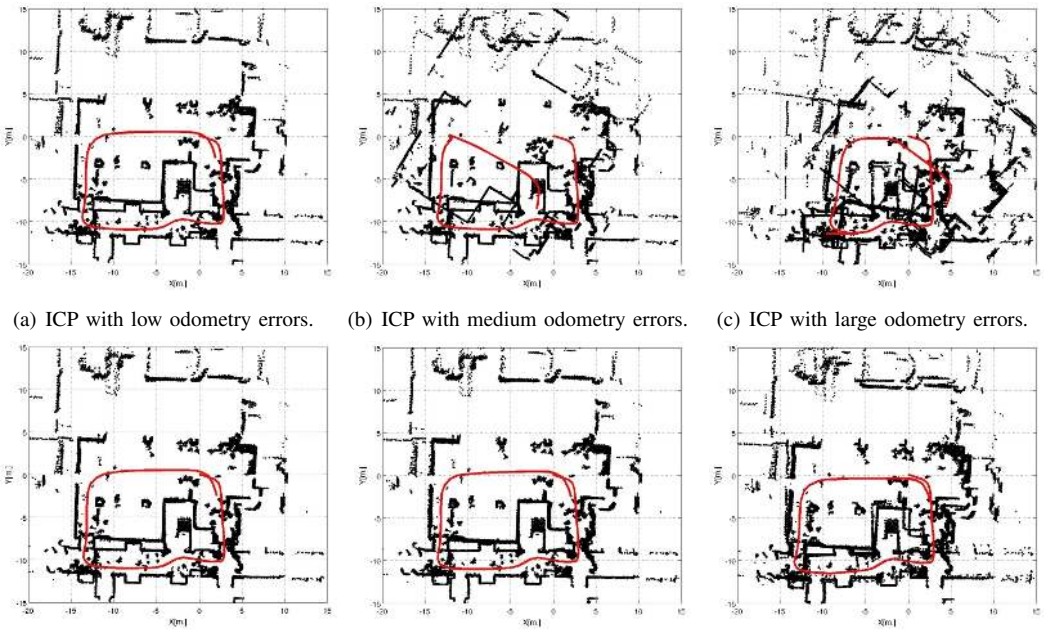
PRECISION ANALYSIS OF SCAN MATCHERS.

Exp.	Method	ICP	MbICP mixed	MbICP pure
#1	H	99.97	100	100
	M	0.02778	0	0
	L	0	0	0
#2	H	99.94	99.97	100
	M	0.05556	0.02778	0
	L	0	0	0
#3	H	99.75	99.89	100
	M	0.25	0.1111	0
	L	0	0	0
#4	H	99.67	99.81	100
	M	0.3337	0.1947	0
	L	0	0	0
#5	H	99.32	99.75	99.94
	M	0.6762	0.252	0.05591
	L	0	0	0
#6	H	99.34	99.83	99.92
	M	0.6609	0.1692	0.08418
	L	0	0	0
#7	H	98.88	99.8	99.94
	M	1.122	0.1999	0.05688
	L	0	0	0
#8	H	98.87	99.45	99.8
	M	1.135	0.5541	0.2013
	L	0	0	0

is well structured and no relevant differences between ICP and MbICP algorithms can be shown. In order to evaluate the performance of ICP and MbICP (pure) under several wheel slippage and terrain conditions or complex scenarios, we have introduce artificial random errors (low, medium and large) to odometry data with magnitudes as in experiments #2, #5 and #7 of Tables I and II, respectively. In this case, it is clearly shown that the MbICP provides good quality maps and estimations even with large errors in odometry data. The ICP degenerates even at medium errors, although, it has to be remarked that, it has just one single failure in rotation (see figure 5(b)). As a consequence of this failure, the robot is not able to properly close the loop. For larger odometry errors ICP performance is clearly degraded and several failures have been found, in translation and rotation. We have tested both algorithms with several parameter configurations, and selected the best ever found in any case. The estimation of the MbICP, through all experiments (#1 to #8), has shown to be more robust and accurate than ICP, specially when rotation errors are large.

IV. CONCLUSIONS

This paper describes the extension of 2D metric-based scan matching (MbICP) to 3D. The focus is on the derivation of the distance and minimization expressions for the 3D case and, in particular, to consider point to plane and point to facet correspondences. Experimental results show that most of the advantages of MbICP are also present in 3D, since it is more robust and more precise than the standard ICP algorithm. The new metric also requires less iterations to converge. However, contrary to the 2D case, the increased cost per iteration is not compensated by the smaller number of iterations and results in slightly higher computational times. Future work will focus on integrating metric based algorithms in a 3D probabilistic framework to take advantage of some of the most recent developments in 3D scan matching algorithms.



(a) ICP with low odometry errors. (b) ICP with medium odometry errors. (c) ICP with large odometry errors.

(d) MbICP with low odometry errors. (e) MbICP with medium odometry errors. (f) MbICP with large odometry errors.

Fig. 5. 2D map views (XY plane) and elevation map of an old bakery obtained with different scan matching algorithms and odometry errors.

ACKNOWLEDGEMENTS

Authors want to thank Pedro Piniés for the 3D dataset.

REFERENCES

[1] L. Montesano, J. Minguez, and L. Montano, "Modeling dynamic scenarios for local sensor-based motion planning," *Autonomous Robots*, vol. 25, no. 3, pp. 223–251, 2008.

[2] A. Nchter, K. Lingemann, J. Hertzberg, and H. Surmann, "Heuristic-based laser scan matching for outdoor 6d slam," in *In Advances in artificial intelligence. 28th annual German Conf. on AI*, 2005, pp. 304–319.

[3] P. Besl and N. McKay, "A method for registration of 3d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, no. 14, pp. 239–256, 1992.

[4] S. Pfister, K. Kreichbaum, S. Roumeliotis, and J. Burdick, "Weighted range sensor matching algorithms for mobile robot displacement estimation," in *IEEE International Conference on Robotics and Automation*, 2002, pp. 1667–74.

[5] J.-S. Gutmann, "Amos: comparison of scan matching approaches for self-localization in indoor environments," in *Proceedings of the First Euromicro Workshop on Advanced Mobile Robots (EUROBOT 96) EURBOT-96*, 1996, p. 61.

[6] O. Bengtsson, "Localization in changing environments - estimation of a covariance matrix for the idc algorithm," in *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems Expanding the Societal Role of Robotics in the the Next Millennium (Cat No 01CH37180) IROS-01*, vol. 4, 2001, p. 1931.

[7] F. Lu and E. Milios, "Robot pose estimation in unknown environments by matching 2d range scans," *Intelligent Robots Systems*, no. 20, pp. 249–275, 1997.

[8] J. Minguez, L. Montesano, and F. Lamiraux, "Metric-based iterative closest point scan matching for sensor displacement estimation," *IEEE Trans. on Robotics*, vol. 22, no. 5, pp. 1047–1054, 2005.

[9] L. Montesano, J. Minguez, and L. Montano, "Probabilistic scan matching for motion estimation in unstructured environments," in *Conference on Intelligent Robots and Systems (IROS)*, 2005, pp. 3499–3504.

[10] L. Montesano, "Detection and tracking of moving objects from a mobile platform. application to navigation and multi-robot localization." Ph.D. dissertation, Universidad de Zaragoza, Spain, 2006.

[11] P. Biber and W. Strafler, "The normal distributions transform: A new approach to laser scan matching," in *IEEE Int. Conf. on Intelligent Robots and Systems*, Las Vegas, USA, 2003.

[12] E. Takeuchi, "A 3-d scan matching using improved 3-d normal distributions transform for mobile robotic mapping," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

[13] A. Censi, "Scan matching in a probabilistic framework," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Orlando, Florida, 2006, pp. 2291–2296. [Online]. Available: http://purl.org/censi/2006/gpm

[14] A. Censi and S. Carpin, "Global 6dof scan-matching in the hough domain," in *IEEE Int. Conf. Robotics and Automation*, 2009.

[15] S. Rusinkiewicz and M. Levoy, "Efficient variants of the icp algorithm," *3D Digital Imaging and Modeling, International Conference on*, vol. 0, p. 145, 2001.

[16] R. Benjemaa and F. Schmitt, "Fast global registration of 3d sampled surfaces using a multi-z-buffer technique," in *Image and Vision Computing*, 1997, pp. 113–120.

[17] A. Censi, "An accurate closed-form estimate of icp's covariance," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, 2007, p. 3167.

[18] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp," in *Proceedings of Robotics: Science and Systems*, Seattle, USA, June 2009.

[19] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," in *Proceedings 1991 IEEE International Conference on Robotics and Automation*, 1991, p. 2724.

[20] S. J. Cunnington and A. J. Stoddart, "N-view point set registration: A comparison," in *British Machine Vision Conference*, 1999.

[21] P. Biber, "nscan-matching: simultaneous matching of multiple scans and application to slam," in *In Robotics and Automation. ICRA Proceedings IEEE International Conference on*, 2006, pp. 2270–2276.

[22] O. Wulf, K. O. Arras, H. I. Christensen, and B. Wagner, "2d mapping of cluttered indoor environments by means of 3d perception," in *Proc. IEEE International Conference on Robotics and Automation (ICRA'04)*, New Orleans, USA, 2004.

[23] G. Araujo and M. Aldon, "Line extraction in 2d range images for mobile robotics," *Journal of Intelligent and Robotic Systems*, no. 40, pp. 267–297, 2004.

[24] D. Chetverikov, "The trimmed iterative closest point algorithm," in *Object recognition supported by user interaction for service robots ICPR-02*, vol. 3, 2002.