Towards Efficient Indoor/Outdoor Registration Using Planar Polygons

Rahima Djahel, Pascal Monasse, Bruno Vallet

TOWARDS EFFICIENT INDOOR/OUTDOOR REGISTRATION USING PLANAR POLYGONS

Rahima Djahel[1], Bruno Vallet[2], Pascal Monasse[1]

[1] LIGM, École des Ponts, Univ Gustave Eiffel, ESIEE Paris, CNRS, Marne-la-vallée, France -firstname.lastname@enpc.fr
[2] LASTIG, Univ Gustave Eiffel, ENSG, IGN, F-94160 Saint-Mandé, France - firstname.lastname@ign.fr

KEY WORDS: Planar polygons, RANSAC, MSAC, Contrario, Sensor topology, Clustering, Registration

ABSTRACT:

The registration of indoor and outdoor scans with a precision reaching the level of geometric noise represents a major challenge for Indoor/Outdoor building modeling. The basic idea of the contribution presented in this paper consists in extracting planar polygons from indoor and outdoor LiDAR scans, and then matching them. In order to cope with the very small overlap between indoor and outdoor scans of the same building, we propose to start by extracting points lying in the buildings' interior from the outdoor scans as points where the laser ray crosses detected façades. Since, within a building environment, most of the objects are bounded by a planar surface, we propose a new registration algorithm that matches planar polygons by clustering polygons according to their normal direction, then by their offset in the normal direction. We use this clustering to find possible polygon correspondences (hypotheses) and estimate the optimal transformation for each hypothesis. Finally, a quality criteria is computed for each hypothesis in order to select the best one. To demonstrate the accuracy of our algorithm, we tested it on real data with a static indoor acquisition and a dynamic (Mobile Laser Scanning) outdoor acquisition.

1. INTRODUCTION

1.1 Context

The Building Indoor/Outdoor Modeling (BIOM) project aims at automatic, simultaneous indoor and outdoor modelling of buildings from heterogeneous data. The heterogeneity is both in data type (image and LiDAR) and acquisition platform (terrestrial indoor/outdoor and aerial). State of the art approaches generally deal with either the indoor or the outdoor, and often use strong priors of parallelism and orthogonality that are not necessarily verified.

Terrestrial laser scans can be acquired in static or dynamic mode, i.e., from static stations or a mobile mapping system. In order to cover all the faces of the object, several points of view are necessary. In this work, we assume the outdoor (mobile mapping) scan to be the reference as its georeferencing system has direct access to GNSS data, so our problem is to register the indoor scans together and with the outdoor scan. The indoor scans are each defined in a local reference frame relative to the laser scanner. The registration consists in referencing all these point clouds in the same coordinate system. The goal is therefore to determine the geometric transformation (rotation, translation) necessary to bring the data in coherence. According to (Monnier et al., 2013) the registration technique can be decomposed in:

1. Feature extraction from each dataset.
2. Feature matching
3. Using matched features to determine the optimal transformation to apply.

The main contribution of this paper is to leverage on the prior that man made object present large planar parts which give two rotational and one translational constraint, such that matching three pairs of polygons (with independent normals) is enough to recover a 6D (rotation+translation) transform between two scans.

1.2 State of the art

The ICP (Iterative Closest Point) algorithm is considered the most used approach in the registration of point clouds (Besl, 1992). ICP starts with two overlapping point clouds and initial guess. The transformation parameters can be iteratively estimated by generating pairs of corresponding points and minimizing the error metric. The major disadvantage of this method is the convergence towards a local solution if the initial data are not spatially close or if the initial transformation is poorly estimated.

Several variants of the ICP algorithm have been proposed to improve its robustness such as using a point-to-plane error metric. At each iteration of the algorithm, the relative pose that gives the minimal point-to-plane error is usually estimated using a standard nonlinear least-squares methods. (Low, 2004) proposes an approximation of the nonlinear optimization problem with a linear least-squares one that can be solved more efficiently. An extension of the ICP framework to nonrigid registration that uses the same convergence properties of the original algorithm was proposed in (Amberg et al., 2007). Another approach (Sharp et al., 2002) proposes to use Euclidean invariant features in a generalization of ICP registration of range images. To find the correspondence of 3D range camera, the authors proposed to use either spherical harmonics or the second order momentum. A new method for detecting uncertainty in pose has been introduced in (Gelfand et al., 2003), where the transformations that can cause unstable sliding in the ICP algorithm have been estimated using a sampling strategy and the points that best contain this sliding have been picked. Other non-iterative methods have dealt with the problem of registration

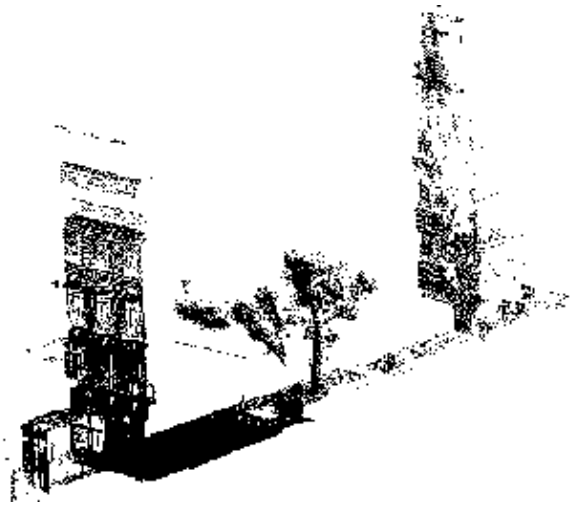

Figure 1. Outdoor MLS scan acquired with an MMS

of 3D point clouds, such as the method proposed in (Theiler et al., 2012), where the authors used virtual tie points generated by the intersection of three non-parallel planes in two different scans. The objective of this registration algorithm is to search for the assignment which preserves the geometric configuration of the largest possible subset of all tie points. The authors of (Forstner, Khoshelham, 2017) have proposed methods to efficiently register point clouds by introducing new optimal and sub optimal direct solutions based on plane-to-plane correspondences for determining the relative motion. A fast 3D point cloud registration method has been proposed in (Xiao et al., 2012), where the objective is to maximize the spherical correlation on S^2. A big planar patches have been employed as attributes to find the maximum using a novel search algorithm. Other works based on deep learning techniques have been proposed in the literature, such as the method proposed in (Elbaz et al., 2017), where the authors selected super points using a Random Sphere Cover Set and then matched them. A deep neural network auto-encoder has been used to encode local 3D geometric structures.

1.3 Data

The data used for this study was acquired by two different means (Mobile LiDAR Scan for the outside and static scans for the inside) on the Zoological Museum of Strasbourg.

1.3.1 Outdoor data The outdoor data used to experiment our method is a Mobile LiDAR Scan (MLS) (cf Fig. 1 acquired with the Stéréopolis II Mobile Mapping System (MMS) (Paparoditis et al., 2012). The acquisition system gives access to the sensor topology inherent to such MLS acquisitions, that is usually lost during export to formats such as .las or .ply. The data was collected from three streets, north, south and east of the museum, the west façade facing a park inaccessible to the MMS. Each outdoor scan contains approximately 3 million points.

1.3.2 Indoor data The indoor data used in our study is composed of 30 static LiDAR scans of the inside of the Musée Zoologique, one or two per room. Each scan consists of roughly 500 million points, which were downsampled to around 2 million points for practical reasons.

1.4 Contributions

In this paper, we propose a new method for the indoor/outdoor registration that consists in first, extracting indoor and outdoor

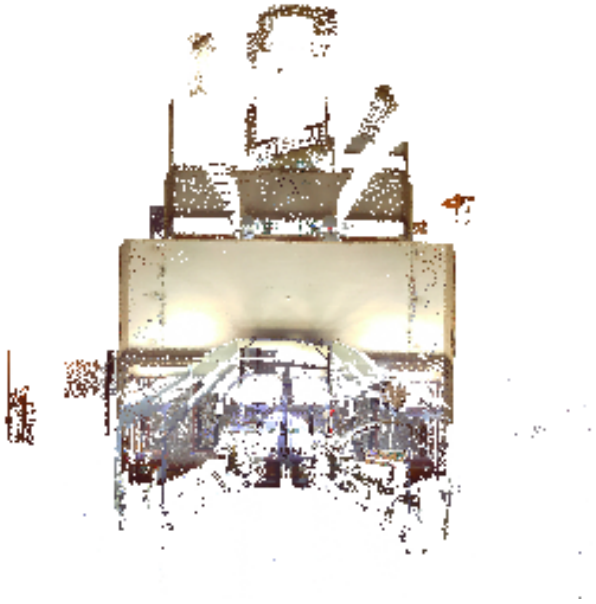

Figure 2. Indoor scan acquired in static mode inside the Zoological Museum of Strasbourg.

Figure 3. Indoor scans of the ground floor of the Zoological Museum of Strasbourg.

polygonal parts from the data, and then matching these polygons. The interest of polygons relative to planes is that they have a spatial extent limited to the areas where they have supporting points in the input data, so they form a simple and compact summary of our LiDAR scans. Given the specificity of LiDAR data that pass through windows, we propose to start with the extraction of the buildings' interior points captured from external scans; these are the points where the laser ray crosses the façades through apertures, mostly windows. Afterwards, we perform the registration. The work carried out has confirmed that the environment and the type of data drive the choice of the registration algorithm. For example, in man-made environments, where most objects are bounded by planar surfaces, the ideal is to choose a method of registration based on plane correspondences (Theiler et al., 2012) or primitive correspondences. The basic idea of our contribution is to perform an extraction of the planar polygons, from both the indoor and outdoor data, then to group them into three clusters according to their normals, which will be used subsequently for the estimation of the transformation.

The remainder of this paper is organized as follows. Section 2

Figure 4. Indoor and outdoor scans acquired in static mode at the Zoological Museum of Strasbourg: blue(outdoor scans), RGB(indoor scans).

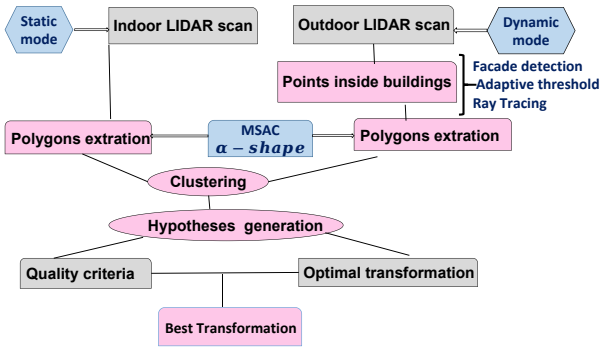

Figure 5. Pipeline Details

describes the two methods used for the extraction of planar polygons according to the nature of the data. Section 3 details the registration method, followed by results and discussion in Section 4. Conclusion and future works are proposed in Section 5.

2. PLANAR POLYGONS EXTRACTION

Due to its robustness to noise and outliers, RANSAC became the most popular method for LiDAR point cloud segmentation. Despite this success, it can generate false segments consisting of points from several nearly coplanar surfaces. False planes made up of points from different planes or roof surfaces represent a real obstacle for RANSAC (Xu et al., 2016). In order to exceed the limitations of RANSAC, we have exploited two methods depending on the nature of the data.

2.1 RANSAC Based on Sensor Topology

For the outdoor scans, we have access to the sensor topology (adjacency between successive pulses in the same line and between lines) so we can use it to enhance the polygon RANSAC detector. We use for that a recent method (Guinard et al., 2020) that exploits the sensor topology to extract compact planar patches instead of planes:

- Sample selection: as we are looking for compact planar patches, once a first sample point is drawn randomly from the point cloud, the other two are drawn in a local neighborhood (defined based on the sensor topology).

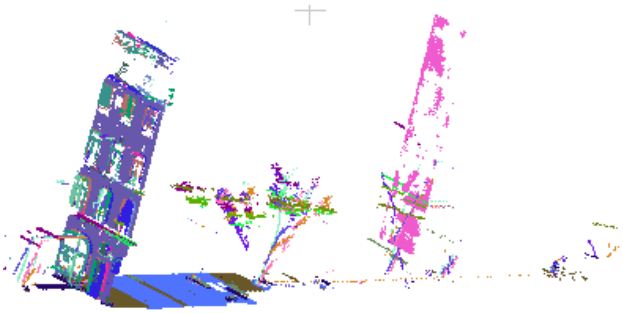

Figure 6. Inliers of the estimated planes from an outdoor scan computed with Sensor topology based RANSAC

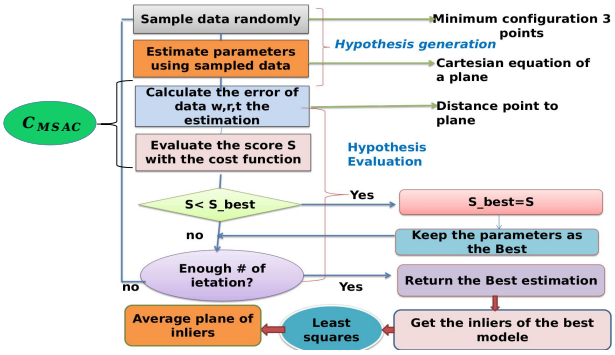

Figure 7. Robust estimation of a plane using MSAC .

- Region growing: instead of computing the distances of all the points of the cloud to the hypothetical plane, a region is grown starting at the first sample in order to recover a compact planar patch.

At each iteration, the planar patch with most inliers is selected and approximated by a single polygon using the α-shape algorithm (Edelsbrunner et al., 1983). Figure 6 shows the inliers of the detected planes from an outdoor scan.

2.2 Polygon MSAC

As we did not have access to the sensor topology for the indoor scans, we could not use the aforementioned method for the extraction of planar polygons. This is why, we have proposed a straightforward adaptation of MSAC (M-estimator Sample Consensus) which is a RANSAC extension that provides a potential solution to the spurious planes problem (Torr, Zisserman, 2000). While RANSAC gives the same unit score to all inliers:

$$C_{RANSAC}(e_i) = \begin{cases} 0 & if \ \ \epsilon_i^2 < t^2 \\ 1 & otherwise \end{cases} \quad (1)$$

MSAC gives each point a penalty score measuring how well the point corresponds to the model:

$$C_{MSAC}(\epsilon_i) = \min(\epsilon_i^2, t^2) \quad (2)$$

where ϵ_i is the distance of LiDAR point P_i to the current hypothetical plane and t is a distance threshold.

As RANSAC, MSAC produces hypothetical planes by randomly selecting three input points. The score of the sample plane is simply the sum of (2) over all the points P_i of the input scan. When the scores have been computed for all planes, the one with the highest score is extracted, its inliers (points P_i such that

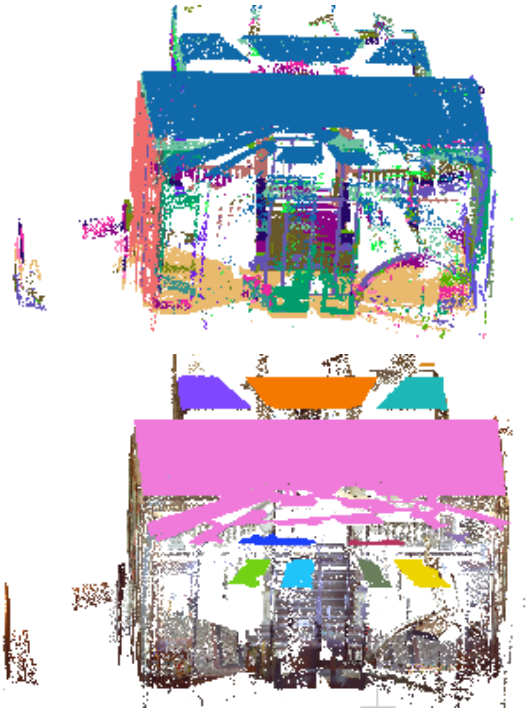

Figure 8. Top: Inliers of the estimated planes from an indoor scan, Bottom: Polygons extracted from the prominent plane.

$\epsilon_i < t$) removed from the point cloud, and the process is iterated until the score of the best plane is below a given threshold, as detailed in Figure 7. For each detected plane, we project its inliers to the plane and extract planar polygons by computing the $\alpha-$shape (Edelsbrunner et al., 1983) of these projected inliers. Figure 8 shows the inliers of the detected planes from an indoor scan and the polygons computed by the α-shape for the plane with the most inliers (corresponding to the ceiling).

3. INDOOR/OUTDOOR REGISTRATION

A crucial step of the BIOM project is achieving a registration of indoor and outdoor scans with accuracy close to the scan accuracy. To achieve such accuracy, we rely on matching planar polygons extracted from both indoor and outdoor data. This is quite difficult because obviously most polygons detected from one scan will not be visible in the other. In order to facilitate this matching, we will detect points from the outside scan that are inside the building and extract planar polygons only from these points.

3.1 Detecting points inside buildings in the outside scan

As the LiDAR beam usually passes through windows as shown in Figure 9, we propose to start by detecting façades from the outdoor scan, then detecting indoor points as points behind a façade by ray tracing.

3.1.1 Façade detection For façade detection, we decided to use the planar polygon detection of Section 2.1 to detect the façades as large vertical polygons. Thus we will keep only the detected planar polygons which are sufficiently vertical (deviation below 3°).

The choice of the threshold on inlier distance is a crucial factor as a bad choice can cause important under or over detection of

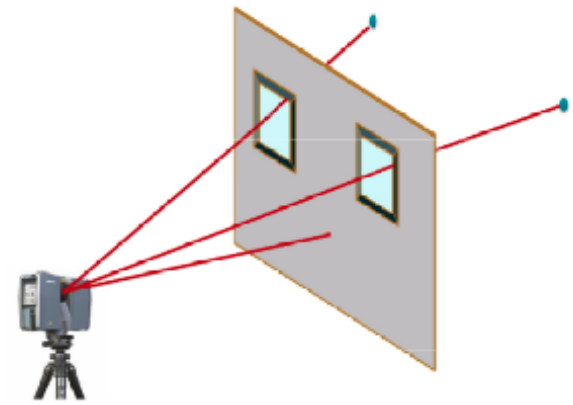

Figure 9. LiDAR rays that pass through windows.

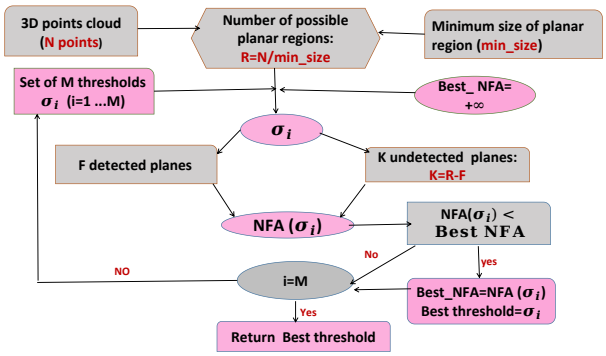

Figure 10. Design of the contrario validation method.

façades, leading to a bad detection of interior points. In order to select automatically an appropriate threshold, we propose an *a contrario* algorithm that takes as input M possible thresholds and selects the threshold with the lowest NFA (Number of False Alarms). Given a 3D point cloud of N points, the definition of the minimum size $\min_{\text{size}}$ of a planar region allows us to deduce the maximum number of possible planar regions:

$$R = \frac{N}{\min_{\text{size}}} \tag{3}$$

For each given threshold, RANSAC estimates $F \leq R$ planes, so at most $K = R - F$ planes are undetected as shown in Figure 10. Inspired by (Bughin, Almansa, 2010), we can define the error as being the probability that a significant existing planar region is not detected. Another possibility is the probability that the number of detected planes is less than the number of undetected planes:

$$NFA(\sigma) = N_{\text{iter}} \times p(K \geq F \mid \sigma) \tag{4}$$

The probability $p(K \geq F \mid \sigma)$, can then be upper bounded by the tail of the binomial law of parameter P defined by:

$$P(\sigma) = \frac{2\sigma \times diag^2}{v} \tag{5}$$

where $diag$ is the diagonal diameter of the bounding box of the 3D point cloud and V its volume. Then:

$$p(K \geq F \mid \sigma) \leq \sum_{j=F}^{K} \binom{K}{j} P(\sigma)^j \left(1 - P(\sigma)\right)^{K-j} \tag{6}$$

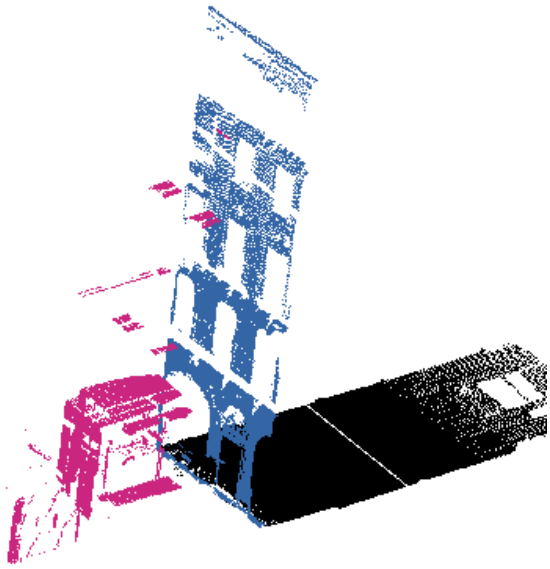

Figure 11. Indoor points detected from an outdoor scan, pink: the detected points, blue : inliers of the vertical plane that represents the façade

3.1.2 Ray Tracing Once the façades are detected, we detect indoor points by ray tracing:

1. For each point P_i, trace a ray R_i from the LiDAR optical center to P_i.
2. Find the intersection points P_i^j of R_i with the supporting planes $\mathcal{P}_j$ of each façade polygon F_j.
3. Test if P_i^j is inside the polygon (using the CGAL library.
4. In order to remove ouliers close to the façade plane, we additionaly require the point to be at least 1 m away from the intersected polygon: $(dist(P_i, P_i^j) > 1m)$.
5. If one P_i^j satisfies both criteria (inside the polygon and sufficiently far from the façade), tag it as an indoor point.

A result of this indoor point detection method is presented in Figure 11. Finally, the polygon detection algorithm of Section 2 is run only on the points detected as indoor from the outdoor scan, yielding a limited number of polygons.

3.2 Matching planar polygons

Selection of correspondences is a crucial step for the registration. If we have at least three correct correspondences of polygons with independent normals, it is possible to find the relative rotation/translation between the indoor and outdoor scans. We will start by presenting a simple threshold based matching in Section 3.2.1 in order to introduce the three main criteria used to match polygons, then propose a more robust matching based on clustering algorithm in Section 3.2.2

3.2.1 Threshold based matching Let us call:

- P_1 and P_2 detected planar 3D polygons from the two LiDAR scans: indoor scan $(Scan_1)$ and outdoor scan$(Scan_2)$
- $\mathbf{n}_1$ and $\mathbf{n}_2$ the corresponding plane normals
- $\mathbf{g}_1$ and $\mathbf{g}_2$ the corresponding centroids.
- $\mathcal{P}_{1,2}$ their bisector plane.
- P_1' and P_2' the projections of P_1 and P_2 on the plane $\mathcal{P}_{1,2}$.

We propose first a simple filter based on three measures:

1. Angle: $Angle(P_1, P_2) = \langle \mathbf{n}_1, \mathbf{n}_2 \rangle$
2. Distance: as there is not a standard definition for this distance, we have chosen to define it as the sum of distances from the centroid of each polygon to the bisector plane:

$$Distance(P_1, P_2) = dist(\mathbf{g}_1, \mathcal{P}_{1,2}) + dist(\mathbf{g}_2, \mathcal{P}_{1,2}) \quad (7)$$

3. Overlap:

$$Overlap(P_1, P_2) = \frac{|P_1' \cap P_2'|}{\min(|P_1'|, |P_2'|)} \quad (8)$$

where $|.|$ denotes the area. Note that we did not use the union on the denominator because a planar part of the inside scene is seen through an opening from the outside so only a very limited portion of it can be detected which would result in very low overlap with a more standard definition.

In practice, finding appropriate thresholds for these three criteria is tedious and leads often to multiple or no matchings. This is why we propose a more robust approach.

3.2.2 Cluster based matching We propose a more robust matching for polygons based on three main principles: clustering the polygons (by direction then by offset), enumerating match hypotheses between the clusters and evaluating which hypothesis is the best.

Algorithm 1 Greedy direction clustering

Input: P_i a set of planes with normals $\mathbf{n}_i$ and number of inliers n_i. ϵ a tolerance angle (typically $\pi/4rad$).
Clusters initialization:
- $C_1 = \{P_1\}$ where P_1 is the plane with most inliers
- $C_2 = \{P_2\}$ where P_2 is the plane with most inliers among planes for which $\mathbf{n}_i.\mathbf{n}_1 < cos(\epsilon)$
- $C_3 = \{P_3\}$ where P_3 is the plane with most inliers among planes for which $\mathbf{n}_i.\mathbf{n}_1 < cos(\epsilon)$ and $\mathbf{n}_i.\mathbf{n}_2 < cos(\epsilon)$

Mark P_1, P_2 and P_3 as processed and all other P_i as unprocessed
Each cluster C_k has a normal $\vec{c}_k$ computed as the weighted centroid of all the normals of the planes in C_k
Let P_{cur} be the unprocessed plane with the most inliers. Mark P_{cur} as processed.
compute $k_{min} = \arg\min_k 1 - |\vec{n}_i.\vec{c}_k|$
compute $d_i = 1 - |\vec{n}_i.\vec{c}_{k_{min}}|$
If $d_i < \epsilon$, add P_i to the cluster $C_{k_{min}}$.
Call C^h and tag as horizontal the cluster for which $1-|\vec{z}.\vec{c}_k|$ is minimum.
Call C^{v_1} and C^{v_2} and tag as vertical the two remaining clusters.

Direction clustering: For each input scan, we greedily cluster planes P_i according to their normals $\vec{n}_i$ by decreasing number

of inliers n_i as detailed in Algorithm 10 that produces 3 clusters C^h, C^{v_1} and C^{v_2} for each scan.

Direction cluster matching We associate C^h for the two scans. We associate each remaining cluster (C^{v_1} and C^{v_1}) to the vertical cluster from the other scan with the smallest angle.

Alignment: we compute the rotation that aligns the three clusters: the vertical C^{v_1} and C^{v_2} and the horizontal clusters using the method described in Section 3.3

Hypotheses enumeration: for each associated direction cluster, we enumerate possible plane matches:

- For each of the three associated main cluster pair $A_m = (C^1_{k_1}, C^2_{k_2})$ we will call I_m the set of pairs of planes $(P^1_{i_1} \in C^1_{k_1}, P^2_{i_2} \in C^2_{k_2})$
- for each (J_1, J_2, J_3) in $I_1 \times I_2 \times I_3$, compute the translation that aligns the planes in (J_1, J_2, J_3) using the method described in Section 3.4.
- note that I_3 is extracted from the horizontal clusters which cover the displacement along z,I_2 is extracted from the vertical clusters which have the smallest angle with the y axis and I_1 is extracted from the vertical clusters which have the smallest angle with the X axis.
- keep the translation calculated by the best hypothesis.

Selection of the best hypothesis: To asses which hypothesis is the best, we define a criterion that (1) is robust to outliers as many planes detected in one scan have no counterpart in the other (2) favors important polygon overlaps and (3) favors small distances over these overlapping parts. For robustness, we need a distance threshold d_{thr} above which a polygon is just considered an outlier. Then for two polygons P_1, P_2 with centroids O_1, O_2 in the same cluster, we define a robust error:

$$E^{d_{thr}}(P_1, Scan_2) =$$

$$\frac{1}{d^2_{thr}} \int_{P \in P_1} min(d^2_{thr}, dist(P, Scan_2)^2) dP \tag{9}$$

We see that the distance is 0 when P_1 is completely overlapped by polygons of $Scan_2$ with a distance of 0 (the perfect case) and increases as distance augments and overlap decreases up to $|P_1|$ when the overlap is empty and/or the distance is over d_{thr}. In practice we make the approximation:

$$E^{d_{thr}}(P_1, Scan_2) \approx$$

$$\sum_{P_2 \in Scan_2} |P^2_1 \cap P^1_2| \frac{max(0, d^2_{thr} - Distance(P_1, P_2)^2)}{d^2_{thr}} \tag{10}$$

Note that we do not use our modified relative $Overlap$ function anymore as a large overlap surface should be favoured to a small one as it correponds to more input points. Finally our global criteria writes:

$$rdist(Scan_1, Scan2) = \sum_{P_1 \in Scan_1} E^{d_{thr}}(P_1, Scan_2)$$

We consider that the best hypothesis is the one that maximizes this criterion.

3.3 Rotation estimation

Assuming that we have an association between three directional clusters of two scans:

- C^h_1 and C^h_2 the horizontal clusters of scans 1 and 2, $\vec{c}^h_1$ and $\vec{c}^h_2$ their normals.
- $C^{v_1}_1$ the first vertical cluster of scan 1 and $C^{v_1}_2$ the associated vertical cluster of scan 2, $\vec{c}^{v_1}_1$ and $\vec{c}^{v_1}_2$ their normals.
- $C^{v_2}_1$ the second vertical cluster of scan 1 and $C^{v_2}_2$ the asociated vertical cluster of scan 2, $\vec{c}^{v_2}_1$ and $\vec{c}^{v_2}_2$ their normals.

We want to find the rotation that best aligns these three pairs of directions. Let $\vec{u}_1$=$\vec{c}^h_1$ and $\vec{u}_2$=$\vec{c}^h_2$. We are looking for the two vector:

$$\{\vec{v}_1, \vec{v}_2\} = \underset{\vec{v}_1 \in \{\vec{c}^{v_1}_1, \vec{c}^{v_2}_1\}, \vec{v}_2 \in \{\vec{c}^{v_1}_2, \vec{c}^{v_2}_2\}}{\arg\min} \langle v_1, v_2 \rangle \tag{11}$$

We want to create a first orthogonal basis M_1 from u_1 and v_1, and a second orthogonal basis M_2 from u_2 and v_2, so we have:

- $\vec{q}_1$=the projection of $\vec{v}_1$ on the orthogonal plane to $\vec{u}_1$
- $\vec{q}_2$=the projection of $\vec{v}_2$ on the orthogonal plane to $\vec{u}_2$

$$Proj_{orth\_Plane_{\vec{u}_1}}(\vec{v}_1) = \vec{v}_1 - \frac{(\vec{v}_1.\vec{u}_1)}{|\vec{u}_1|^2}|\vec{u}_1| \tag{12}$$

$$Proj_{orth\_Plane_{\vec{u}_2}}(\vec{v}_2) = \vec{v}_2 - \frac{(\vec{v}_2.\vec{u}_2)}{|\vec{u}_2|^2}|\vec{u}_2| \tag{13}$$

$$\vec{l_1} = \vec{u}_1 \times \vec{q}_1 \tag{14}$$

$$\vec{l_2} = \vec{u}_2 \times \vec{q}_2 \tag{15}$$

we have:

$$M_1 = \begin{bmatrix} u_1(x) & q_1(x) & l_1(x) \\ u_1(y) & q_1(y) & l_1(y) \\ u_1(z) & q_1(z) & l_1(z) \end{bmatrix} \tag{16}$$

and

$$M_2 = \begin{bmatrix} u_2(x) & q_2(x) & l_2(x) \\ u_2(y) & q_2(y) & l_2(y) \\ u_2(z) & q_2(z) & l_2(z) \end{bmatrix} \tag{17}$$

We can calculate the rotation as the base change matrix between M_1 and M_2:

$$M_2 = RM_1 \tag{18}$$

$$R = M_2 M_1^{-1} \tag{19}$$

3.4 Translation estimation

We have defined a hypothesis as three pairs of planes. Each pair is extracted from two matched clusters ($C_i \in scanA, C_j \in scanB$).
For each hypothesis we have:

$pair_l = \{Plane_{i,l}(N_{i,l}, d_{i,l}), Plane_{j,l}(N_{j,l}, d_{j,l})\}$.
where l represents the pair index in the hypothesis and i,j represent the indices of the two matched clusters. For each pair we have the following properties:

$$N_l = RN_{i,l} \tag{20}$$

$$N_l T = d_{i,l} - d_{j,l} \tag{21}$$

therefore according to the two equations above and the definition of a hypothesis, we can deduce the following linear system:

$$\begin{bmatrix} N_1(x) & N_1(y) & N_1(z) \\ N_2(x) & N_2(y) & N_2(z) \\ N_3(x) & N_3(y) & N_3(z) \end{bmatrix} \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = \begin{bmatrix} d_{i,1} - d_{j,1} \\ d_{i,2} - d_{j,2} \\ d_{i,3} - d_{j,3} \end{bmatrix} \tag{22}$$

by solving this linear system, we can find the translation.

4. EVALUATION AND DISCUSSION

Our proposed method has been tested on real data to evaluate its effectiveness. The proposed algorithm works without any constraints on the initial position of the two scans, unlike iterative methods which require the correct estimation of the initial position to be able to converge to a global solution. The key step of our algorithm is the estimation of planar polygons which was carried out using two methods depending on the nature of the data.

The first evaluation of our algorithm was carried out on two interior scans from which a perfect registration was done manually. Each input scan was subsampled to around 260 000 points. Starting from this perfect position, we have altered one of the two scans with an initial translation error ranging from a few centimeters to a few meters and an initial rotation error ranging from a few degrees to a few tens of degrees as shown in Table1. The chosen setting allowed us to estimate 22 planes from the first scan, and 23 planes from the second scan. The obtained results show that our algorithm is able to register the two scans within a reasonable calculation time regardless of the initial error. Afterwards, we performed the registration of indoor and outdoor scans. As we do not know the ground truth we only considered the visual results. We consider that the obtained results demonstrate our algorithm's ability to efficiently register indoor and outdoor data as shown in Figure 12.

Achieving precise results requires fine-tuning of the algorithm's parameters. The iterations number of MSAC must be calculated according to the number of points of the scan to ensure the robustness of the algorithm. The number of estimated planes depends on the inliers threshold and the minimum size of a planar region, whereas the number of extracted polygons is reliant on the value of α. If we estimate more planes, we can generate more hypotheses and use more polygons for the evaluation of each hypothesis, and therefore we can find a more precise result.

Implementation: Our algorithm was implemented in C++. All geometric calculations were carried out using the CGAL library `https://www.cgal.org/`). In addition we used the Boost library `https://www.boost.org/` to calculate polygons intersection, Eigen library `https://eigen.tuxfamily.org/dox/` for matrix calculation and The Point Cloud Library (PCL) `https://pointclouds.org/` for point cloud processing.

Initial error		Final error		Time
Metric error	Angular error	Metric error	Angular error	
0.099 m	0.001°	0.00793 m	0.0008°	263 s
0.44 m	0.528°	0.0213 m	0.06°	244 s
1.27 m	3°	0.0241 m	0.043°	269 s
4.15 m	12°	0.0402m	2°	294 s

Table 1. Result of indoor/indoor registration tests

5. CONCLUSION AND FUTURE WORKS

In the present work, we have proposed a method based on polygon detection and matching to address the challenging problem of indoor/outdoor registration whereas state of the art approaches tackle either the indoor or the outdoor registration problem. To the best of our knowledge, no method has proposed a joint indoor/outdoor registration with a unified formalism. Our preliminary tests have highlighted the potential of the proposed method. Our main perspective concerning this work is to extract outdoor points from the indoor scans in order to get additional information, and to precisely detect windows outline in order to introduce additional constraints in the registration.

6. ACKNOWLEDGMENT

This work has been supported by the Building Indoor/Outdoor Modelling ANR-17-CE23-0003 BIOM project. The authors would like to thank Mrs. Tania LANDES who provided them with the indoor data.

REFERENCES

Amberg, B., Romdhani, S., Vetter, T., 2007. Optimal step nonrigid icp algorithms for surface registration. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 1–8.

Besl, P. et McKay, N., 1992. A method for registration two 3-D shape bacterial availability. *IEEE Trans Pattern Analysis and Machine Intelligence*, 14, 232–256.

Bughin, E., Almansa, A., 2010. Planar patch detection for disparity maps. *Proc. 3DPVT*, Citeseer.

Edelsbrunner, H., Kirkpatrick, D., Seidel, R., 1983. On the shape of a set of points in the plane. *IEEE Transactions on information theory*, 29(4), 551–559.

Elbaz, G., Avraham, T., Fischer, A., 2017. 3d point cloud registration for localization using a deep neural network auto-encoder. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4631–4640.

Forstner, W., Khoshelham, K., 2017. Efficient and accurate registration of point clouds with plane to plane correspondences. *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2165–2173.

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

Guinard, S. A., Mallé, Z., Ennafii, O., Monasse, P., Vallet, B., 2020. Planar Polygons Detection In Lidar Scans Based On Sensor Topology Enhanced RANSAC. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 5(2).

Figure 12. Result of the test of our algorithm on two pairs of indoor/outdoor scans, Top: the initial position of each pair, Bottom: The position of each pair after registration,pink points represent the indoor points detected from the outdoor scans.

Low, K.-L., 2004. Linear least-squares optimization for point-to-plane icp surface registration. *Chapel Hill, University of North Carolina*, 4(10), 1–3.

Monnier, F., Vallet, B., Paparoditis, N., Papelard, J.-P., David, N., 2013. Mise en cohérence de données laser mobile sur un modele cartographique par recalage non-rigide. *Revue Française de Photogrammétrie et de Télédétection*, 202, 27–41.

Paparoditis, N., Papelard, J.-P., Cannelle, B., Devaux, A., Soheilian, B., David, N., Houzay, E., 2012. Stereopolis II: A multi-purpose and multi-sensor 3D mobile mapping system for street visualisation and 3D metrology. *Revue française de photogrammétrie et de télédétection*, 200(1), 69–79.

Sharp, G. C., Lee, S. W., Wehe, D. K., 2002. ICP registration using invariant features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1), 90–102.

Theiler, P., Schindler, K. et al., 2012. Automatic registration of terrestrial laser scanner point clouds using natural planar surfaces. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3, 173–178.

Torr, P. H., Zisserman, A., 2000. MLESAC: A new robust estimator with application to estimating image geometry. *Computer vision and image understanding*, 78(1), 138–156.

Xiao, J., Adler, B., Zhang, H., 2012. 3d point cloud registration based on planar surfaces. *2012 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, IEEE, 40–45.

Xu, B., Jiang, W., Shan, J., Zhang, J., Li, L., 2016. Investigation on the weighted ransac approaches for building roof plane segmentation from lidar point clouds. *Remote Sensing*, 8(1), 5.