Odometry from Planar Landmarks

Keerthi Narayana and Bruno Steux

***Abstract*— This paper presents a new perception odometry approach using extracted stationary planar features to resolve 5 degrees of freedom of the robot motion. The approach exploits the geometrical properties of the extracted features to determine the transformation of the moving robot, which has perceived these landmarks. This way of localizing can help several applications in indoors and outdoors such as urban canyons, with plenty of planar features. The paper presents the concept and the algorithm, and validates them using a simulated scenario.**

I. INTRODUCTION

Several approaches are proposed in the recent years to provide assistance and solutions to the 3D localization (pose estimation) problem. Most solutions use Global Navigation Satellite Systems (GNSS) and Inertial Measurement Units (IMU). However, pose computed from GNSS receivers degrade especially in urban and indoor environments, where satellite signal reception is perturbed by manmade structures. Consequently, Inertial Measurement Units (IMU) are a good alternative to fill the short gaps in GNSS-based pose. However, the pose computed from these sensors subjected to drift errors, and a good quality IMU costs a lot.

In the indoor robotic navigation, the technique of Simultaneous Localization and Mapping (SLAM) is used for localization. It uses perception sensors to create the map of the environment, and this information is used either to correct the estimated pose or compute the robot transformation. The first implementation is a Bayesian filter approach ([1], [2], [3]), where the predicted pose from the additional sensors such as odometers are corrected using the constructed maps, containing mainly a set of selected features (landmarks) from the scene. Alternatively, robot transformation can also be computed by comparing two overlapping scans taken at two different instants of time. A featureless scan correlation ([4], [5]) is performed to resolve transformation using an optimization formulation.

However, the Bayesian approach already incorporates the errors of the additional sensors. All the accumulated uncertainties over a period time, causes erroneous associations between landmarks in the constructed map and the new observations, resulting in the failures of the approach ([2]). On the other hand, featureless approaches are computationally heavy and the optimization formulation is nondeterministic.

K. Narayana was with the center of robotics (CAOR), Mines ParisTech, 60 Bld Saint Michel, 75006 Paris, France keerthi2narayan@gmail.com

B. Steux is with the center of robotics (CAOR), Mines ParisTech, 60 Bld Saint Michel, 75006 Paris, France Bruno.Steux@mines-paristech.fr

Therefore, we propose odometry (dead reckoning) approach, where certain landmarks are extracted from the scene, and from the geometrical properties of these stationary landmarks observed at two instants of time, the relative pose of the robot is computed. A similar approach is presented in [7], where a 2D transformation is estimated from the circular landmarks. These landmarks are extracted from the tree trunks in a parking area. We choose planar features as the landmarks, since they are one of the most recurring features of the manmade environments. Moreover, their time invariant geometrical properties can be used for both identifying the correspondence between two observation sets (Data Association), and to compute the undergone robot 3D transformation.

II. 3D TRANSFORMATION

Most SLAM problems are tackled only in the 2D space. This is mainly due to the associated complex 6 Degrees of Freedom (6DoF) in the 3D space, and the unknown associations between the landmarks. Moreover, there is a lack of 3D pose sensors, apart from GNSS receivers and Inertial Measurement Units (IMU). Our approach addresses the 3D pose resolution using the geometrically invariant properties of the extracted features, i.e. planar landmarks.

We tackle the localization problem in a 3D space with abundant stationary planar features. These features are observed by the moving observation platform such as a robot. A way to extract such planar features from the mobile platform is published in our previous work [6].

A. 3D Rotation

In order to compute a 3D rotation of a observing platform, a 3D reference frame is needed which can depict the undergone rotation. Planar features, such as building facades can be characterized by their normal vectors pointing in the direction perpendicular to their surface. By using two non parallel planes, the desired 3D reference frame can be constructed, as shown in figure 1.

Once the 3D rotation frame is determined at a given observation epoch $\mathbf{k}$, and the same can be constructed in another epoch $\mathbf{k}+\delta$, from the same pair of planes $\mathbf{P_1}$ and $\mathbf{P_2}$, the undergone rotation $\mathbf{R_k}$ of the observing moving platform can be determined, as shown in equation 1. The assumption is that the planes are stationary with respect to the moving mapping platform between the two observation epochs.

$$\left.\begin{aligned} &\mathbf{U_{k+\delta}} = \mathbf{R_k^{-1}} \mathbf{U_k} \\ &where \\ &\mathbf{U} = \begin{bmatrix} \overrightarrow{n_1} & \overrightarrow{n_2} & \overrightarrow{n_3} \end{bmatrix} \end{aligned}\right\} \quad (1)$$

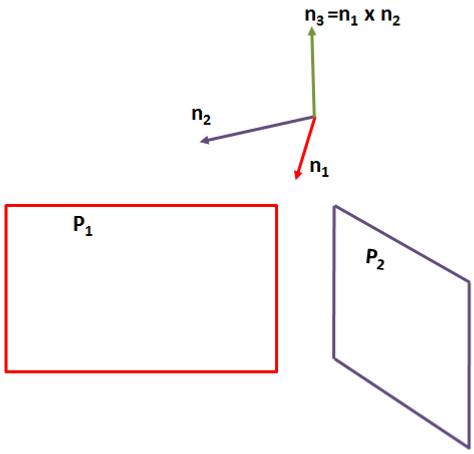

Fig. 1. Construction of 3D rotational frame from two non parallel planes, $\mathbf{P_1}$ and $\mathbf{P_2}$ using their respective normal vectors $\mathbf{n_1}$ and $\mathbf{n_2}$. Third normal $\mathbf{n_3}$ is computed by the vector cross product between the two normals.

$\mathbf{U_{k+\delta}}$ and $\mathbf{U_k}$ given in equation 1, are constructed from the observed and associated pair of planes of the two observation epochs, as shown by $\mathbf{U}$.

Manmade environments often have two or more non parallel planes in the observable scene. In such scenarios, the 3D rotation of the observing sensor platform can be computed from a pair of non parallel planes identified across the two observation epochs. The advantage of the this rotation computation is, it remains independent to the translation of the platform. Therefore, it can be resolved independently.

B. 3D Translation

3D translation computation demands a stationary 3D reference point in a fixed reference frame. However, for the moving observing platform, this point appears to have undergone a translation same in magnitude but reverse in direction.

Planar landmarks are uniform surfaces and not necessarily comes with a reference 3D point. At many instants even the observed patches of these landmarks vary in their border length and surface areas, depending on the pose of the platform. A way to determine the reference point is to use the intersection points of the planar features. As shown in figure 2, a 3D reference point can be constructed only from limited combinations of the planar intersections.

Moreover, observing non parallel planes to the ground is feasible only in some rare scenarios. Most manmade environments have erect walls, and the horizontal roofs are not visible for a ground vehicle. However, an alternative way to determine the 3D reference point is to use a sensor such as a digital camera.

In our present approach which depends only on a 2D laser scanner setup, as presented in [6], we decided to opt to resolve the translation in 2D space. The 2D reference point ϖ is determined by projecting the intersecting line on to the horizontal plane as shown in the last sub figure in the

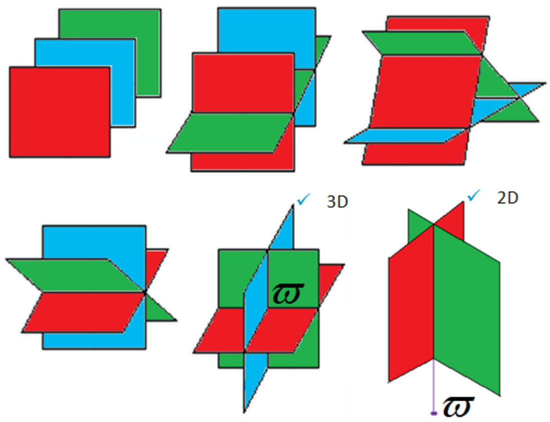

Fig. 2. Construction of 3D rotational frame from two non parallel planes, $\mathbf{P_1}$ and $\mathbf{P_2}$ using their respective normal vectors $\mathbf{n_1}$ and $\mathbf{n_2}$. Third normal $\mathbf{n_3}$ is computed by the vector cross product between the two normals.

figure 2. The translation vector $\mathbf{T_k}$ of the platform, can be expressed using equation 2.

$$\varpi_{\mathbf{k}+\delta} = \varpi_{\mathbf{k}} - \mathbf{T_k} \tag{2}$$

Again, to determine the 2D translation, we need two non parallel planes, and such scenarios are often present in manmade environments. The critical issue is to determine the corresponding pair of non parallel planes across the two observation epochs.

III. DATA ASSOCIATION

As explained, to compute 5DoF transformation (3DoF rotation and 2DoF translation) in 3D space, a pair of non parallel planes needs to be identified across the two observation sets. The identification of the corresponding element across two sets is widely known as Data Association (DA).

Uniform planar features do not come with any uniquely identifiable tags. Moreover, planar features appear and disappear from the observable scene, as the robot moves, making the identification problem non trivial. Additionally, we opted to compute the robot pose independent of other sensors; therefore, restricting the association to be based only on the geometrical properties of the observed landmarks.

A. Relative Constraints

The dihedral angles between a pair of observed planes is conserved regardless the observing position of the robot. This property is computed for each pair of the observed planes in an epoch. The first level of association is done by comparing this property of the pair of planes for the two observation sets. All parallel plane pairs are discarded. However, results often contain several associations including some ambiguous ones, as it is likely to have the same angular relationships for more than one pair of planes in the scene.

B. Absolute Constraints

To minimize these erroneous associations, a constraint based on the knowledge that a land vehicle on a manmade roads do not move abruptly, can be applied. This implies, for the two observation sets done within a short time, the corresponding planes appear closer to each other. Moreover, if they are projected on one another, they have some overlapping regions. This conditions are tested using a distance criteria (spatial neighborhood) and a spatial alignment test for overlapping regions.

C. Lenient and Strict Data Associations

The planar features help separate the rotation computation from translation resolution. Once rotation is resolved the associations can be done relatively easier compared to the state, where the entire undergone transformation is unknown. Therefore, even the DA can be performed in two stages.

The absolute constraints are applied on observations across the two sets, therefore, the margin for the criteria varies depending upon the motion of the robot. If no information is available about the undergone motion of the robot, a higher magnitude threshold needs to be applied to provide margins for all the unknowns. In such case we call the DA, Lenient Data Association (LDA).

If the rotation is resolved prior computing translation, at this stage, a more stringent condition can be applied to reduce ambiguities in associations. This step we call it as Strict Data Association (SDA). It also employs algorithms to ensure injective associations, meaning, an associated plane is retained only with the most suitable match in the other set.

IV. HANDLING OUTLIER ASSOCIATIONS

As mentioned before, the proposed approach, simplifies the 6DoF pose problem by not only separating the rotation from translation, but also by splitting the Data Association in two, LDA and SDA. We term this as a Divide and Conquer (D&C) approach.

However, the lenient constraints can lead into several ambiguous associations. Additionally, different associated planar landmark pairs can compute results with varying precision, mainly due to the distance of these planes and their orientation to the observing platform. Therefore, a strategy to handle multiple associations is essential. In the presence of multiple associated planar pairs, several transformation solutions can be computed. However, in the presence of ambiguous or outlier associations the common choice of mean-based approach often fails. Therefore, we devised a new algorithm, to choose an associated pair of planes which optimally describes the undergone transformation of the robot.

The algorithms, Optimal Candidate Selection by Consensus (OCSC) and its weighted variant (WOCSC) are inspired from the RANSAC algorithm [8]. The algorithm OCSC like RANSAC tries to choose the best candidate solution by a maximizing function. However, RANSAC is nondeterministic, as it tries to select the best fitting model parameters to a given large population by randomly sampling a minimal set, till a satisfactory solution can be computed using this random set. However, in the deterministic OCSC each candidate minimal set generates a solution, and if the solution is the one best fits all the members of the associated set then the solution is retained. The OCSC algorithm has two steps, an expectation step and a consensus step. The expectation step generates a candidate transformation from a pair of associated planes, and the consensus step, applies this result on all the associated planes, from the set of planes of one of the epoch. The optimal transformation is the one which produces the best fit (minimum error) for all the associated planes across two epochs.

The problem is formulated using equation 3.

$$\mathbf{S_{k+\delta}} = \mathbf{K_k} \bigtriangledown \mathbf{S_k} \Rightarrow \forall \mathbf{i} \in \{1 \dots p\}, \mathbf{x^i_{k+\delta}} = \mathbf{K_k} \bigtriangledown \mathbf{x^i_k} \quad (3)$$

where, $\mathbf{S}$ is a varying population observed at two epochs $\mathbf{k}$ and $\mathbf{k}+\delta$. In our case, it corresponds to the set of planar landmarks. The variation can be described using an operand $\mathbf{K_k}$ and operator $\bigtriangledown$. Each member of the population $\mathbf{x}$ undergoes the same transformation. For the rotation, $\mathbf{K_k}$ is the rotation matrix $\mathbf{R_k^{-1}}$ and $\bigtriangledown$ is matrix vector multiplication. For the translation, $\mathbf{K_k}$ is the translation vector $-\mathbf{T_k}$ and $\bigtriangledown$ is vector addition.

Expectation step generates the candidate $\mathbf{K^j_k}$ from a minimal subset, in our case, an associated pair of non parallel planes $\mathbf{j}$. In the consensus step, the optimal candidate $\mathbf{K^*_k}$ is selected, as shown in equation 4.

$$\mathbf{K^{\star}_k} = \underset{\mathbf{j} \mapsto \mathbf{K^j_k}}{\arg\min} \sum_i ((\mathbf{K^j_k} \bigtriangledown \mathbf{x^i_k}) - \mathbf{x^i_{k+\delta}})^2 (\mathbf{w^i_k} \cdot \mathbf{w^i_{k+\delta}}) \quad (4)$$

The improvement of WOCSC with respect to OCSC is the term $(\mathbf{w^i_k} \cdot \mathbf{w^i_{k+\delta}})$, which takes for each element, the confidence $\mathbf{w}$ into account. Each planar landmarks are considered to have an associated confidence (refer CPEF on [6]). It shall be noted that only the associated pairs $\mathbf{i}$ of observations $(\mathbf{x^i_k}, \mathbf{x^i_{k+\delta}})$ are used in the consensus, and not all the observed planes from epochs $\mathbf{k}$ and $\mathbf{k}+\delta$.

The algorithm, handles outliers and reduces noise by choosing the optimal solution. In its general form, it can be applied to solve any overdetermined system.

V. DIVIDE AND CONQUER APPROACH

The overall approach used is summarized in figure 3.

VI. EXPERIMENTAL RESULTS

A. Test Scenario

The concept and the algorithms are validated using a simulation platform. The platform helps to know the true trajectory traversed by the robot, and it provides the laser scanner range measurements(refer [6]) from each known pose of the robot, which is the input for our algorithms. Another advantage is that these inputs can be altered by adding configured noises. These algorithms are at present implemented in MATLAB. The used trajectory and the scene are shown in figure 4.

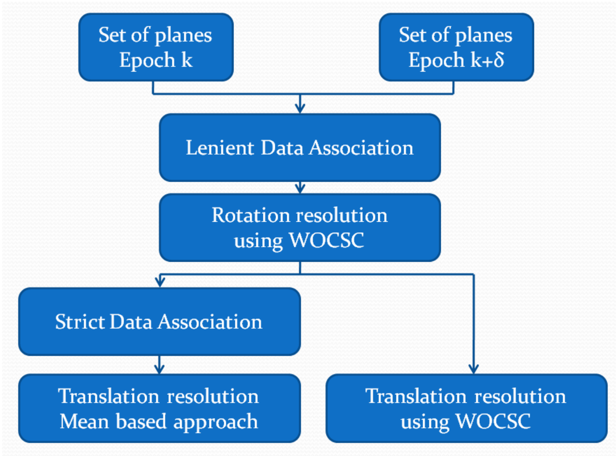

Fig. 3. This provides the overall view of Divide and Conquer (D&C) approach, where rotation is separated from translation. First a Lenient Data Association (LDA) is performed to tackle all unknowns of the undergone transformation. Therefore, rotation is resolved using Weighted Optimal Candidate Selection by Consensus (WOCSC) algorithm. Translation can be resolved using two ways, either using a mean based approach or again by applying WOCSC. If mean based approach is applied, it is a must to ensure that there is no outlier associations, and this is done by using Strict Data Association (SDA). SDA is optional if we apply WOCSC algorithm.

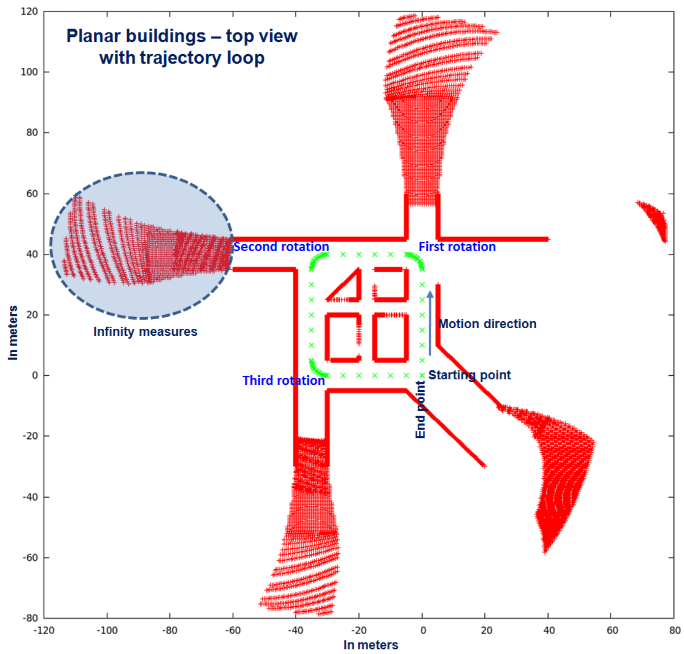

Fig. 4. Top view of the scene generated from the measured laser scans, shown in red dots. Infinity measures are the laser scan measurements for far distance, we discard in our algorithm. Straight red lines shows the walls, corresponding to planar features. Green dots, traversing a loop, is the true trajectory.

We generated test data with four different noise levels, which is added to each laser range measurement. First, a no noise case $\sigma = 0$. Second $\sigma = 6mm$, corresponding to the precision of the SICK LMS 221 laser scanners, we tested inside our laboratory. Third, $\sigma = 2cm$, corresponding to the precision of a low quality laser scanners, and the fourth $\sigma = 10cm$, a worst case scenario.

B. Test Results

Figure 5 shows the four trajectory results obtained for the four different input data generated using the same scene described in figure 4 with explanation in its caption.

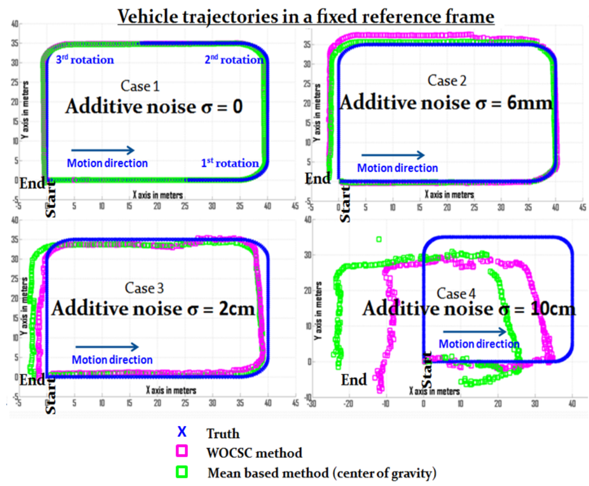

Fig. 5. Four trajectories generated using four data sets with different noise levels are shown in the four sub figures. All the trajectories follow the basic shape but as expected, higher the noise higher the trajectory error. In a fixed reference frame (trajectory) the accumulated error appears as a drift. Each of the four cases depicts the truth in blue cross, trajectory generated from the algorithm sequences (LDA,WOCSC Rotation,SDA,Mean Translation) in green, and (LDA,WOCSC Rotation,WOCSC Translation) in magenta. As discussed before, WOCSC approach (second sequence) performs better as the noise level increases, as shown by the plots in magenta. The error in the mean based approach in the last sub figure originates from the erroneous data association, result of the higher noisy data. Higher noises alters the estimated planar properties, which then contributes to the robot pose error. While computing the trajectory, these errors (non systematic if originates from DA ambiguities) gets accumulated resulting in drift errors.

The error characteristics obtained for the 5DOF relative pose estimation using our approach is summarized in figures 6 and 7. In the table the standard deviation (σ) is the important information providing the precision of the computed relative pose. σ for the pitch rate increases rapidly with the noise level, compared to the other two rotation rates. This is because, the planar landmarks in direction perpendicular to the vehicle motion (incidence angle $\mathbf{0}^{\circ}$ for the laser range measurements in our setup [6]), gets the full magnitude of the laser measurement noise. σ for the cog method gets higher compared to WOCSC approach showing its sensitive to the outlier associations. A noise with $\sigma = \mathbf{10}$cm, added to each laser range measurement, alters the estimated planar landmark properties. This change in the planar properties leads to a higher σ for the computed pose, especially in the cluttered environments.

	Mean roll rate error (deg/epoch)	Standard deviation (deg/epoch)	Mean pitch rate error (deg/epoch)	Standard deviation (deg/epoch)	Mean yaw rate error (deg/epoch)	Standard deviation (deg/epoch)
No noise	-2.85506E-15	9.51929E-14	1.03948E-14	1.83304E-13	-6.35649E-04	1.62722E-02
Noise σ= 6mm	-5.53156E-03	9.68693E-02	-2.72707E-03	1.75711E-01	-8.34166E-04	1.14980E-01
Noise σ= 2cm	-2.12593E-02	3.18210E-01	-3.00451E-02	5.53417E-01	8.14208E-04	2.72763E-01
Noise σ=10cm	-7.16991E-02	1.12940E+00	-4.70674E-02	2.38289E+00	8.36368E-03	7.65798E-01

Fig. 6. Angular Rate Precision - 3DoF

	Mean velocity errors (meters/epoch)				Std deviation (meters/epoch)			
	WOCSC		Cog		WOCSC		Cog	
	X	Y	X	Y	X	Y	X	Y
No noise	1.46903E-03	1.75199E-03	1.44865E-03	2.16769E-03	1.48305E-02	1.38290E-02	1.54236E-02	1.52860E-02
Noise σ= 6mm	5.97738E-03	-6.33657E-03	5.94436E-03	-1.41951E-04	9.64136E-02	1.09878E-01	7.19298E-02	7.19869E-02
Noise σ= 2cm	4.98491E-03	7.96491E-04	7.79074E-03	7.90890E-04	1.68949E-01	2.00265E-01	1.98067E-01	1.68686E-01
Noise σ=10cm	4.02695E-02	2.79718E-02	7.82631E-02	5.10067E-03	6.99082E-01	7.16818E-01	1.33231E+00	1.27200E+00

Fig. 7. Translation Precision - 2DoF

At present, we have applied no noise reduction filters, neither for the laser range measurements nor for the trajectory estimation, except the estimation of the plane is done by a least squares sense (refer [6]). Additional noise reduction filters, higher precision and higher rate perception sensors can help in a better estimation of the trajectory.

VII. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

The concept of odometry approach using planar landmarks is promising to address the current short comings of the localization techniques. It can address the 5DoF 3D pose problem without using any other additional sensor information. To determine the 5DoF pose, it is sufficient to have a pair of non parallel planes, observed in both epochs between which transformation is estimated. We also presented a new algorithm called Optimal Candidate Selection by Consensus (OCSC) and its weighted variant (WOCSC). They can be applied to resolve ambiguities in associations by handling outliers and noisy associations, while doing so chooses the optimal transformation undergone by the robot. The algorithms in its general form can be applied to solve any overdetermined system. We validated the concept using a simulated data with different levels of additive noise.

B. Future Works

Additional validations of the algorithm can be done including real data scenarios. The computation of the pose, solely from the perception sensor information means a higher trust on them. Therefore, a better quality perception sensors can help. There are higher quality 3D laser scanners such as Velodyne High Definition Lidar (HDL), which can facilitate better landmark feature extractions, thus reducing the pose errors. Acceleration and vibration of the robot can be the other major sources of laser measurements errors. Higher perception rate or lowering the speed of the robot or a 'stop and go' motion, can help to minimize such errors.

For resolving the remaining vertical translation (6^{th} DoF), an additional digital camera can be used. The typical drift error observed in the scenarios with noisy inputs can be addressed by using a higher value for $\delta > \mathbf{1}$. That means, the Data Association is performed not between the immediate two sets of data (i.e. $\delta = \mathbf{1}$) but with the data observed with a certain delay. This increased delay ensures a higher signal to noise ratio, which when averaged by the used time delay, acts as a filter, helping to reduce the noise, and therefore, the drift. To use a higher δ value, landmarks need to remain observable for a longer period. The two isoclinal pairs of laser scanner setup, pointing forward and backward direction of the vehicle achieves this (refer [6]). The other major obstacle for this approach is the absence of non parallel planar landmarks in the observing scene. The additional camera or a 3D laser scanner with ability to process landmarks other than the planes can be very useful. Even the empty spaces (where there is no detectable objects present) can be used as landmark patterns.

VIII. ACKNOWLEDGMENTS

The authors gratefully acknowledge the Center of Robotics (CAOR) from Mines ParisTech for providing the platform and resources to realize this work, and the reviewers' comments.

REFERENCES

[1] M.W.M.G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte, & M. Csorba. "A solution to the simultaneous localization and map building (SLAM) problem", *IEEE Transactions on Robotics and Automation*, vol. 17, no. 3, pp 229-241, 2001.

[2] M. Montemerlo, S. Thrun. "Simultaneous localization and mapping with unknown data association using FastSLAM", *In Proceedings of IEEE International Conference on Robotics and automation*, 2003, vol. 2.

[3] J. Weingarten, R. Siegwart "EKF-based 3D SLAM for structured environment reconstruction", *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005). Proceedings*, 2005, pp 2089 2094.

[4] F. Lu & E. Milios, "'Robot pose estimation in unknown environments by matching 2d range scans", *Journal of Intelligent and Robotic Systems*, vol. 18, no. 3, pp 249-275, 1997.

[5] B. Steux & O. El Hamzaoui, "CoreSLAM: a SLAM Algorithm in less than 200 lines of C code", *Proceedings of the 11th Control Automation Robotics & Vision (ICARCV)*, pp 1975–1979, 2010.

[6] K. Narayana, F. Goulette & B. Steux, "Planar Landmark Detection Using a Specific Arrangement of LIDAR Scanners", *Proceedings of the Position Location and Navigation Symposium 2010*, USA, 2010.

[7] T. Bailey, "Mobile robot localisation and mapping in extensive outdoor environments", *PhD thesis, Australian Center for Field Robotics, University of Sydney*, 2002.

[8] M.A. Fischler & R.C. Bolles. "Random sample consensus: A paradigm for model
tting with applications to image analysis and automated cartography", *Communications of the ACM*, vol. 24, no. 6, pp 381-395, 1981.