Fast Keypoint Features From Laser Scanner for Robot Localization and Mapping

Fast Keypoint Features from Laser Scanner for Robot Localization and Mapping

Fabjan Kallasi[1] and Dario Lodi Rizzini[1] and Stefano Caselli[1]

***Abstract*—Detecting features in sensor measurements and distinguishing among them is an important capability for robot localization and navigation. Despite the wide diffusion of range finders, there are few works on keypoint features for 2D LIDAR and there is potential for improvement over the existing methods. This paper proposes two novel keypoint detectors for the stable detection of interest points in laser measurements and two descriptors for robust associations. The features defined by combining keypoints and descriptors allow stable and efficient place recognition. Experiments with standard benchmark datasets assess the performance of the detectors and descriptors investigated. One of the proposed features, termed FALKO-BSC, achieves higher repeatability score and similar descriptor performance compared with the FLIRT state-of-the-art feature. FALKO-BSC is also shown to enable effective localization.**

***Index Terms*—Range Sensing; Localization; Mapping**

I. INTRODUCTION

RECOGNIZING distinctive regions, local patches and points in sensor data is a fundamental capability for many tasks including robot localization, mapping, registration, scene description, and object detection. In computer vision and point cloud processing, the feature approach [1]–[3] has become an established paradigm to address these tasks. The goal of the feature approach is the detection of interest points, also known as *keypoints*, from the sensor data and the computation of a distinctive signature for each them. The term keypoint feature or, briefly, feature refers to an interest point with the corresponding descriptor. Invariance from viewpoints and repeatability are the main requirements of keypoint detection. The standard signature of an interest point has the form of a vector of values (usually dependent from the point neighborhood) called *descriptor*. Descriptors enable more robust correspondence matching among points acquired from different viewpoints than simple geometric matching.

Several keypoint features have been proposed for images and point clouds, but only few keypoint features exist for sensor data acquired from planar range finders [4]–[6]. While images and point clouds provide quantitatively and qualitatively rich information, laser scans consist of relatively few points, which convey only geometrical data. The peculiar nature of laser measurements makes the extraction of stable and distinguishable interest points more difficult. Apart from standard geometric features like segments, there are few works related to 2D range keypoint features; these can be classified into two categories. The first category includes methods that convert the laser scan into an image and compute keypoints using an existing computer vision method [5], [6]. This approach enables the re-use of many algorithms proposed in computer vision, but it is computationally expensive, requires rasterization of laser measurements, and suffers from inconsistencies due to obstacle projection in an image. The second category includes keypoint features explicitly designed for the laser scan peculiarities. Indeed, the scan radial order is taken into account in all the operations involving neighbor points, e.g. smoothing, downsampling, curvature computation. This second category includes FLIRT [4], which detects interest points according to properties such as range-differences or curvature and associates a descriptor to them.

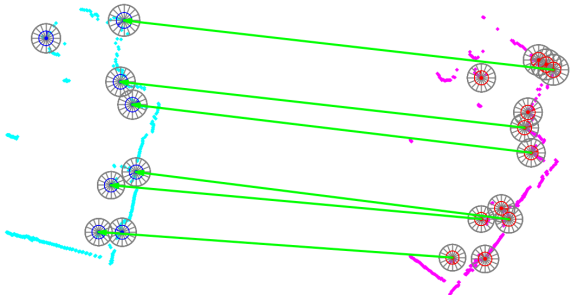

Fig. 1: Example of feature association with FALKO detector and BSC descriptors.

In this paper, we propose two novel keypoint detectors, FALKO (*Fast Adaptive Laser Keypoint Orientation-invariant*) and *Orthogonal Corner* (OC), and two novel descriptors, *Binary Shape Context* (BSC) and *Cumulative Gaussian Histogram* (CGH). Like with the FLIRT curvature detector, the two proposed detectors extract high curvature points in the laser scans. However, FALKO and OC have been designed to detect more stable points like corners rather than general elements like gaps and isolated points that are dependent on viewpoint. FALKO selects interest points through a careful neighborhood selection and an effective cornerness scoring. OC exploits the orthogonal alignment, which often holds in artificial environments. *Binary Shape Context* (BSC) is a variant of the shape context [4] that is less sensitive to point concentration (Figure 1). The *Cumulative Gaussian Histogram* (CGH) counts the polar distribution of neighbor points w.r.t. the main orientation of the interest points. FALKO and OC have been compared with the state-of-the-art approach FLIRT

Manuscript received: August, 31, 2015; Revised October 29, 2015; Accepted December 26, 2015.

This paper was recommended for publication by Editor Cyrill Stachniss upon evaluation of the Associate Editor and Reviewers' comments.

[1]Authors are with RIMLab - Robotics and Intelligent Machines Laboratory, Dipartimento di Ingegneria dell'Informazione, University of Parma, Italy, {kallasi,dlr,caselli}@ce.unipr.it

on standard benchmark datasets used in the robotic community to assess feature stability and localization performance.

II. RELATED WORK

During the last two decades, several keypoint features have been proposed for both images and point clouds. SIFT [7] is often considered the first example and the prototype of keypoint features with descriptor. SIFT achieves scale invariance through point extraction at different scales and distinguishes points using descriptors. In later years, several other keypoint methods in computer vision have been proposed [1], [2]. The keypoint feature paradigm has also been successfully applied in the context of 3D perception and shape recognition [3]. Extraction of significant local features from planar range finder data is more difficult due to the limited number of points in a scan compared to an image, and the peculiar distribution of neighbor points. Hence, the occupancy grid approach is prevailing in works on laser scan processing [8], [9], and elementary geometric features like segments [10], curvature points [11], and B-splines [12] have been used in association with laser scanners.

One of the first attempts to define laser keypoints with descriptor is illustrated in [13], where an orientation histogram obtained from the local submap is associated to each selected high-curvature point. This approach is improved in a subsequent work [14], but the global descriptors still represent a complete submap. The only true examples of keypoint features with descriptors for laser scans are proposed in [4]–[6]. Li and Olson [5] propose to rasterize laser scans and to extract interest points from the obtained image using Kanade-Tomasi corner detector. The SIFT descriptor is then computed on the neighborhood of each point. The method is improved in the successive work [6] through scan smoothing pretreatment, more efficient corner extraction and candidate suppression. The main disadvantage of this approach lies in the inaccuracies resulting from rasterization.

Tipaldi and Arras [4] proposed FLIRT, the first keypoint feature explicity designed for planar laser scans. The input laser scan is rescaled several times and keypoints are detected at different scales according to parameters like curvature, normal and range. Next, the local shape is encoded by two descriptors, one representing the polar points distribution and the other representing occupancy. While FLIRT is a practical and conceptual improvement over the previous approaches, the detection is based on a single parameter, which may result into the identification of multiple and potentially unstable interest points. For example, it can find keypoints where discontinuities due to specific viewpoint occlusions occur, and also in unstable isolated points. Furthermore, the multiscale approach, which reduces the dependency from neighbor point distribution due to angular resolution and distance, is computationally expensive and not always effective, since it is performed only through local smoothing without downsampling.

III. KEYPOINT DETECTORS

In this section, we illustrate two novel keypoint detectors: OC, which exploits orthogonal world hypothesis, and FALKO, which extracts stable points through invariance from scan density and orientation.

A. Orthogonal Corner Detector

Human-made indoor environments often consist of straight linear walls and architectural elements arranged along orthogonal directions. Laser scans obtained in such environments are composed of several points aligned along two orthogonal dominant directions. This condition can be exploited to find stable keypoints at the intersection of orthogonal line pairs.

To find the dominant direction of the laser scan points, the algorithm computes the *Hough Spectrum* [15] of such points. Let $\mathcal{S}$ be the point set from an input laser scan and $[\theta, \rho]$ be the Hessian parameters of a line $p_x \ \cos\theta + p_y \ \sin\theta = \rho$. Let us define a subdivision of parameter space into $n_\theta \times n_\rho$ cells centered in $[\theta_t, \rho_r]$, where $\theta_t = t \ \Delta\theta$ with $\Delta\theta = \pi/n_\theta$ and $\rho_r = \Delta\rho(r - n_\rho/2)$ with $r = 0, \ldots, n_\theta - 1$ and $t = 0, \ldots, n_\rho - 1$. The Hough Transform $HT(\theta_t, \rho_r)$ is a bidimensional histogram that associates to each bin cell centered on (θ_t, ρ_r) the number of scan points $p \in \mathcal{S}$ lying on the line with parameters (θ_t, ρ_r). The Hough Spectrum $HS(\theta)$ of $\mathcal{S}$ is defined as $HS(\theta_t) = \sum_{\rho_r} HT(\theta_t, \rho_r)^2$. The value of $HS(\theta_t)$ depends on the number of points aligned to a line belonging to the pencil of parallel lines with parameter θ_t. To consider the contribution of two orthogonal directions, the *Orthogonal Hough Spectrum* can be defined as $OHS(\theta_t) = HS(\theta_t) + HS(\theta_{t+n_\theta/2})$. The dominant direction $\bar{\theta}$ is found as the absolute maximum of $OHS(\cdot)$.

Once the dominant direction $\bar{\theta}$ is found, all the scan points are rotated by $-\bar{\theta}$ (call $\bar{\mathcal{S}}$ the set of rotated points) and tested as candidate keypoints. The neighborhood of each scan point p_i (i index in the scan) has radius r_i increasing with its distance from the sensor origin, $\|p_i\|$, to have an approximatly uniform number of neighbors. In our implementation, the radius is computed as $r_i = a \exp(b\|p_i\|)$ where the parameters $a = 0.2$ and $b = 0.07$ are chosen s.t. $0.20\ m \lesssim r_i \lesssim 0.40\ m$ for ranges in $1 - 10\ m$. For each point p_i, the following sets have been defined

$$C(p_i) = \{p_j \in \bar{\mathcal{S}} : \|p_j - p_i\| < r_i\} \quad (1)$$

$$C_x(p_i) = \{p_j \in C(p_i) : |p_{j,x} - p_{i,x}| < w \wedge \ |p_{j,y} - p_{i,y}| > w\} \quad (2)$$

$$C_y(p_i) = \{p_j \in C(p_i) : |p_{j,y} - p_{i,y}| < w \wedge \ |p_{j,x} - p_{i,x}| > w\} \quad (3)$$

where w is the tolerance on point alignment. The neighbor points of p_i, constituting set $C(p_i)$, can be efficiently found by exploiting the point radial order and the regular angular resolution (index i is increasing counterclockwise). In particular, if $\Delta\alpha$ is the angular resolution, the points $p_j \in C(p_i)$ must have an index j s.t. $|j - i| < w_i$ where $w_i = \lfloor \arcsin(r_i/\|p_i\|)/\Delta\alpha \rfloor$.

A good corner is defined by points aligned on both directions, i.e., with high values of $n_x = |C_x(p_i)|$ and $n_y = |C_y(p_i)|$. A score is defined as

$$score(p_i) = \frac{n_x + n_y}{\epsilon + |n_x - n_y|} \quad (4)$$

The keypoints are chosen as the maxima of $score(p_i)$. In order to avoid points that are too close and ambiguous, a non-maxima suppression (NMS) procedure on $0.20\ m$ range is applied to select the most stable features.

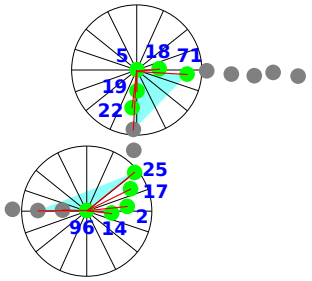

Fig. 2: An example of FALKO keypoint extraction with scores from eq. (14) (blue) of each candidate point (green), corresponding triangle approximation (light blue), and orientation vectors (red).

B. FALKO

Laser scan data consist in a set of points which approximate the perceived environment shape. In LIDAR with low angular resolution, the approximation is less accurate. Accurate detection of interest points from noisy data can be performed only after higher level inference than detection of simple geometric shapes like segments or circles. FALKO exploits the concept of edge intersection in 2D range data. This detector has been designed to be orientation invariant and point density independent.

Like for OC keypoints, the set of neighbor points $C(p_i)$ of a candidate scan keypoint p_i is defined by equation (1). The first step of the proposed method avoids evaluation of points which cannot be considered as candidate corners due to the lack of neighbors, geometric inconsistency or artifacts originated by range discontinuities. The point set $C(p_i)$ is divided in two subsets

$$C_L(p_i) = \{p_j \in C(p_i) : j < i\} \tag{5}$$
$$C_R(p_i) = \{p_j \in C(p_i) : j > i\} \tag{6}$$

If the cardinality $|C_L| < 2$ or $|C_R| < 2$, the point is discarded from the candidate corner set. The two endpoints x_L and x_R of each candidate point neighborhood are defined as

$$x_L = p_{j_{min}} : \quad j_{min} = \arg\min_j \{p_j \in C_L(p_i)\} \tag{7}$$
$$x_R = p_{j_{max}} : \quad j_{max} = \arg\max_j \{p_j \in C_R(p_i)\} \tag{8}$$

The triangle $\triangle p_i x_L x_R$ is then geometrically evaluated as a rough approximation of the corner. Let $\overline{x_L x_R}$ be the base of the triangle. If the base length $\|\overline{x_L x_R}\|$ or the triangle height is less than $\frac{r_i}{\beta}$ the point is discarded. Parameter β is chosen taking into account specific limitations on the aperture and subtended angle of the corner. In particular, greater values of β allow both wider and sharper corners to be selected as candidates. These conditions efficiently filter unsuitable corner candidates based on simple geometric properties. Then, for each candidate point a *cornerness* score is computed. A polar grid, which quantizes the space in circular sectors, is centered on the candidate point p_i. For each point $p_{j,L} \in C_L(p_i)$ and $p_{j,R} \in C_R(p_i)$, a quantized orientation w.r.t. the candidate point is computed as

$$\phi_{j,L} = \left\lfloor \frac{s_n}{2\pi} \tan^{-1}\left(\frac{p_{j,y} - p_{i,y}}{p_{j,x} - p_{i,x}}\right) \right\rfloor, \quad \forall p_j \in C_L(p_i) \tag{9}$$
$$\phi_{j,R} = \left\lfloor \frac{s_n}{2\pi} \tan^{-1}\left(\frac{p_{j,y} - p_{i,y}}{p_{j,x} - p_{i,x}}\right) \right\rfloor, \quad \forall p_j \in C_R(p_i) \tag{10}$$

where s_n is the number of circular sectors in the polar grid. Let

$$d_\theta(\phi_1, \phi_2) = \left((\phi_1 - \phi_2) + \frac{s_n}{2}\right) (\text{mod } s_n) - \frac{s_n}{2} \tag{11}$$

be a distance function between the quantized orientations in circular sector units. The score for a candidate point is defined as

$$score_L(p_i) = \sum_{h=i-1}^{j_{min}} \sum_{k=h-1}^{j_{min}} |d_\theta(\phi_h, \phi_k)| \tag{12}$$
$$score_R(p_i) = \sum_{h=i+1}^{j_{max}} \sum_{k=h+1}^{j_{max}} |d_\theta(\phi_h, \phi_k)| \tag{13}$$
$$score(p_i) = score_L(p_i) + score_R(p_i) \tag{14}$$

This score function measures the alignment of the two point sets, respectively C_L and C_R, and it is orientation invariant. For each set, the more points are aligned in the same *direction* the smaller is the score value. Figure 2 shows an example of score computation. Keypoints are then chosen as local minima (or maxima with score inversion) of the score function in (14) with a NMS procedure. Like for OC keypoints, NMS range is set to $0.20\ m$.

IV. KEYPOINT DESCRIPTORS

In this section, two novel descriptors are illustrated: the first is derived from shape context descriptor [16], while the second represents the relative orientation between each keypoint and its neighborhood.

A. Binary Shape Context

Shape context is a local descriptor which represents the points distribution in a linear-polar histogram. The bins of the histogram centered on the keypoint count the neighbor points lying inside the region corresponding to the bin. The main drawback of this approach is its dependency from point density. Indeed, due to the angular quantization of the laser scanner, the same shape detected from two different viewpoints may results in two different histogram distributions. Figure 3(a) shows an example of shape context descriptor for the same shape viewed from two near poses. The two descriptors exhibit different peaks in the linear-polar histogram which can cause mismatch in the recognition of the keypoint. We propose a binary version of the local shape descriptor which is less sensitive to neighbor points density than the original one. Formally, given r_n and α_n, respectively the number of radial and polar quantization in the descriptor grid, for each keypoint kp the *BSC* descriptor is defined as the grid

$$BSC_{mj} = \begin{cases} 1 & \text{if } \exists\ p_i : (p_i - kp) \in \overline{BSC_{mj}} \\ 0 & otherwise \end{cases} \tag{15}$$

where p_i is a point of kp neighborhood, $m = 1, \ldots, r_n$, $j = 1, \ldots, \alpha_n$ and $\overline{BSC_{mj}}$ is a cell in the linear-polar grid. A

distance function for the BSC descriptor can be defined as the Hamming-like distance function

$$d_{BSC}(BSC_1, BSC_2) = \sum_{m=1}^{r_n} \sum_{j=1}^{\alpha_n} BSC_{1,mj} \oplus BSC_{2,mj} \quad (16)$$

where $\oplus$ is the XOR operator. Figure 3(b) shows the resulting descriptors computed for the same keypoint viewed from different poses.

B. Cumulative Gaussian Histogram

Laser scan data provide only geometric information of the perceived environment. Thus, geometric shape or neighborhood point distribution are commonly used as interest point *description*. However, these approaches are viewpoint and density dependent. Therefore, we propose a relative orientation-based descriptor that exploits the viewpoint and density invariant property of orientation between points.
We compute a histogram based on the same radial grid presented in section III-B. Given α_n, the number of histogram bins, for each point p_i in the keypoint kp neighborhood, a relative quantized orientation ϕ_i is computed as

$$\phi_i = \left\lfloor \frac{\alpha_n}{2\pi} \ \tan^{-1}\left(\frac{p_{i,y} - kp_y}{p_{i,x} - kp_x}\right) \right\rfloor \quad (17)$$

Then, for each point a discrete Gaussian distribution $\mathcal{N}(\phi_i, \sigma)$ is constructed. The *CGH* is defined as

$$CGH_j = \sum_{\phi_i} \mathcal{N}(j - \phi_i, \sigma) \quad (18)$$

where $j = 1, \ldots, \alpha_n$ and CGH_j is a cell in the radial histogram. Figure 3(c) shows two examples of CGH descriptor. Histograms have the same radial distribution although the two scans are perceived from two different robot poses. A normalization of the histogram makes the descriptor less sensitive to neighbor points density, and a symmetric χ^2 test can be used to evaluate the distance between two CGH descriptors.

C. Computing Corner Orientation

Descriptor matching can be computed through the evaluation of a distance function. Since the same corner can be perceived from different view-angles, the resulting descriptors are not aligned w.r.t. the same start angle. Thus, for matching, descriptors are rotated according to the *corner orientation* estimated by the keypoint detector. For OC keypoints, the corner orientation is given by the bisector angle between two points aligned to the dominant direction $\bar{\theta}$ and its orthogonal direction $\bar{\theta}+\pi/2$. For FALKO keypoints, the corner orientation is computed with a variant of the *Intensity Centroid* [17]. First, for both sets C_L and C_R given by equations (5) and (6), the centroid is computed w.r.t. the interest point. The average of the two centroids is used to compute the orientation vector $\mathbf{o}$. The corner orientation is estimated as $atan2(\mathbf{o}_y, \mathbf{o}_x)$. This method results in good orientation estimation without point density dependency. Figure 3 shows some examples of corner orientation vectors which can be used to align the descriptors w.r.t. the same starting angle.

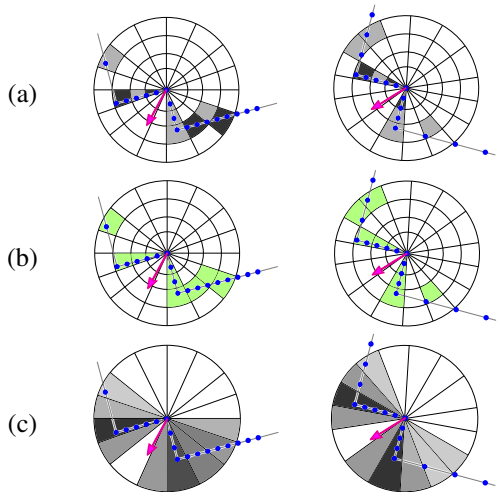

Fig. 3: Descriptor grids for example data with two different viewpoints and relative *corner orientation vector* (purple arrow): (a) Local Shape Context proposed in [4]; (b) *Binary Shape Context*; (c) *Cumulative Gaussian Histogram*. Darker grey level represents higher value in the histogram bin.

V. EXPERIMENTS

The aim of the experiments presented in this section is the evaluation and comparison of the proposed keypoint detectors and descriptors w.r.t. the state of the art. The most important property of keypoints is their stability, i.e., the detection of the same point in the environment after changing viewpoint and regardless of laser scanner properties like noise level and resolution. On the other hand, the performance of descriptors depends on their capability to distinguish among different places of the environment. In order to make the comparison reproducible, we adopted the criteria employed in the experimental assessment of [4], which in turn follows the approach of Mikolajczyk et al. [1], [18]. Thus, we compared features by changing the viewpoint, the noise level and the number of measurements of a scan. Furthermore, global localization and loop closure capabilities have been exploited comparing the proposed features to FLIRT.

A. Experimental Setup

The proposed keypoint detectors and descriptors are assessed on five datasets: *fr-079*, *fr-clinic*, *intel*, *mit-csail* and *victoria-park*. All log files are freely available from Radish and other online repositories[1]. However, the assessment about features requires a groundtruth where all scan viewpoints are referred to a single global reference frame. We used the datasets with the registered scans provided with [4], which consists of both the original scans and the corresponding adjusted pose graph. For each test, a total of 100 random scans, 20 from each dataset, have been evaluated. Results are then averaged over the 100 samples.

The scans in each dataset are used to assess the performance of both the proposed keypoint detectors and the descriptors. The experimental assessment is based on the comparison

[1]For example, http://kaspar.informatik.uni-freiburg.de/~slamEvaluation/datasets.php

between two similar features sets. The stability of keypoints has been measured in two ways. In the first one, the keypoints have been extracted from all the scans of the datasets and associated together according to the procedure described in [6]. The stable interest points are detected several times. The covariance matrices of the associated keypoints provide a measure of their stability.

The second metric for keypoint stability is the *repeatability* of the points detected from two similar scans. In particular, the sets of keypoint $\mathcal{P}_R$ and $\mathcal{P}_S$ are detected respectively from a given reference scan and from another similar scan. The keypoints appearing in both sets can be found by associating the points, e.g., according to nearest neighbor criterion. Like in [18], the repeatability is defined as the percentage of common keypoints over the smaller of the two keypoint sets, i.e. $|\mathcal{P}_R \cap \mathcal{P}_S| / \min\{|\mathcal{P}_R|, |\mathcal{P}_S|\}$. In our experiments we performed the four different tests proposed in [4] that assess the feature robustness to changing conditions. A given scan is compared with another scan obtained either by changing viewpoint or by transforming the original scan.

1) *Viewpoint*. Each reference scan $\mathcal{S}_{Ri}$ of the dataset is compared with all the scans $\mathcal{S}_{Sj}$ with a viewpose in $1\ m$ range. Then, the set of close scan pairs is partitioned according to their level of similarity. Since the points of a scan geometrically delimit a free region around the laser, the similarity of two scans $\mathcal{S}_{Ri}$ and $\mathcal{S}_{Sj}$ is measured by their *overlap*, which is the ratio between the area of intersection $\mathcal{S}_{Ri} \cap \mathcal{S}_{Si}$ and the area of their union $\mathcal{S}_{Ri} \cup \mathcal{S}_{Sj}$.
2) *Noise*. Each scan $\mathcal{S}_{Sj}$ is obtained by adding increasing level of gaussian noise to the range measurement of the original scan $\mathcal{S}_{Ri}$. The standard deviation of the noise (measured in meters) represents the level of noise.
3) *Oversampling*. The scan pairs used in this test are obtained by adding new points to the original scan through linear interpolation. The oversampling level corresponds to the number of points obtained by interpolating two adjacent points.
4) *Subsampling*. The scan pairs used in this test are obtained by removing points from each scan of the dataset. The subsampling level corresponds to the number of points removed between two remaining points.

Descriptor performance is assessed using the same scan pairs used in the four keypoint detector tests illustrated above. Each pair of scans has been processed as follows. The keypoints sets $\mathcal{P}_R$ and $\mathcal{P}_S$ are detected respectively from the reference scan $\mathcal{S}_R$ and the similar scan $\mathcal{S}_S$ as in the tests discussed above and the corresponding sets of descriptors, respectively $\mathcal{D}_R$ and $\mathcal{D}_S$, are computed. Next, the sets $\mathcal{D}_R$ and $\mathcal{D}_S$ have been associated measuring the distances between descriptors. Two descriptors match if their distance is below a given threshold. Several values of threshold have been used in order to compute an evaluation through precision-recall curves. The ground truth for these tests is obtained with a geometric association of the keypoint used for descriptors computation.

Dataset	Keypoint	Num. Points	$\bar{\lambda}_{mean}$	$\bar{\lambda}_{max}$	Single Points
fr079	FALKO	**1127**	0.027	0.039	**21.9**%
	OC	698	**0.022**	**0.032**	30.2%
	FLIRT	1094	0.026	0.037	24.8%
fr-clinic	FALKO	**39077**	0.028	0.040	**49.5**%
	OC	1804	**0.021**	**0.033**	59.1%
	FLIRT	19314	0.024	0.035	52.6%
intel	FALKO	1153	0.028	0.039	**24.8**%
	OC	405	**0.019**	**0.028**	33.1%
	FLIRT	**1155**	0.025	0.035	26.2%
mit-csail	FALKO	**1463**	0.025	0.037	**33.5**%
	OC	716	**0.021**	**0.032**	48.5%
	FLIRT	1446	0.024	0.035	34.4%
victoria-park	FALKO	**9764**	0.027	0.039	**61.3**%
	OC	760	**0.021**	**0.036**	72.5%
	FLIRT	7120	0.026	0.038	70.5%

TABLE I: Average of the geometric mean and maximum of the keypoint covariance matrix square-rooted eigenvalues. Best values are in bold font.

B. Keypoint Detector Assessment

Evaluation of keypoints repeatability has been performed with default parameters for each dataset. In particular, the parameters used for FLIRT are those indicated in [4]. For OC, we set $n_\theta = 360$, $\Delta\rho = 0.05$, $n_\rho = 1200$ and $w = 0.04$. In FALKO, we set $\beta = 4.0$ and $s_n = 16$. These parameter values have been used with both indoor and outdoor environments for a fair comparison among methods. However, to achieve better performance a fine tuning of the parameters is recommended. Figure 4 shows the average results of repeatability tests. In each test, FALKO outperforms both FLIRT and OC, except for high noise deviation levels where, anyway, the scan data are no longer reliable (above $0.3\ m$). In outdoor datasets, where the scan is more scattered and there are stronger biolations of the orthogonal hypothesis, OC results in poor performance, which drops its average repeatability value over the five datasets.

Table I summarizes the results of the global stability tests of keypoint detectors for all the datasets. As in [6], a global map of keypoints is built for each dataset by adding a new landmark, when a new point is distant at least $d_{min} = 0.2\ m$ from any landmark in current map, and by associating a point to the closest landmark l in the range $d_{max} < 0.05\ m$. The covariance matrix Σ_l is computed using the k_l keypoints associated to the map landmark l, when $k_l > 1$. If $k_l = 1$, l is computed from a single point and is, therefore, an ephemeral landmark. The maximum $\lambda_{max,l}$ and the geometric mean $\lambda_{mean,l}$ of the eigenvalues of Σ_l are used to measure the uncertainty of l. The average of maximum eigenvalues $\bar{\lambda}_{max}$ and of eigenvalue geometric means $\bar{\lambda}_{mean}$ are reported in Table I with the percentage of landmarks obtained from single points. As can be seen, OC has the lowest values of $\bar{\lambda}_{max}$ and $\bar{\lambda}_{mean}$ in all the datasets, though with the higher percentage of ephemeral points and lower number of detected keypoints. FALKO obtains slightly higher values of $\bar{\lambda}_{max}$ and $\bar{\lambda}_{mean}$ than FLIRT, but it tends to detect more stable points as shown by the lowest percentage of single points, i.e. ephemeral landmarks.

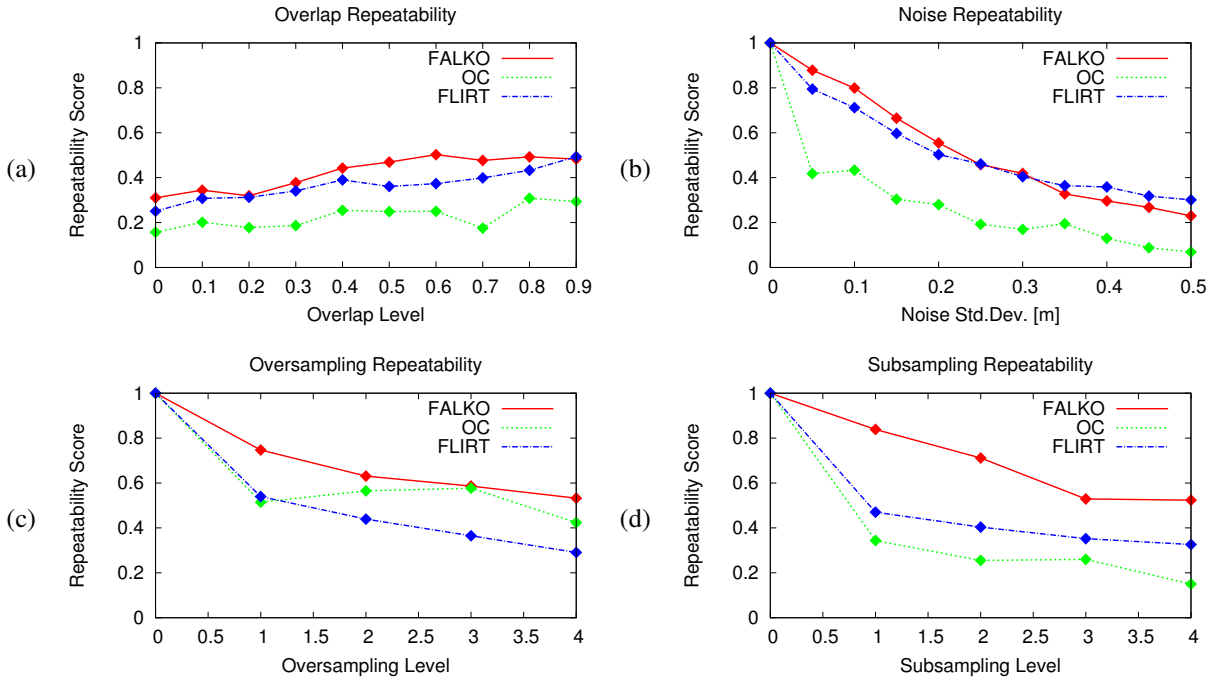

Fig. 4: Repeatability values of keypoint detectors FALKO, OC and FLIRT in four tests: (a) viewpoint with different scan overlap value; (b) noise level with different noise variance; (c) oversampling of scan measurement; (d) subsampling of scan measurement.

Keypoint	Time [μs]
FLIRT	23469
OC	5362
FALKO	567

Descriptor	Time [μs]
β-grid	1106
BSC	34
CGH	27

TABLE II: Average computation time of keypoints and descriptors

C. Descriptor Assessment

Like in section V-B, descriptors have been evaluated with the same parameters for all datasets. *β-grid* was set with the default parameters from [4]. For BSC and CGH descriptors we set $r_n = 8$ (only for BSC), $\alpha_n = 16$ and a neighbors radius equal to $0.5\ m$. For CGH we set standard deviation $\sigma = 0.6$. Figure 5 shows the 1-Precision Recall curves computed for the four scan transformations. The results of a pure descriptor-association are poor for all the evaluated descriptors due to the intrinsic lack of information in laser scan data, as shown also in [1], [4]. A pure descriptor-association without position or geometry information is not recommended for pose estimation purpose. Descriptors can be used as a gating rule in geometric associations like Nearest-Neighbors, Hungarian algorithm or RANSAC. BSC and CGH perform almost the same or even slightly better as β-grid in each test, while requiring much less computation. Indeed, table II shows the registered computation time. Performance is computed over 1000 samples of the same dataset (fr079) on an Intel Core i7-4770 CPU @ 3.40GHz, 8GB RAM. Results show that FALKO, BSC and CGH significantly outperform the other methods in computation performance. OC performs better than FLIRT, but is 10 times slower than FALKO due to rotation alignment. Computational times of FALKO and OC are compatible with real-time execution.

D. Data Association with Keypoints

Estimation of robot pose given a single measurement and a map is an important capability for global localization and loop closure in SLAM. The experiments illustrated in this section are designed to assess the contribution of keypoint features to the execution of these tasks. The tests have been performed on four datasets, three indoor (*mit-csail*, *fr079*, *intel*) and one outdoor (*fr-clinic*), using the three keypoint detectors (FALKO, OC, FLIRT) combined with pure geometric matching (Geom) or with descriptor (BSC, CGH, β-grid). Each scan $\mathcal{S}_R$ of a dataset has been compared to the other scans $\mathcal{S}_i \neq \mathcal{S}_R$ of the the same dataset and aligned to the best matching one among the $\bar{\mathcal{S}}_i$. The candidate scan is estimated as follows. The keypoints detected from each scan are used to compute the GLARE signature of the scan [19]. The initial set of candidates consists of the 10 scans $\mathcal{S}_i$, whose GLARE signatures are closer to the GLARE signature of $\mathcal{S}_R$ according to the L1-norm. The keypoints of each $\mathcal{S}_i$ and $\mathcal{S}_R$ are the associated using maximum common subgraph data association [20] and are used to compute the robot pose in least-squares sense. The selected scan $\bar{\mathcal{S}}_i$ is the scan that, after the alignment, has the greatest number of keypoints with a neighbor point in $\mathcal{S}_R$ within the $0.10\ m$ range. If a descriptor is used, only the matching points with descriptor distance above a threshold are accepted. The robot is considered localized if the associated points are at least N_{min}.

Figure 6 shows the 1-precision-recall curves w.r.t. N_{min} for the different datasets and the different keypoint features. A localization is considered correct when the position error of the aligned scan is less than $0.50\ m$ and the angular error less than 10°. The curves show that FALKO with geometric matching

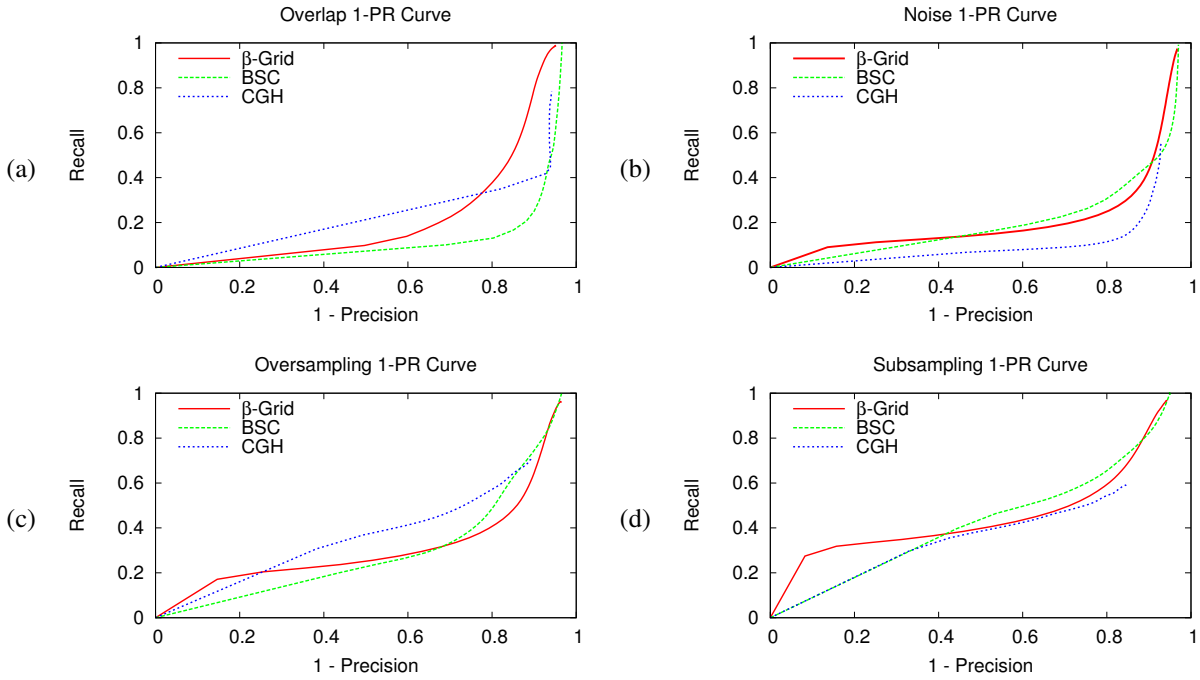

Fig. 5: 1-Precision Recall curves for the descriptors β-grid, BSC and CGH in four tests: (a) viewpoint (scan overlap value in interval $50\% - 90\%$); (b) noise level; (c) oversampling of scan measurements; (d) subsampling of scan measurements.

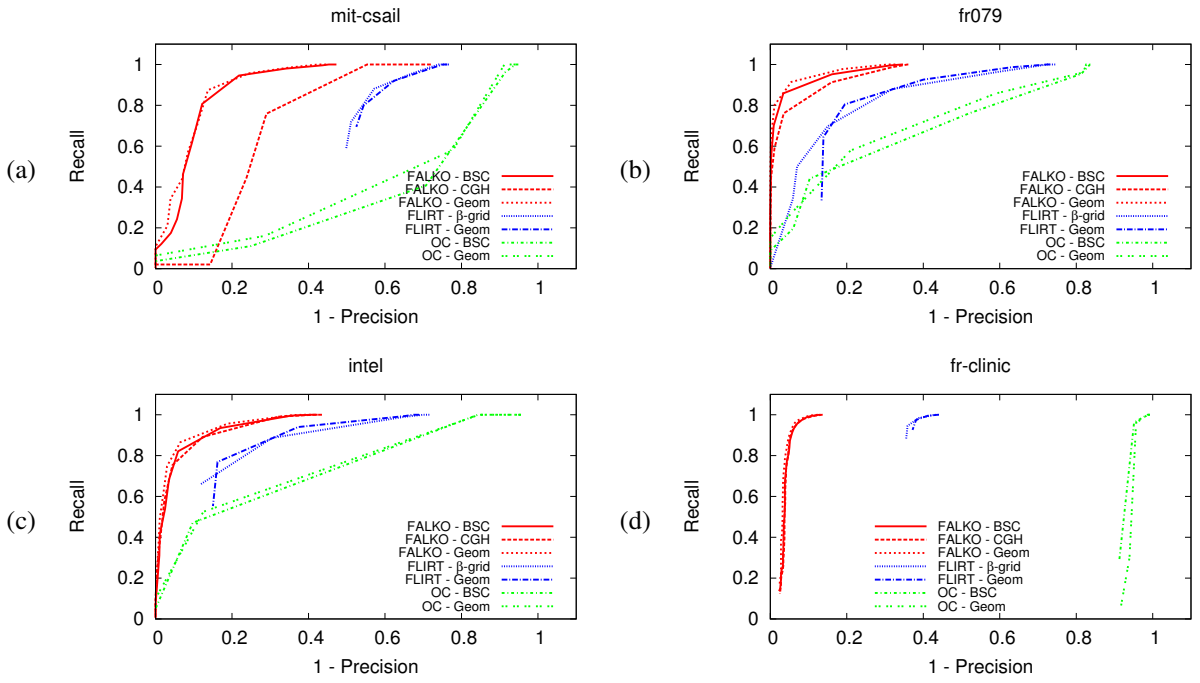

Fig. 6: 1-Precision Recall curves of global localizazion with FALKO-{BSC, CGH, Geom}, FLIRT-{β-grid, Geom} and OC-{BSC, Geom} for datasets (a) mit-csail, (b) fr079, (c) intel, and (d) fr-clinic.

or descriptor validation dominates those of the other detectors. The performance of FALKO-Geom and FALKO-BSC are comparable, while FALKO-CGH has lower performance, in particular in mit-csail. The proposed global localization method strongly relies on the stability and robustness of the keypoint detector, which is the main peculiarity of FALKO.

Figure 7 shows the values of recall, precision, position and angular errors at a break even point. The average position and angular errors are computed only for correctly localized trials. The localization error is less than $10\ cm$ and almost always less than 1° for all features for correct localization. FALKO outperforms almost always the other methods with a

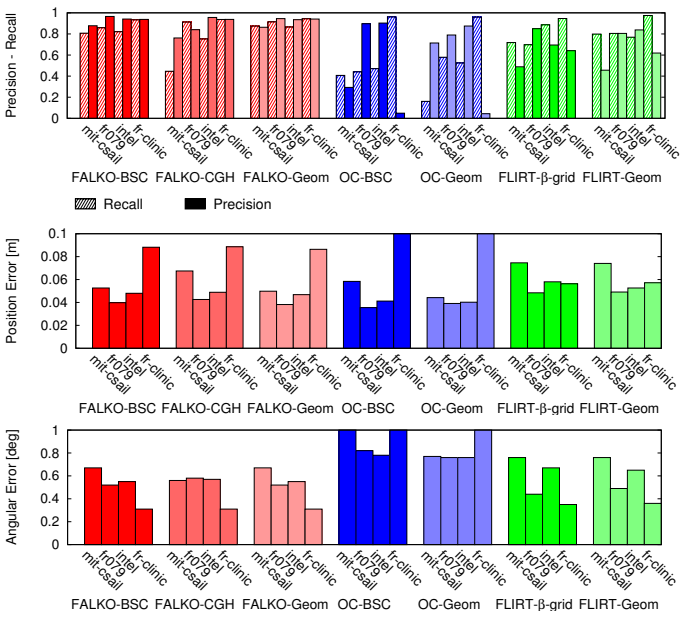

Fig. 7: Values of recall, precision, and average position and angular errors of the tested features for each dataset.

high reliability of the association. In the outdoor dataset both FALKO and OC obtain higher positional error even though FLIRT obtains much lower precision in data association and place recognition. The lowest position errors are achieved by OC, but these values are computed on less corrected associations.

VI. CONCLUSIONS

In this paper, we have proposed two keypoint detectors, OC and FALKO, with two novel descriptors designed to detect features from range finder measurements, BSC and CGH. Like other state-of-the-art detectors, OC and FALKO are conceived to find stable high curvature points in a laser scan, and to be invariant to sensor viewpoint and point density. OC and FALKO satisfy these properties through an efficient evaluation of the neighbor point distribution instead of relying on a computationally expensive multi-scale approach. The time required to process a laser scan with OC and FALKO is one or two orders of magnitude less than other available keypoint detectors. The proposed descriptors BSC and CGH provide an equally efficient signature for the keypoints. The performance of the proposed algorithms has been tested on widely used benchmark datasets and compared with FLIRT, the state-of-the-art feature for LIDARs. The results about detection show that FALKO achieves higher repeatably score and extracts less ephemeral points than the other keypoint detectors. Moreover, the precision-recall curves of the proposed detectors are consistent with the achievable results obtained from computer vision and laser scan data descriptors. The results also show that geometric association with FALKO performs better than other methods for global localization and loop closure purpose. Using BSC as a final gating rule results in slightly better precision in data association.

REFERENCES

[1] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.

[2] S. Krig, "Interest point detector and feature descriptor survey," in *Computer Vision Metrics*. Springer, 2014, pp. 217–282.

[3] Y. Guo, M. Bennamoun, F. Sohel, M. Lu, and J. Wan, "3D Object Recognition in Cluttered Scenes with Local Surface Features: A Survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2270–2287, Nov 2014.

[4] G. D. Tipaldi and K. O. Arras, "Flirt-interest regions for 2d range data," in *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2010, pp. 3616–3622.

[5] Y. Li and E. Olson, "Extracting general-purpose features from LIDAR data," in *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2010.

[6] ——, "Structure tensors for general purpose lidar feature extraction," in *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2011, pp. 1869–1874.

[7] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[8] G. Grisetti, C. Stachniss, and W. Burgard, "Improved Techniques for Grid Mapping with Rao-Blackwellized Particle Filters," *IEEE Trans. on Robotics*, vol. 23, no. 1, pp. 34–46, 2007.

[9] D. Lodi Rizzini and S. Caselli, "Metric-topological maps from laser scans adjusted with incremental tree network optimizer," *Robotics & Autonomous Systems*, vol. 57, no. 10, pp. 1036 – 1041, 2009.

[10] V. Nguyen, A. Martinelli, N. Tomatis, and R. Siegwart, "A comparison of line extraction algorithms using 2d laser rangefinder for indoor mobile robotics," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2005, pp. 1929–1934.

[11] P. Nunez, R. Vazquez-Martin, J. Del Toro, A. Bandera, and F. Sandoval, "Feature extraction from laser scan data based on curvature estimation for mobile robotics," in *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2006, pp. 1167–1172.

[12] L. Pedraza, D. Rodriguez-Losada, F. Matía, G. Dissanayake, and J. Miró, "Extending the limits of feature-based SLAM with B-splines," *IEEE Trans. on Robotics*, vol. 25, no. 2, pp. 353–366, 2009.

[13] M. Bosse and R. Zlot, "Map Matching and Data Association for Large-Scale Two-dimensional Laser Scan-based SLAM," *Int. Journal of Robotics Research*, vol. 27, no. 6, pp. 667–691, Jun 2008.

[14] ——, "Keypoint design and evaluation for place recognition in 2D LIDAR maps," *Robotics & Autonomous Systems*, vol. 57, no. 12, pp. 1211–1224, 2009.

[15] A. Censi, L. Iocchi, and G. Grisetti, "Scan Matching in the Hough Domain," in *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2005.

[16] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, 2002.

[17] P. L. Rosin, "Measuring corner properties," *Computer Vision and Image Understanding*, vol. 73, no. 2, pp. 291–307, Feb. 1999.

[18] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Gool, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1-2, pp. 43–72, 2005.

[19] M. Himstedt, J. Frost, S. Hellbach, H.-J. Boehme, and E. Maehle, "Large scale place recognition in 2D lidar scans using geometrical landmark relations," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2014, pp. 5030–5035.

[20] T. Bailey, E. Nebot, J. Rosenblatt, and H. Durrant-Whyte, "Data association for mobile robot navigation: a graph theoretic approach," in *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2000, pp. 2512–2517.