FSCD and BASD: Robust Landmark Detection and Description on Radar-Based Grids

Matthias Rapp and Klaus Dietmayer
Institute of Measurement, Control and Microtechnology
Ulm University, Germany
Email: {matthias.rapp, klaus.dietmayer} at uni-ulm.de

Markus Hahn, Frank Schuster,
Jakob Lombacher and Jürgen Dickmann
Daimler AG, Germany
Email: {markus.hahn, frank.fs.schuster,
jakob.lombacher, juergen.dickmann} at daimler.com

***Abstract*—This paper presents a detector and descriptor combination for robust landmarks on grids from radar measurements. Landmarks are fundamental for localization, a task of great importance in the field of mobile robotics and essential for autonomous driving. A fast detector is proposed which uses a rotational invariant pattern to locate scattering centers. These scattering centers occur as local maxima on a measurement grid, where detections from radar sensors for each cell are incremented. For association, a binary version of a descriptor designed especially for radar data is used. Experiments show that for radar data, the proposed combination improves performance compared to state-of-the-art algorithms.**

I. Introduction

New generations of radar provide an increased level of accuracy and number of detections. Because radar is a substantial part of the modern vehicle sensor setup, it is moving into the focus of research such as tracking, ego-motion estimation and localization. Beside *Simultaneous Localization and Mapping* (SLAM) the field of localization can be divided into two major categories. Grid-based localization uses a grid as reference for localization. This approach performs well on small environments. But the storage of the grid increases immensely if the considered environment gets larger. Landmark-based localization approaches are desirable as the memory requirement is greatly reduced. These algorithms such as EKF-SLAM, FAST-SLAM and GraphSLAM have been proven to provide advanced localization robustness and accuracy for mobile robotics [1]. However, the landmark-based localization result is sensitive to the input data, as landmarks of bad quality may lead to inaccurate localization performance or may end up in high localization error due to false landmark association.

For laser-based measurements, the selection of appropriate landmarks has already been investigated in the past [2]. For cameras, Shi et al. [3] propose visual landmarks which are shown to be appropriate for camera-based SLAM [4].

The contribution of this paper is a novel combination of landmark detection and description. The detection of landmarks based on radar data uses the idea of intensity comparison to locate scattering centers which allows a robust selection of landmarks. For the association task of landmarks, this paper presents a binary descriptor which describes the surrounding by comparison of statistic measures. This descriptor is designed for radar-based grids and shows competitive performance compared to state-of-the-art algorithms.

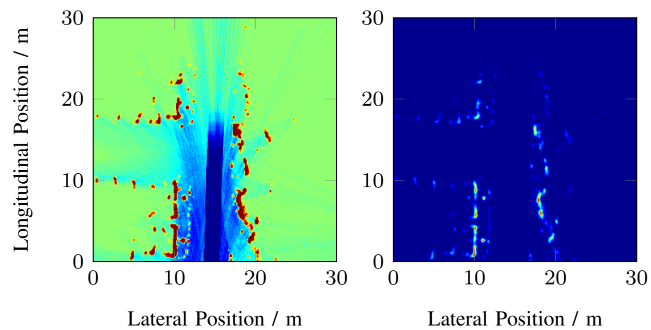

Fig. 1. This figure illustrates the difference between the occupancy grid (left) and the measurement grid (right). The measurement grid features a less saturated behavior compared to the occupancy grid. The red color indicates a high occupancy respectively many detections per cell on the measurement grid, whereas cells with low value are colored in blue.

The paper is structured as follows: Section II introduces the principles of grid representation using radar measurements. The detector for robust landmarks is presented in Section III. These landmarks are described by a binary descriptor in Section IV. In Section V, experiments are carried out to show the performance on real world data. The most important points of this paper are summed up in the conclusions in Section VI.

II. Grid Representation of Radar Measurements

An intuitive representation of range-based measurements is given by a grid where the surrounding environment of the robot is quantized into equally-sized cells. The idea is to transform the detections into a global coordinate system using the mounting position of the sensor and the odometry of the moving robot.

A. Occupancy Grid

For a grid representation of the environment, occupancy grids are the most common representation of continuous range measurements [5], see Fig. 1. For radar data, the measurement model has to be adapted because radar can detect an object behind other occluding objects due to mirroring or multi-path effects [6].

B. Measurement Grid

The most robust landmark centers of radar measurements are the scattering centers of objects because the sensor

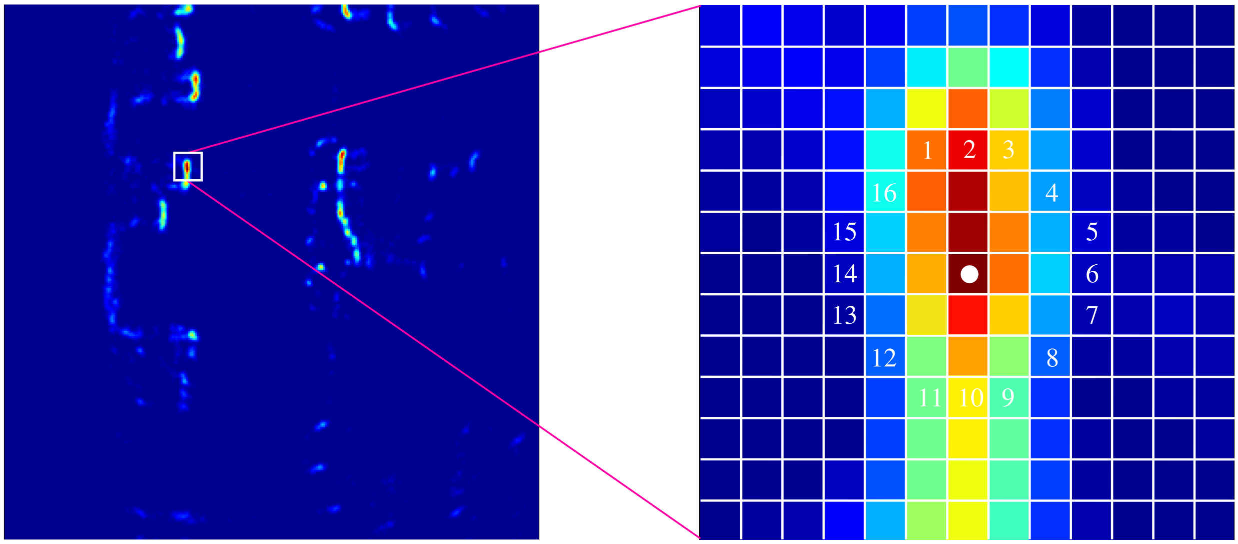

Fig. 2. This figure shows the pattern for landmark detection on the measurement grid (left). This radial pattern is depicted in the right figure, where the value of each cell of the pattern is compared to the center cell (white dot). If the value of all cells on the pattern is lower than the center cell, this cell identified as a landmark candidate.

receives multiple detections from them. In urban environments, highly reflective objects are typically street lights, metallic garbage bins, road signs and these are thus desired landmarks for localization. In order to detect these scattering centers, occupancy grids are not suitable due to the saturated behavior of the occupancy grid mapping. This subsection presents a mapping approach to locate scattering centers of objects. The idea is to increment the number of measurements for each cell. Experiments show that a smoothing factor is beneficial in order to compensate the quantization into cells. This mapping approach shows diminished saturated behavior than occupancy grid mapping, making this representation a qualified candidate for landmark detection (see Fig. 1). As cells with many detections result in a high value, this representation is feasible for identifying strong reflective objects. In the remaining sections this grid is used for landmark detection and description.

III. FSCD Detector

This section presents the algorithm for robust landmark detection on radar-based grids. The proposed detector operates on the measurement grid representation as introduced in Subsection II-B. The key idea is to locate the scattering centers. The scattering center is the position of an object from which a majority of measurements is obtained from different angles.

The proposed method for scattering center detection is inspired by the *Features from Accelerated Segment Test* (FAST)-detector [7]. The authors use a designed template for intensity comparisons for fast corner detection. In the following, a template is designed to locate local maxima of the measurement grid.

As in case of multiple drives through an environment, each observation may be conducted from different directions, the detector has to be invariant to rotation. This is the reason why a radial template as illustrated in Fig. 2 is used in which the cells of this pattern are denoted by c_j in the following. Using this template, a candidate cell $\widehat{c}$ of the measurement grid M (visualized by a white dot in Fig. 2) has to satisfy

$$|\{M(\widehat{c}) > M(c_j) : j = 1, \dots, 16\}| = 0 \qquad (1)$$

to be classified as a landmark, where $M(c_j)$ are the values evaluated for the pattern elements. This means that the candidate must have the highest measurement grid value with respect to the testing template. At this stage, even clutter may be detected. To select only strong reflecting cells, the candidate cell has to be higher than a specific lower boundary to be classified as a landmark. This is the only parameter of the proposed landmark detector. As there are only 16 comparisons in total, the implementation of this detector is very fast. A single loop over all grid cells is required and the computational cost is linear with only a very small constant. The detector observes many landmarks which are close to each other due to the pattern design. Because of the iterative looping through all cells, the landmarks are ordered. Therefore, a simple clustering with linear complexity is used in post-processing to combine these adjacent landmarks into a single landmark.

IV. BASD Descriptor

For association, the position information of each detected landmark is enriched by a feature vector containing features, i.e information of the landmark and the surroundings. This allows the robot to recognize primarily seen landmarks by determining the similarity of the feature vectors. In the field of image processing, descriptors are an established way of obtaining appropriate feature vectors. Prominent descriptors are *Fast REtinA Keypoints* (FREAK) [8], *Speeded-Up Robust Features* (SURF) [9] and *Scale-Invariant Feature*

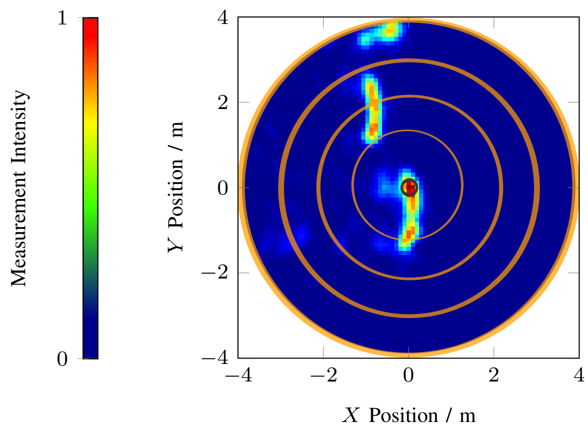

Fig. 3. This figure illustrates the basic radial partitioning of the surrounding environment of the *Binary Annular Statistics Descriptor* (BASD). This descriptor uses statistic measures for the elements of each ring. The feature vector consists of comparisons of these measures with regard to the outer rings.

Transform (SIFT) [10] which were used for many tasks such as image registration.

In this section, the *Annular Statistics Descriptor* (ASD) [11] for radar-based occupancy grids is extended to a binary version. In the field of image processing, binary descriptors are popular because they provide feature vectors which compare rapidly using the Hamming distance and require less memory. This descriptor uses an annular partitioning of the surrounding environment of a landmark, see Fig. 3. The features which describe the landmark are statistical measures for the elements of each ring, n_{Rings} in total. These measures are the maximum, minimum, mean, standard deviation and the median for the values of the corresponding ring, which are $n = 5$ in total. This descriptor shows very good description performance on radar-based occupancy grids [11]. However, the problem is the high computational cost of the association of these feature vectors.

The original descriptor ASD can be extended to a binary descriptor (BASD) as follows: Instead of using the statistical measures as features, the new binary feature vector consists of the comparison of these statistical measures. For each ring, these statistical measures are calculated as in the original ASD. For the binary description, each of these statistical measures is now compared to the corresponding statistical measure in the remaining outer rings, which are $n_{\text{Rings}} \cdot (n_{\text{Rings}} + 1)/2$ comparisons. These comparisons serve as new binary features. This leads to a feature vector of size $n \cdot n_{\text{Rings}} \cdot (n_{\text{Rings}} + 1)/2$, which requires less memory for storage since the entries are binary. Experiments show that many of these statistical measures do not contribute to the description of landmarks. It turns out that the mean and standard deviation are the most descriptive features.

V. Experiments

Experiments were conducted to display the performance of the proposed method in landmark detection and description. The experiment used eight rides through a parking lot. A test vehicle was equipped with four short range radars mounted around the vehicle in order to provide a 360° perception of the surroundings. The driven trajectory through the parking lot for each ride was about 400 m. For each ride, a measurement grid was built as mentioned in Section II-B. For ground truth, a deeply coupled system consisting of a DGPS receiver with an accuracy up to 2 cm and an INS system (iMAR iTRACE F400-E) served as the reference system. So, each cell in every measurement grid of each ride corresponds to the same location in the parking lot due to the accurate global reference system. In the following experiments, the detection and the description performance for landmarks are shown.

A. Landmark Detection

In the first experiment, the performance of landmark detection was evaluated on the parking lot data set. The data set of eight grids was split into two different subsets.

One of these subsets was used to tune the parameters of the proposed FSCD detector, *Speeded-Up Robust Features* (SURF) [9], *Features from Accelerated Segment Test* (FAST) [7] and *Scale-Invariant Feature Transform* (SIFT) [10] detectors.

For two observations, the *repeatability* measure served as the indicator for the performance of the detectors. For each landmark in the first measurement grid, any landmark located less than 0.3 m away in the second measurement grid was located. If such a pair was determined, the landmark was said to be found on the second observation. If landmarks in the first observation are located very close to each other, these landmarks may be associated with the same landmark of the second observation. To avoid this effect, a landmark is constrained to be only part of one pair.

The repeatability score is simply the ratio of all detected landmarks and the amount of found landmarks according to the previous criterion.

To determine the best parameters, three measurement grids were used and thus six grid pairs in total were given (comparing the identical grid was skipped). These parameters were used on the remaining grids for the evaluation, which consisted of 20 observation pairs. The result of the repeatability is shown in Fig. 4. All detectors provided reliable landmarks whereas SURF and FAST have the worst reliability with a mean below 50 %. On the data set, the slower SIFT detector achieves best reliability compared all image processing detectors. It is remarkable that the proposed FSCD detector attains best reliability on the data set which makes it a qualified candidate for landmark detection.

B. Landmark Description

In this section, the performance of the landmark description is evaluated. For comparison, state-of-the-art algorithms were applied on the measurement grid, *Fast REtinA Keypoints* (FREAK) [8], *Speeded-Up Robust Features* (SURF) [9] and *Scale-Invariant Feature Transform* (SIFT) [10]. For these algorithms, the implementations in OpenCV 2.4.7 were applied.

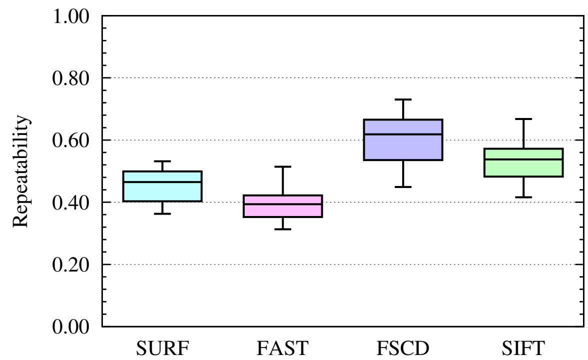

Fig. 4. This boxplot illustrates the performance of the landmark detection algorithms. The proposed FSCD detector performs best compared to the other algorithms.

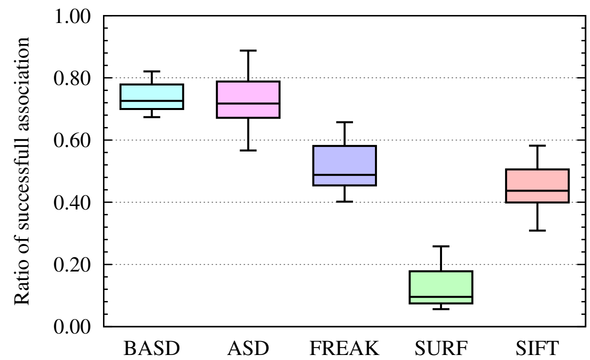

Fig. 5. For the evaluation of the descriptors, this boxplot shows the performance of the feature vectors provided by each descriptor. It is remarkable that the proposed BASD and ASD show the best performance for the dataset.

The FSCD detector outputs were used as landmarks for the experiment. These landmarks were detected on the measurement grids of the corresponding observations. For each landmark, a feature vector was calculated by each descriptor. Next, for each landmark of the first observation the landmark of the second observation with the smallest error according to an appropriate measure was determined. This pair was deemed to be an association. As a ground truth was available, an association was said to be correct if the position of the landmarks of this association was the same. The ratio of correct associations and possible associations served as the measure of the performance of the descriptors. The ratios of correct associations for each descriptor are shown in Fig. 5. Surprisingly, the SURF descriptor fails on this data set. This might be the reason, because the measurement grid differs to much from visual images for which SURF is optimized. In contrast to SURF, the FREAK and SIFT descriptors attain satisfying ratios of successful associations. Whereas the ASD descriptor and the proposed modification BASD perform best on the data set and outperforms the image processing descriptors. The reason is, that both descriptors are specially designed for grids from radar data where the benefit is shown in this experiment.

VI. Conclusions

This paper presents a novel combination of detection and description of landmarks designed for radar data. For representation, a so-called *measurement grid* is used where the cell values are incremented with regard to the position of the measurement. Using this grid representation, the proposed detector uses a rotational invariant pattern for landmark detection which is very fast due to simple comparisons. For the association of landmarks, the feature vector for each landmark is calculated using a modified ASD. The modification BASD is a binary version of ASD which allows faster association of landmarks. Experiments were carried out on real-world data, where the proposed method shows the best landmark detection performance regarding state-of-the-art algorithms. The description of the proposed BASD displays performance comparable to ASD, which performs best. However, BASD provides binary feature vectors requiring less memory. Due to the binary features, the feature vectors of this descriptor are fast for association finding.

Experiments show that the proposed combination of FSCD for detection and BASD for description is a qualified candidate for subsequent applications such as localization and SLAM.

References

[1] Sebastian Thrun, Wolfram Burgard, and Dieter Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*, The MIT Press, 2005.

[2] Gian Diego Tipaldi and Kai Oliver Arras, “Flirt-interest regions for 2d range data,” in *2010 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2010, pp. 3616–3622.

[3] Jianbo Shi and Carlo Tomasi, “Good features to track,” in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR ’94., 1994 IEEE Computer Society Conference on*, Jun 1994, pp. 593–600.

[4] Andrew J Davison, “Real-time simultaneous localisation and mapping with a single camera,” in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 2003, pp. 1403–1410.

[5] Hans P. Moravec and Alberto Elfes, “High Resolution Maps from Wide Angle Sonar,” in *IEEE International Conference on Robotics and Automation*, March 1985, pp. 116 – 121.

[6] Klaudius Werber, Matthias Rapp, Jens Klappstein, Markus Hahn, Jürgen Dickmann, Klaus Dietmayer, and Christian Waldschmidt, “Automotive Radar Gridmap Representations,” in *2015 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM 2015)*, Heidelberg, Germany, Apr. 2015.

[7] Edward Rosten and Tom Drummond, “Machine learning for high-speed corner detection,” in *Computer Vision–ECCV 2006*, pp. 430–443. Springer, 2006.

[8] Alexandre Alahi, Raphael Ortiz, and Pierre Vandergheynst, “Freak: Fast retina keypoint,” in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 510–517.

[9] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool, “Speeded-up robust features (SURF),” *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[10] David G Lowe, “Distinctive image features from scale-invariant keypoints,” *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[11] Matthias Rapp, Tilmann Giese, Markus Hahn, Jürgen Dickmann, and Klaus Dietmayer, “A Feature-Based Approach for Group-Wise Grid Map Registration,” in *Intelligent Transportation Systems Conference 2015*. IEEE, 2015.