NARF: 3D Range Image Features for Object Recognition

Bastian Steder Radu Bogdan Rusu Kurt Konolige Wolfram Burgard

***Abstract*— We present our findings regarding a novel method for interest point detection and feature descriptor calculation in 3D range data called NARF (Normal Aligned Radial Feature). The method makes explicit use of object boundary information and tries to extract the features in areas where the surface is stable but has substantial change in the vicinity.**

I. INTRODUCTION

In object recognition or mapping applications, the ability to find similar parts in different sets of sensor readings is very important. Most methods use *features* that describe a chunk of data in a compressed representation and that can be used to efficiently perform comparisons between different data regions. The task mainly consists of two parts, the identification of appropriate points, often referred to as *interest point*s or *key point*s, and the way in which the information in the vicinity of that point is encoded in a *descriptor* or *description vector*.

In this work we focus on single range scans, as obtained with 3D laser range finders or stereo cameras, where the data is incomplete and dependent on a viewpoint. We chose range images as the way to represent the data since they reflect this situation and enable us to borrow ideas from the vision sector.

We present the normal aligned radial feature (NARF), a novel interest point extraction method together with a feature descriptor for points in 3D range data. The interest point extraction method has been designed with two specific goals in mind: i) the selected points are in positions where the surface is stable (to ensure a robust estimation of the normal) and where there are sufficient changes in the immediate vicinity; ii) since we focus on partial views, we want to make use of object borders, meaning the outer shapes of objects seen from a certain perspective. The outer forms are often rather unique so that their explicit use in the interest point extraction and the descriptor calculation can be expected to make the overall process more robust.

Figure 1 shows an outline of our interest point extraction procedure. The implementation of our method is available under an open-source license[1].

II. NARF INTEREST POINT EXTRACTION

The detection of interest points is an important step to reduce the search space for feature extraction and focus the attention on significant structures. We have the following requirements for our interest point extraction procedure: i) it must take information about borders and the surface structure into account; ii) it must select positions that can be reliably detected even if the object is observed from another perspective; and iii) the points must be on positions that provide stable areas for normal estimation or the descriptor calculation in general.

Stable interest points need significant changes of the surface in a local neighborhood to be robustly detected in the same place even if observed from different perspectives. This typically means, that there are substantially different dominant directions of the surface changes in the area. To capture this, we

- find borders in the range image, meaning non-continuous traversals from foreground to background,

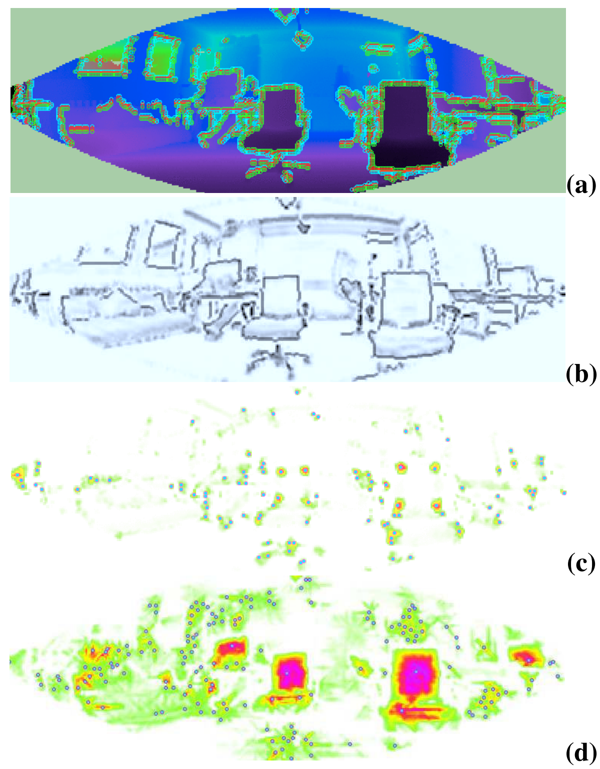

Fig. 1. The interest point extraction procedure. **(a)** Range image of an office scene with two chairs in the front with the extracted borders marked. **(b)** Surface change scores according to borders and principle curvature. **(c)** Interest values with marked interest points for a support size of 20cm. Note how the corners of the chairs are detected as interest points at this scale. **(d)** Interest values with marked interest points for a support size of 1m. Note how, compared to (c), the whole surface of the chair's backrests contain one interest point at this scale.

B. Steder and W. Burgard are with the Dept. of Computer Science of the University of Freiburg, Germany. {steder,burgard}@informatik.uni-freiburg.de; R. Rusu and K. Konolige are with Willow Garage Inc., USA. {rusu,konolige}@willowgarage.com

[1]This work is part of the ROS point cloud library. See http://www.ros.org/wiki/pcl/Tutorials

by looking for substantial increases in the 3D distances between neighboring image points.

- look at the local neighborhood of every image point and determine a score how much the surface changes at this position and a dominant direction for this change, incorporating the information about borders,
- look at the dominant directions in the surrounding of each image point and calculate an interest value that represents i) how much these directions differ from each other and ii) how much the surface in the point itself changes (meaning how stable it is),
- perform smoothing on the interest values, and
- perform non-maximum suppression to find the final interest points.

III. The NARF descriptor

Feature descriptors describe the area around an interest point in a way that makes efficient comparison regarding similarity possible. Our goals in the development for the NARF descriptor were i) that it captures the existence of occupied and free space, so that parts on the surface and also the outer shape of an object can be described, ii) that it is robust against noise on the interest point position, and iii) that it enables us to extract a unique local coordinate frame at the point.

To compute the NARF descriptor in an interest point, we

- calculate a normal aligned range value patch in the point, which is a small range image with the observer looking at the point along the normal,
- overlay a star pattern onto this patch, where each beam corresponds to a value in the final descriptor, that captures how much the pixels under the beam change,
- extract a unique orientation from the descriptor,
- and shift the descriptor according to this value to make it invariant to the rotation.

The unique orientation together with the normal then defines a local 6DOF coordinate frame at the point. Please consider Figure 2 for a visual example of the process and application.

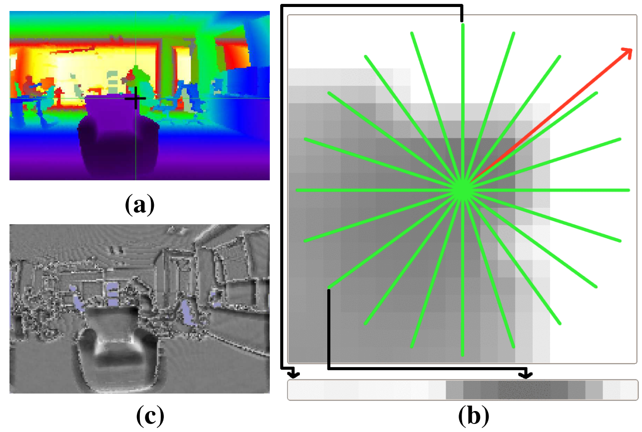

Fig. 2. **(a)**: A range image of an example scene with an armchair in the front. The black cross marks the position of an interest point. **(b)**: Visualization how the descriptor is calculated. The top shows a range value patch of the top right corner of the armchair. The actual descriptor is visualized on the bottom. Each of the 20 cells of the descriptor corresponds to one of the beams (green) visualized in the patch, with two of the correspondences marked with arrows. The additional (red) arrow pointing to the top right shows the extracted dominant orientation. **(c)**: The descriptor distances to every other point in the scene (the brighter the higher the value). Note that mainly top right rectangular corners get low values.

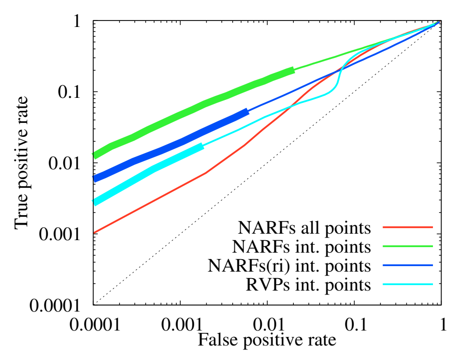

Fig. 3. This graph shows ROC curves for feature matches where the threshold for the descriptor distance increases from the bottom left to the top right. Please note, that the axes are logarithmic. The higher a line in this graph, the better is the performance.

IV. Results

We wanted to test the capability of our NARF-descriptors to match features from an object model to the features in a scene containing the model. Figure 3 shows the results for the matching process of seven object models (armchair, cart, chair, cup, pioneer robot, stapler, and teddy bear) against scenes with known ground truth positions. To match the objects against the scenes, we sampled poses from our models that differed from the ground truth poses between 0° and 50°.

The results of true positives and false positives are summarized as ROC (Relative Operating Characteristic) curves. Please note the logarithmic scales, which show the interesting part of the plot, the bottom left corner, better. The absolute number of false positives is much higher than the absolute number of true positives, which makes areas with a high ratio of false positives less useful. The thicker plots mark areas, where the number of true positives to false positives is lower than 1:10. The plot marked *NARFs all points* is for our NARF feature descriptors extracted at every image point, without using the interest points. The plot marked *NARFs int. points* shows the performance of the NARF features together with the interest points. The plot marked *NARFs(ri) int. points* is for the rotational invariant version of the NARF descriptor. The plot marked *RVPs int. points* is using the range value patches that we used in our former work [1] as feature descriptors. As can be seen, the interest points are a definite improvement compared to random point positions. Additionally, the rotationally invariant version and the rotationally variant version of the NARF descriptor outperform the RVPs we used before.

References

[1] B. Steder, G. Grisetti, M. Van Loock, and W. Burgard. Robust on-line model-based object detection from range images. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, St. Louis, MO, USA, October 2009.