A Confidence Measure for Segment Based Maps

Rolf Lakaemper
Temple University
Philadelphia, PA, USA
lakamper@temple.edu

ABSTRACT

Map confidence, or map quality based on regional consistency is an important measure to evaluate the quality of robot maps. It is classically handled analyzing occupancy grids, which is an unnatural choice if the map is not represented by data points, but by line segments. We define a map-confidence measure that is tailored for segment based maps, without leaving the compact data representation by segments. The presented confidence measure is not based on comparison to ground truth data, but evaluates the map (ground truth free) based on map consistency.

1. INTRODUCTION AND APPROACH

The interest in robot mapping based on higher geometric structures like linear elements is currently growing. Obvious advantages in runtime, memory efficiency and simpler mid-level analysis capability make such mapping approaches powerful competitors to the classic, point based techniques. Computing the confidence in scans, and evaluating a global map based on the confidence of aligned single scans in the context of the global map is an important task in robot mapping. This paper gives an example, how map confidence can be computed purely based on line segments. The presented approach evaluates the confidence in each single segment; it can be used to delete inconsistent segment data ('map cleaning'), as well as to score the quality of a given segment map based on segment consistency.

The core algorithm was originally designed as a processing module of a segment based robot mapping system (description of this system is part of a future publication). Its purpose there is to clean intermediate mapping results, consisting of a low number of aligned, segment based local maps, from inconsistent or noisy segments. In a straightforward manner, such a module can be extended to a global (or regional) confidence measure: the more consistent segments (in a certain region), the better.

An important design paradigm of the presented research is not to leave the very efficient and compact data representation by segments. Such an approach leads to multiple advantages compared to point/grid based methods:

- The segment based approach captures structural information. This information goes significantly further than the information of object presence, contained in raw point data. Figure 3, (b,c,d) shows examples: the red segments are detected as noise. In point density based confidence approaches like occupancy grids [2], the original data points would not only be labeled as correct (since the point density in this area is high), but would have enhanced the confidence in this region.
- The segment based approach is fast. In indoor environments or urban outdoor environments, a typical scan consists of $n < 20$ segments of sufficient length, while the number of data points is typically one to two orders of magnitude (factor $10 - 100$) higher. This becomes especially important when point relations between different scans have to be evaluated, which usually implies algorithms with runtime between $O(nlogn)$ and $O(n^2)$.
- The segment based approach is memory efficient. Compared to occupancy grids, the memory consumption is significantly lower.
- The segment based approach is precise. Segment endpoints don't have to be adjusted to a resolution parameter, hence there are no quantization errors. This is in contrast to grid based approaches.

The basic idea of the approach is to cluster segments, based on an inter segment-distance measure. The quality of clusters defines the confidence in the participating segments, which in turn defines the confidence in the entire map. The main steps are i) the definition of a perceptually consistent segment-distance measure, ii) the adaption of a classic clustering technique (hierarchical clustering) to gain a parameter free clustering system, and iii) a new measure for intra cluster consistency, which directly leads to the final goal, the confidence measure.

2. RELATED WORK

To the authors's best knowledge, there are no publications available about generic map evaluation, the reason being that map evaluation is highly task specific. Task specific map evaluation is usually performed in the broader environment of robot competitions, such as RoboCup [1] or the US Department of Energy Grand Challenge [13]. Test arenas,

developed by the National Institute of Science and Technology (NIST) exist [4], as an effort to create robot maps in standard environments. These arenas were used in various events, e.g. the RoboCup Rescue competition and the Response Robot Evaluation Exercise [11].

An occupancy grid based evaluation tool, the Jacobs Map Evaluation Toolkit [3], was utilized in the RobocupRescue competition 2008. Aside from functionalities like ground truth map creation, it consists in its core of a metric comparing the (grid/pixel based) maps. In short, correspondences between foreground points of the evaluated map and a ground truth map are established. The correspondence quality is computed using the spatial distance of the corresponding points.

In contrast, the presented evaluation method does not perform a comparison to a ground truth map, but aims to analyze the consistency of a single map. Working on a higher data structure, line segments, it tries to capture regional structural properties. These are evaluated based on their ambiguity of representation: a single cluster represents a single feature, a high intra cluster distance can be interpreted as ambiguity, or low confidence.

A more general introduction and overview of benchmarking and evaluation in robotics is given in [10].

3. INTER SEGMENT DISTANCE

This section introduces a distance measure between pairs of line segments s_1, s_2. The basic idea of the distance measure is to merge two line segments to an 'average' segment $\bar{s}$. The distance is the merging cost, which consists of three parts:

- the angular distance between s_i and $\bar{s}$, $i = 1, 2$
- the spatial distance between s_i and $\bar{s}$, $i = 1, 2$
- the spatial distance between s_1 and s_2.

The first two parts penalize the amount of 'non collinearity' of the segments, the third part penalizes spatial distance. Although used as a distance measure between two segments, the design is based on comparison to a 'virtual' average segment. This is motivated by certain experiments, suggesting that human perception assigns or connects line segments to larger structures under certain circumstances. For example, two collinear, overlapping line segments are perceived as one line, i.e. both segments represent the same element and should therefore have a distance of zero (which is the case for our distance measure). Please note that such a distance measure is no metric. It already disobeys the most 'intuitive' axiom of the metric axioms, the *identity of indiscernibles* ($d(a, b) = 0 \leftrightarrow a = b$), since two non identical collinear segments s_1, s_2 with $s_1 \cup s_2 \neq \emptyset$ have a distance $d(s_1, s_2) = 0$. This fact becomes important for the choice of the clustering algorithm, see section 4

The definition of the measure is out of scope of this paper and will be part of a future publication. Figure 1 gives examples of segment configurations and resulting distances.

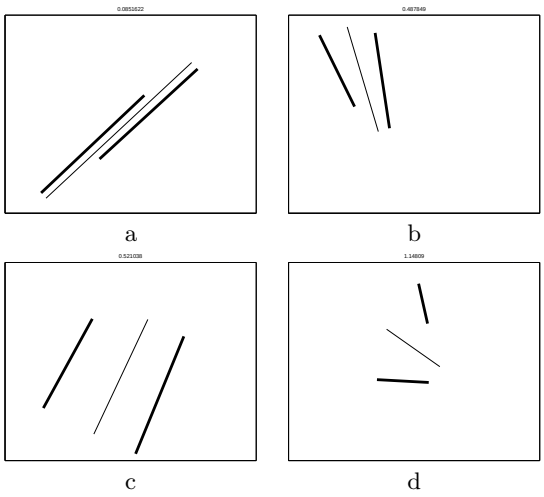


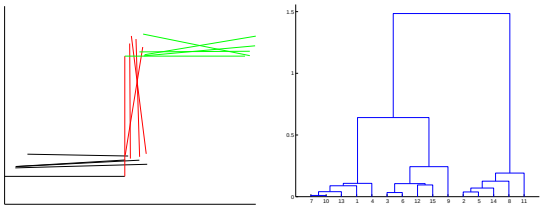
Figure 1: Segment configurations with increasing distance. a) 0.09 b) 0.49 c) 0.52 d) 1.14. The thin line is the merged segment. The increase in a)-c) results from larger intra segment distance, while d) results from angular distance.

4. CLUSTERING

For the clustering, agglomerative hierarchical clustering in 'single' mode is utilized. This method seeks to build a bottom up hierarchy of clusters, starting with each segment being a single cluster, ending in a single cluster containing all segments. Pairs of clusters are merged as one moves up the hierarchy. The merge is determined in a greedy manner: the two clusters with minimal distance are merged to a single one. Hierarchical clustering allows for different strategies to determine the distance of the newly emerged cluster to the remaining elements. In our case, we use the 'single mode' strategy: the distance between two clusters is the minimum distance between their elements. There is a geometric motivation for the use of this mode: in the example of collinear, slightly overlapping segments single mode clusters these segments to a single group — intuitively, single mode clustering acts like a connected components algorithm, the necessary topology being defined through the distance measure (small distance = neighbors).

Hierarchical clustering has two main properties which suggest its use in the segment merging context: first, it is, in its first stage, parameter free, i.e. no pre-defined number of clusters has to be determined. Parameters might be introduced later in a follow up stage, which selects the level of clustering (agglomerative hierarchical clustering always ends in a single cluster). Second: it is simply based on mutual distances between the data points (here: line segments), yet without the need to embed them in a metric space. This means, hierarchical clustering can deal with any distance measure (especially non-metrics, as in the given case).

We want to illustrate the segment clustering by a simple example, see Figure 2. The data set of this example consists of 15 segments, which can intuitively be combined to 3 clusters. Figure 2,b), shows the resulting *dendrogram*. Each horizontal bar shows the *linkage* L_i between two clusters, L_i is assigned the minimal distance between elements of the left and right subtree of the linkage; in the dendrogram this cost is displayed by the height of the bar. In this simple case, the dendrogram clearly suggests the three clusters. The critical step in hierarchical clustering is to define the step to end the clustering process. We do so if a potential merge de-



a b

Figure 2: Clustering example. Left: Segments. Right: Dendrogram. See text for details.

creases the intra cluster consistency significantly. To determine clusters, we assign a *consistency value* $c(L_i)$ to each linkage L_i. $c(L_i)$ compares the linkage distance L_i with all linkage distances L_i^l, L_i^r of the left and right subtree:

$$L_i = \frac{L_i - \text{mean}(L_i^l \cup L_i^r)}{\text{mean}(L)} \quad (1)$$

with L being the set of all linkages. Data elements s_1, s_2 (segments) belong to one cluster if all linkages connecting s_1 and s_2 do not exceed a certain threshold T_c. Normalizing by mean(L) makes the approach scale independent. Our consistency measure has a clear geometric motivation, and performed well in different examples (see results in section 7). Determining clusters from dendrograms can be performed in different ways. More details about hierarchical clustering can be found e.g. in [5].

5. CLUSTER QUALITY: THE CONFIDENCE MEASURE

The main step in our evaluation is to determine the consistency of each cluster. Please observe that we already computed an intra cluster consistency value $c(L_i)$ to determine the clusters. $c(L_i)$ has certain drawbacks handling outliers, it is not necessarily consistent with the perceptual consistency. We therefore introduce a new intra-cluster-consistency measure $\mathcal{C}$ which is stronger perceptually motivated and adjusts better to the specific problem. $\mathcal{C}$ is used to re-evaluate each cluster, it is, however, too expensive to be utilized in the clustering process itself. It is therefore only applied after the clustering process is finalized.

In $\mathcal{C}$, collinear structures are favored, while clusters containing wide-spread segment sets are penalized. Similar to the segment distance measure, each segment in the cluster is compared to an average cluster segment, the cluster representative. In analogy to classic intra cluster consistency measures, the angular and spatial distance to this representative is taken into account to determine the cluster consistency. Intuitively, all angular distances of segments in one cluster to the average cluster representative are computed, as well as the transitional distances. For angular and translational distances, two separate confidence measures $\mathcal{C}_a, \mathcal{C}_t \in [0..1]$ (angular/translational respectively) are computed (see details below). The final confidence $\mathcal{C}$ is computed as

$$\mathcal{C} = \min(\mathcal{C}_a, \mathcal{C}_t). \quad (2)$$

A high confidence (1) is therefore only assigned if both, angular and translational confidence are high. Additionally, clusters must contain a certain minimal number of segments (in the current system: three segments), otherwise they are assigned a confidence of $\mathcal{C} = 0$. Figure 3 shows examples for clusters and their consistency value $\mathcal{C}$. Please note that especially Figure 3 shows the superiority of a segment based evaluation to point based occupancy grids. Figure 4 is a

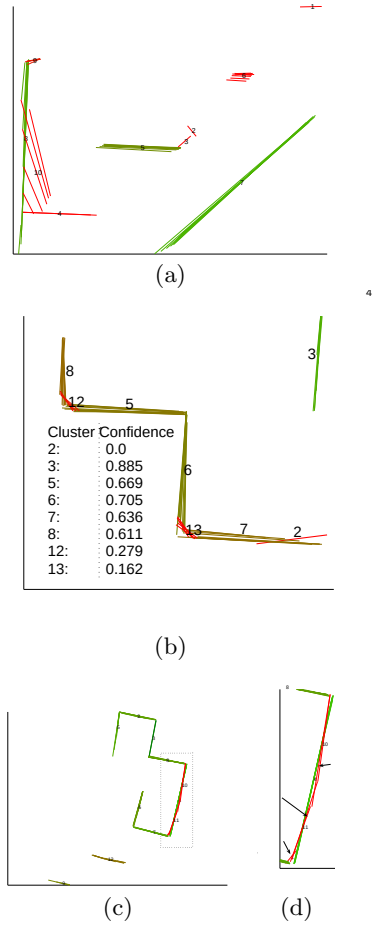


Figure 3: Evaluating clusters using the confidence measure $\mathcal{C}$. The red/green-ness is determined by confidence (the greener, the more confident). a) regions with non matching angles, widespread structures and areas of insufficient density are marked as non confident. b) segment based confidence detects structural inconsistency: the 45 degree corner scans are detected as inconsistent. d) a magnified view of the marked part of c): the correctly detected inconsistent segments have a huge overlap with consistent segments: detection of such areas is not possible with occupancy grids, but only with methods detecting underlying structural information.

comparative example showing the performance of the two confidence measures $c(L_i)$ and $\mathcal{C}(C_i)$: the tendency of both measures is approximately equal (this is why we can use the computationally cheaper $c(L_i)$ during the clustering), yet $\mathcal{C}$ yields more perceptually consistent results.

5.1 Angular Confidence $\mathcal{C}_a$

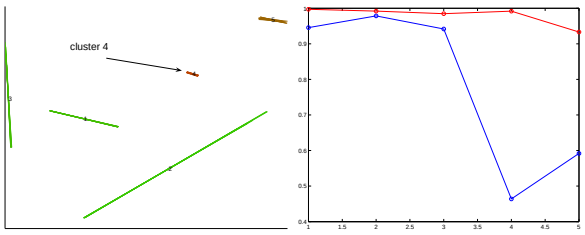


Figure 4: Left: A simple example of a map with high confidence in all 5 regions (clusters). Right: (blue) cluster quality using the distance matrix based method $c(L_i)$, evaluating cluster 4 (arrow in left figure) to be of lower quality; (red)$\mathcal{C}$, which is in accord with the perceived high consistency of all clusters.

For all segments s_i of a single cluster L, we compute their angles $a_i \in [-\pi/2..\pi/2]$ with the x-axis. We define the cluster's average angular direction θ using the weighted circular mean wcm

$$\theta = wcm(2a_i, |s_i|)/2 \tag{3}$$

where $|s_i|$ denotes the length of segment s_i, used as the corresponding angle's weight (multiplication/division by 2 ensures correct handling of segment directions). The angular distance d_a is computed as the weighted (by length) sum of distances to θ:

$$d_a = \frac{\sum_i \min((|a_i - \theta|) \mod \pi,\ \pi - ((|a_i - \theta|) \mod \pi))\ l_i}{\sum_i l_i}. \tag{4}$$

Finally, the angular confidence is computed as

$$\mathcal{C}_a = \exp \frac{d_a^2}{2\sigma_a^2} \tag{5}$$

with a parameter $\sigma_a = 0.05$ which was experimentally determined and fixed.

5.2 Translational Confidence $\mathcal{C}_t$

For the translational confidence, we compute the maximal distance t_i of each segment to a cluster-representative line S. S is defined by θ and a point $P = \sum(p_i)/\#L$, the average center point (p_i: center point of s_i, $\#L$: number of segments in cluster L). The translational distance is defined by

$$d_t = \frac{\sum t_i}{\#L} \tag{6}$$

Finally, the angular confidence is computed as

$$\mathcal{C}_d = \exp \frac{d_t^2}{2\sigma_t^2} \tag{7}$$

with a parameter $\sigma_a = 0.1$ which was experimentally determined and fixed. Observe that σ_t is scale dependant. The current value is determined for robot maps with scale unit of one meter.

6. MAP EVALUATION

It is a small step from regional evaluation of single clusters to global map evaluation. Given all clusters C_i along with their confidence measure $\mathcal{C}(C_i)$, we define the global confidence $\mathcal{M}$ of a map by

$$\mathcal{M} = \frac{\sum_i \#C_i\ \mathcal{C}(C_i)}{\sum_i \#C_i} \tag{8}$$

with $\#C_i$ denoting the cardinality of C_i. $\mathcal{M}$ computes the average consistency of all segments, defining the confidence of a segment by the consistency of the cluster it participates in.

7. RESULTS

7.1 Random Distortion

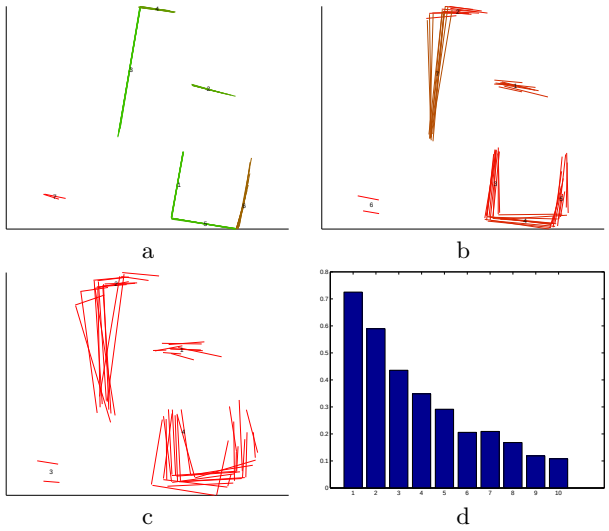


Figure 5: Global map evaluation using $\mathcal{M}$. A map with high confidence was randomly increasingly distorted in 10 steps (Figure shows step 1(a),5(b) and 10(c)). (d) shows the global confidence diagram for the resulting maps, x-axis: step 1-10, y-axis: confidence $\mathcal{M}$.

In this experiment, a map with high confidence was randomly increasingly distorted in 10 steps. Figure 5 shows steps 1, 5 and 10 and the confidence measures $\mathcal{M}$ for each of the 10 distortion levels. Expectedly, the results show decreasing confidence.

7.2 Map Comparison

In this experiment, we compare results of two mapping algorithms. The first algorithm [6] is a point based alignment (not segment based) algorithm. However, it results in corrected poses of single scans. We used an algorithm explained in [8] to extract segments from these single scans, and superimposed them, using the corrected poses (Figure 6,a). The second map (Figure 6,b) was computed by a new, segment based algorithm, which will be topic in a future publication. Both output maps consist of the same single scans' segments, yet aligned using different poses. It can clearly be seen that the first map is less consistent. Our evaluation algorithm does not only capture the overall difference in quality (quality of first map: $\mathcal{M} = 0.2769$, quality

Figure 6: Evaluation of two maps of the data set 'Freiburg082'. a) Quality of map $\mathcal{M} = 0.2769$. b) $\mathcal{M} = 0.4355$. Colors: level of green (vs red) shows confidence: The greener, the more confident, the more red, the worse. The higher regional confidence in (b) leads to the better total confidence value.

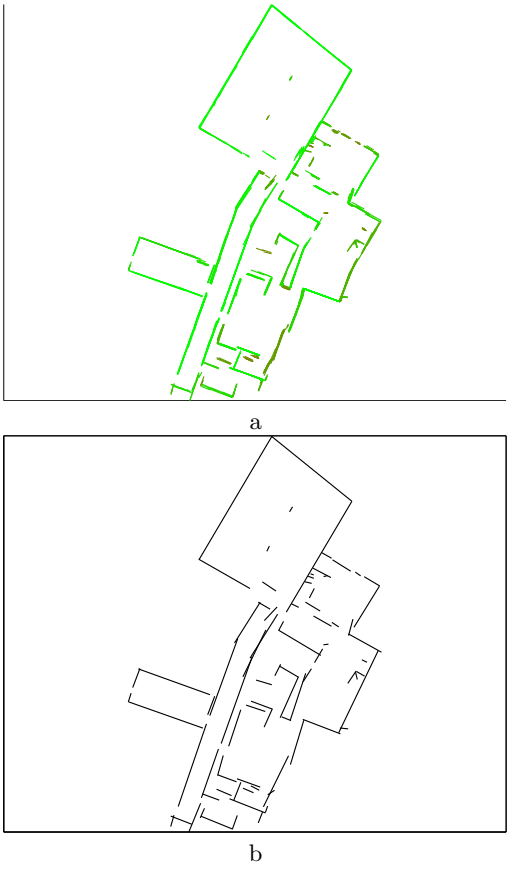


Figure 7: Using segments of high confidence only yields structural de-noising. a) segments of Figure 6,(b), belonging to clusters L_i with a confidence $c(L_i) > 0.3$ (80 clusters, overall confidence $\mathcal{M} = 0.6554$). b) clusters of (a) represented by single representative merged segments (80 segments).

of second map: $\mathcal{M} = 0.4355$), but also identifies the confidence of single clusters (regions). It is interesting to show segments above a certain confidence level only (Figure 7). In this data set, this leads to structural de-noising of the map: usually, large and static (in contrast to smaller and/or moving) objects in the environment yield high confidence representation. Therefore, the main structure of the environment is highlighted (of course, using higher quality clusters only, the map quality based on our evaluation measure increases). Additional merging of the clusters to single segments yields a clear map in a very compact representation (here: 80 segments). We performed a second comparative experiment on a different data set (data set NIST), using the mapping algorithms FFS [9] and FFS with Virtual Scans [7]. The latter one is an extension of the first, and leads to (visually inspected) improved results. Numerical evaluation of the results using the presented measure is consistent with the visual impression, see Figure 8. The maps only differ slightly in certain regions. However, the overall visual impression of (b) is slightly better than the one of (a), which is also expressed in the evaluation. The experiments leading to the respective maps are documented in [7].

8. RUNTIME

The presented algorithm has an order of magnitude of $O(n^2)$, n = total number of segments, which results from computation of the pairwise segment distance matrix. The MATLAB implementation of the algorithm needed 1 second for the experiment using data set NIST (332 segments), and 5 seconds for the experiment using data set Freiburg082 (1975 segments), both on a 1.8GHz laptop PC.

9. CONCLUSION AND OUTLOOK

The presented confidence measure evaluates maps in consistency with visual perception. In its core, it uses a classical clustering algorithm, hierarchical clustering, which is adapted to the current problem utilizing a segment distance measure and a segment based cluster confidence measure. Since segment based representation captures structural features better than its lower representation counterpart, point based maps, erroneously mapped/aligned features can be detected even if they overlap with correct features. This leads to detection of structural consistency, which is the main property evaluated by the presented approach. With a re-definition of segment distance and cluster confidence, the

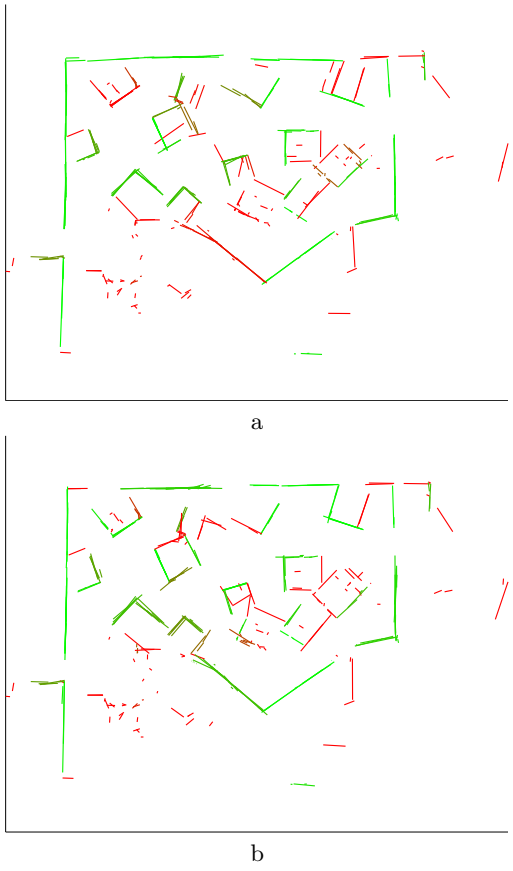


Figure 8: Mapping of data set NIST using algorithms FFS (a) and FFS with Virtual Scans (b). The evaluation leads to values of $\mathcal{M} = 0.3386$ (a) and $\mathcal{M} = 0.3876$ (b), reflecting the slight visual improvement of (b) over (a).

approach is extendable to 3D, which makes it interesting for 3D mapping algorithms based on planar elements, e.g. [12].

10. ACKNOWLEDGEMENTS

Thanks to Alexander Kleiner, University of Freiburg, for the data set 'Freiburg082'.

11. REFERENCES

[1] S. Balakirsky, S. Carpin, A. Kleiner, M. Lewis, A. Visser, J. Wang, and V. A. Ziparo. Towards heterogeneous robot teams for disaster mitigation: Results and performance metrics from robocup rescue. *Journal of Field Robotics*, 24(11-12):943–967, 2007.

[2] A. Elfes. Using occupancy grids for mobile robot perception and navigation. *Computer*, 22(6):46–57, 1989.

[3] I. Varsadan, A. Birk, and M. Pfingsthorn. Determining Map Quality through an Image Similarity Metric. In *Proceedings of the RoboCup Symposium*, July 2008.

[4] A. Jacoff, E. Messina, B. Weiss, S. Tadokoro, and Y. Nakagawa. Test arenas and performance metrics for urban search and rescue robots. In *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, volume 4, pages 3396–3403 vol.3, Oct. 2003.

[5] S. Johnson. Hierarchical clustering schemes. *Psychometrika*, 32(3):241–254, September 1967.

[6] A. Kleiner and C. Dornhege. Real-time localization and elevation mapping within urban search and rescue scenarios: Field reports. *J. Field Robot.*, 24(8-9):723–745, 2007.

[7] R. Lakaemper. Improving sparse laser scan alignment with virtual scans. In *International Conference on Intelligent Robots and Systems (IROS08)*, Nice, France, September 2008. IEEE.

[8] R. Lakaemper. Simultaneous multi-line-segment merging for robot mapping using mean shift clustering. In *International Conference on Intelligent Robots and Systems (IROS09)*, St Louis, MO, USA, September 2009. IEEE.

[9] R. Lakaemper, N. Adluru, L. Jan Latecki, and R. Madhavan. Multi robot mapping using force field simulation: Research articles. *J. Field Robot.*, 24(8-9):747–762, 2007.

[10] R. Madhavan, R. Lakaemper, and T. Kalmar-Nagy. Benchmarking and standardization of intelligent robotic systems. In *14th International Conference on Advanced Robotics (ICAR 2009)*, Munich, Germany, June 2009.

[11] NIST. NIST Response Robot Evaluation Exercise. Search and Rescue: Texas Engineering Extension Service (TEEX), November 2008.

[12] K. Pathak, N. Vaskevicius, J. Poppinga, M. Pfingsthorn, S. Schwertfeger, and A. Birk. Fast 3d mapping by matching planes extracted from range sensor point-clouds. In *International Conference on Intelligent Robots and Systems (IROS09)*, St Louis, MO, USA, September 2009. IEEE.

[13] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, K. Lau, C. Oakley, M. Palatucci, V. Pratt, P. Stang, S. Strohband, C. Dupont, L.-E. Jendrossek, C. Koelen, C. Markey, C. Rummel, J. van Niekerk, E. Jensen, P. Alessandrini, G. Bradski, B. Davies, S. Ettinger, A. Kaehler, A. Nefian, and P. Mahoney. Stanley: The robot that won the darpa grand challenge: Research articles. *J. Robot. Syst.*, 23(9):661–692, 2006.