Template Matching for Radar-Based Orientation and Position Estimation in Automotive Scenarios

Johannes Schlichenmaier*, Niranjan Selvaraj*, Martin Stolz†, and Christian Waldschmidt*

*Institute of Microwave Engineering, Ulm University, 89081 Ulm, Germany
Email: johannes.schlichenmaier@uni-ulm.de
†Robert Bosch GmbH, Advanced Engineering Sensor Systems, P.O. Box 16 61, 71226 Leonberg, Germany

Abstract—**Future high-resolution radars enable new functionalities in advanced driver assistance systems, such as estimation of contour, position, and orientation of vehicles on the road. However, straightforward approaches like that of Oriented Bounding Box generally fail in challenging automotive scenarios, when only one side of the vehicle is visible to the radar. In this paper, an estimation approach based on Generalized Hough Transform matching is presented and examined for its use in such automotive scenarios. An optimization method is discussed and finally, the algorithm is compared against a classic approach.**

I. Introduction

Upcoming trends in automotive technologies, like Advanced Driver Assistance Systems (ADAS) and even autonomous driving heavily rely on multiple sensors (e.g., LiDar, radar, video based systems), to accurately estimate properties like position and orientation of other vehicles on the road.

While sensor fusion and high-level processing like object tracking can lead to good results on their own, given enough time and processing power, an accurate estimation based on a single-shot measurement is of very high importance, because it greatly improves the initial accuracy and performance of the subsequent processing steps.

Because it is mostly unaffected by environmental properties like rain, snow, and darkness, radar as a very robust sensing method is of special interest for acquiring accurate single-shot estimations.

Currently, most of the research regarding position and orientation estimation based on radar is limited to simplistic approaches like that of an Oriented Bounding Box (OBB) [1].

While computational efficient, these simple approaches fail to fully cover all possible scenarios. Due to the complex interaction of radar signals with the environment, scattering centers on automobiles tend to become unstable depending on incident angle. This is especially true if only one side of the car is fully visible to the radar.

II. Problem Formulation

Classic estimation approaches like OBB generally fail to produce a meaningful contour in scenarios where only one side of the vehicle is visible to the radar, or not enough target points are available to accurately estimate both width and length. In these cases, a different approach is needed to provide subsequent algorithms (e.g., tracking) with a good estimation of orientation and position. In the following, such an algorithm based on template fitting is proposed.

III. Generalized Hough Transform

Hough transform (HT) is an analysis method used to extract certain features (e.g., a line, a curve, etc.) from a set of test points (e.g., pixels, target points, etc.) [2]. The basis for HT is to span a (discretized) space of parameters used to describe the desired feature and to numerically accumulate possible fits for every parameter set valid for every test point in a so called *accumulator matrix* $A(\zeta_1, \zeta_2, \ldots, \zeta_N)$. In the simplest case, the desired feature is a line, often represented in Hesse normal form $\vec{x} \cdot \vec{n}_0 = d$ and described by parameters $\zeta_1 = d$ and $\zeta_2 = \angle \vec{n}_0$. After all test points have been taken into consideration, the parameter set with the highest value in the accumulator matrix is considered the best fit for the desired feature.

Generalized Hough Transform (GHT) is an extension of standard Hough Transform, which was developed to use the principles of HT on arbitrary shapes instead of analytical functions like lines and curves. It too uses an accumulator matrix to find most likely parameter sets for the shape [3]. It is almost exclusively used in image processing, where it can very effectively find various shapes and forms in pictures, like circles, rectangles, or polygons.

A. Template Pre-Analysis

Instead of an analytical function, the feature to be extracted in GHT is an arbitrary shape, the template. The template itself is a pixel image, in the simplest case describing the desired shape with zeros and ones along its edges.

Before the actual GHT analysis, the template is examined pixel by pixel for *pixel gradient*, as well as the *relative angle* and *relative distance* to a central reference point of the template, the so called *centroid* [3]. This information fully describes the template in a way that can be subsequently used more effectively in the actual GHT analysis, to find the best matching *orientation* and *position* in the given parameter space (i.e. the accumulator matrix $A(x, y, \theta)$).

B. Accumulator Matrix

For GHT, the accumulator matrix $A(x, y, \theta)$ has the same x-y discretization as the template and an additional discretization
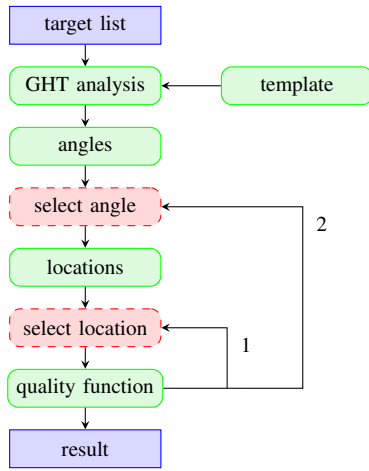
Fig. 1. Algorithm chain for contour estimation with GHT.

Fig. 2. Maximum scores of accumulator matrix for every orientation angle.

for the orientation θ, all of which directly influence the matching error margin for target points. In a too finely discretized accumulator matrix the chance of two misaligned target points voting for the same position and orientation is very slim, so some amount of discretization is necessary. On the other hand, the discretization generally limits the attainable accuracy, so it should not be chosen too coarse.

After every target point has voted for their probable centroid positions and orientations, the accumulator matrix can be evaluated. In standard GHT, the entry with the highest accumulated value is considered the correct position and orientation set. This, however, for the most part is only valid for top-down images with clear edges of high resolution.

In radar data this is usually not the case. Most of the time, radial 2D radar measurements are made in the same plane as the resulting image ($r=\sqrt{x^2 + y^2}$), which means that only the objects' near edges are visible at any given time. Also, neither do radar target points usually form closed edges, nor are they lined up ideally at the objects' edges, which leads to discontinuous and, compared to optical images, low resolution radar images.

To account for this circumstance, a quality function is employed in this work as described in Section IV-B. This quality function is used to better fit the template into the target points based on the characteristic behavior of radar signals in automotive scenarios. This extends GHT to Generalized Hough Transform with Quality Function (GHT/QF).

IV. Matching the Vehicle Box

Fig. 1 shows the algorithm chain used to match the template to the provided target list. The target list may be clustered using an arbitrary clustering algorithm to get rid of unwanted noise or separating multiple objects before feeding it into the chain. However, other than classic approaches like OBB, the template based approach does not rely on clustering and, generally, will also work well, if more than the target points belonging to the object are being fed into the chain.
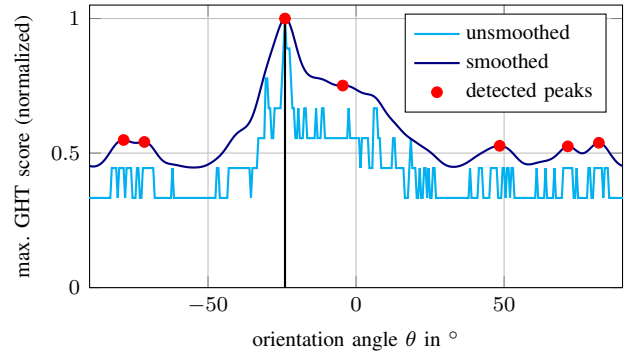
Together with the pre-analyzed template, the target list first undergoes a GHT analysis as described in Section III, after which the maximum score for every orientation angle θ is determined (see Fig. 2).

The template itself has to be known before applying the algorithm, e.g., from earlier measurements or a-priory knowledge of the sought-after object.

To mitigate computational cost, but also to avoid missing peak values on plateaus of similar value, the angle distribution first is smoothed with a Gaussian kernel before probable peak orientation angles are selected for further investigation.

For every one of the angles of interest, the 2D accumulator matrix gets investigated for local maxima, each indicating a possible matching position. For each of those local maxima, defining a full parameter set of x, y, and θ, the quality function evaluates the fit and calculates a score. At the end, the parameter set with the highest score is selected as most probable position and orientation for the template.

A. GHT Parameters

As already discussed in Section III, the parameters of the GHT space, namely the discretization in x, y, and θ are very important for a good result of the GHT analysis. For the radar data examined here, the following values were chosen:

$$\Delta x = 10\,\text{cm}, \quad \Delta y = 10\,\text{cm}, \quad \Delta\theta = 0.5°. \quad (1)$$

These values are in the same magnitude as the radar resolution, which proved to result in a viable amount of accumulation without unnecessarily sacrificing accuracy of position, for the discretization being only marginally coarser than the resolutions of the radar.

As basis for the template, a rectangular polygon with length $l=1.8\,\text{m}$ and width $w=4.8\,\text{m}$ was chosen in accordance with the known dimension of the used test car. For smoothing the GHT scores, a Gaussian kernel with $\sigma=\frac{\pi}{64}$ was used.

B. Quality Function

Because of the aforementioned behaviour of radar signals to primarily interact with near edges of objects, a subsequent investigation of the selected orientations and positions is necessary, to overcome shortcomings of the GHT principle when being used on radar data.

The quality function takes into consideration a multitude of geometric properties about the relative position of target points to the template's edges under test, as well as the orientation and position of the template itself in regards to the position of the radar sensor. Fig. 3 shows the involved geometric entities and their different properties as they are used in the quality function for a sample target point P and the edge $m=3$.

1) LOS Factor γ_m: The Quality Function differentiates between line of sight (LOS) and non-line of sight (NLOS) edges. To accomplish this, for each edge m a line from the midway point M_m to the position of the radar sensor R is drawn. Edges whose line $\overline{M_m R}$ (- - -) does not intersect any of the other edges are considered fully visible to the sensor (LOS edge). If an edge m has an intersecting point I_m, it is considered not (fully) visible to the radar (NLOS edge).

Because points from more visible edges are of more importance to the quality function, a LOS factor γ_m is determined for every edge, ranging from 1 (fully visible) to 0 (fully concealed). LOS edges get assigned $\gamma_m=1$, whereas the LOS factor of an NLOS edge is determined by how close the intersection point I_m is situated to the midway point of the edge opposite to m. Thus, the ratio of i_m (—) to e_m (—) determines the LOS factor in such a way, that $\gamma_m=0$ for $i_m=e_m$ with a linear progression to $\gamma_m=1$ for $i_m=0$.

2) Prominence Factor p_n: Every corner point C_n is assigned a prominence factor p_n which is either 1, if the corner point is connecting to one or more LOS edges, or 0, if not. This prominence is used to assign a higher score to points near a visible corner point (i.e. C_2) in contrast to points near a non-visible corner point (i.e. C_4).

3) Projection Score S_p: The projection score is closely related to the prominence factor and describes the probability for a target point to be situated at a certain distance between the two corner points of a certain edge of the vehicle. It is uniformly 1 along LOS edges and falls with an raised cosine along NLOS edges to gradually punish target points away from prominent corners. Additionally, a grace distance of $0.5\,\text{m}$ allows for points being near a corner, but not above the edge.

The projection score of a point with the scalar projection value d_p (—) for the edge m of length l_m, enclosed by corner points C_m and C_{m+1} is calculated as follows:

$$S_p = \begin{cases} p_m, & \text{for } -0.5\,\text{m} \leq d_p < 0\,\text{m} \\ p_{m+1}, & \text{for } l_m < d_p \leq l_m + 0.5\,\text{m} \\ \mathcal{R}\left(\frac{d_p}{l_m}\right), & \text{for } 0\,\text{m} \leq d_p \leq l_m \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

with p_m and p_{m+1} being the prominence factors of the enclosing corner points and $\mathcal{R}$ being a modified, inverted raised cosine with roll-off factor $\beta = 0.01$, assuming values between the prominence factors of the two respective corners:

$$\mathcal{R}(x, v_1, v_2) = \left(1 - \text{si}(x)\cos(\pi\beta x)\left(\frac{1}{1 - 4\beta^2 x^2}\right)\right)$$
$$\cdot |p_m - p_{m+1}| + \min(p_m, p_{m+1}). \quad (3)$$
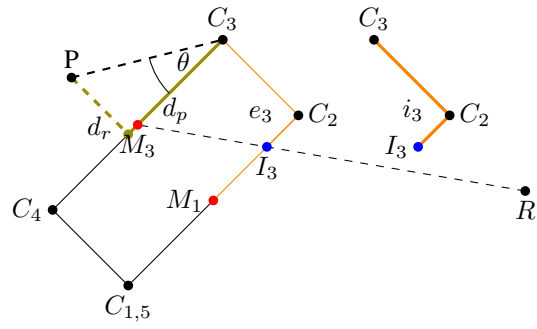

Fig. 3. Exemplary visualization of the geometric properties considered in the quality function for a target point P and the edge $m=3$ ($\overline{C_3 C_4}$) belonging to a possible position and orientation of the template, denoted by the rectangle $\overline{C_1 C_2 C_3 C_4 C_5}$ as well as the radar sensor position R. Edge $\overline{C_3 C_2 I_3}$ is redrawn to denote i_3.

4) Rejection Score S_r: The rejection score describes the probability for a target point to appear at a certain distance from an edge. As target points inside the boundaries of the vehicle are very possible due to bounces from the floor, points inside the template rectangle are considered up to a rejection distance of $1.5\,\text{m}$. Points outside are considered up to a rejection distance of $0.2\,\text{m}$.

The rejection score of a point with the scalar rejection value d_r (- - -) to the edge m is calculated as follows:

$$S_r = \mathcal{G}(d_r, 0.55\,\text{m}, 0\,\text{m}) \cdot \text{rect}\left(\frac{d_r + 0.85\,\text{m}}{1.7\,\text{m}}\right) \quad (4)$$

with $\mathcal{G}(x, \sigma, \mu)$ being a standard Gaussian function.

5) Confidence Factor $\mathcal{K}$: The confidence factor accounts for the most likely event, that most target points T will be situated within the boundaries of the template, and is calculated using the following formula for a given position and orientation:

$$\mathcal{K} = \left(\frac{T_{\text{inside}}}{T_{\text{total}}}\right)^2. \quad (5)$$

Like the rejection score, it honors a grace border of $0.2\,\text{m}$ around the template, in which target points are considered inside points.

6) Total Score: The total quality score Q combines all the aforementioned factors and point scores for a specific position and orientation using the following formula:

$$Q = \mathcal{K}\sum_m^M \left(\sum_t^T S_{p,t} S_{r,t}\right)\gamma_m, \quad (6)$$

where M is the number of edges and T the number of points.

For every edge, the projection and rejection scores for all target points are multiplied to represent the probability of the points being near that edge. This edge score is put into perspective by weighting it with the respective edge LOS factor to avoid NLOS edges having too much impact on the decision. The summation of the weighted edge scores is then qualified by the confidence score to force the algorithm to prefer more inside points.
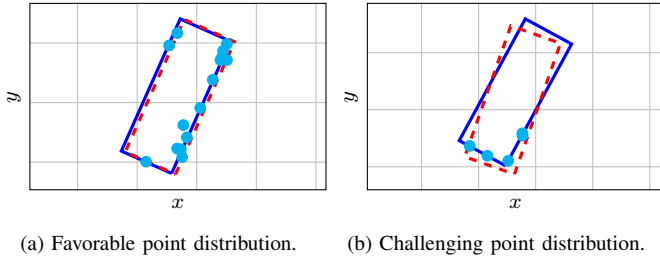
(a) Favorable point distribution. (b) Challenging point distribution.

Fig. 4. Examples of template based position and orientation estimation (———) for different target point distributions (●) with ground truth rectangles (- - -).

(a) Orientation error distributions. (b) Position error distributions.

Fig. 5. Normalized PDF (———) and CDF (- - -) examination for orientation and position errors in challenging scenarios.

V. MEASUREMENT DATA

To verify the effectiveness of the extended GHT approach, a total of 5,679 radar measurements were analyzed with the presented template based estimation (GHT/QF) and tested against a more classic approach using OBB, similar to [4].

The radar measurements were conducted with a 1x8 SIMO radar using a bandwidth of $B=2\,\mathrm{GHz}$ operating in the 77 GHz frequency band. Using a chirp sequence modulation with a high number of ramps and short ramp duration, a range resolution of $\Delta R=0.08\,\mathrm{m}$ and a velocity resolution of $\Delta v=0.11\,\frac{\mathrm{m}}{\mathrm{s}}$ were achieved.

Using a car of type BMW Model E61 equipped with an Automotive Dynamic Motion Analyzer (ADMA) sensor unit as test vehicle, figure eight driving maneuvers as baseline for comparison against the OBB approach, as well as driving away from the radar sensor and transversal to radar sensor were conducted.

VI. RESULTS

To get a first impression of the performance of the GHT/QF approach, all the measurements were processed with both the GHT/QF and OBB approach in regards to orientation and the 2D position error. The mean error, standard deviation, and median error are shown in Tab. I.
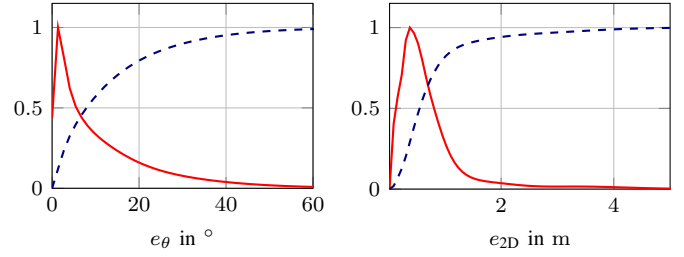
As can be seen, both algorithms perform very similarly with only a slightly higher precision for the OBB. This stems from the very good and easy clustering due to the high number of target points per object which are well situated along the edges with very little clutter (see Fig. 4 (a)). This makes the construction of an OBB very easy, whereas the template based approach is limited by its discretization in both position and orientation.

A. Challenging Scenarios

For 690 of the measurements, the OBB algorithm could not find a bounding box of viable width and length (see Fig. 4 (b)).

TABLE I
RESULTS FOR THE BASELINE TEST.

	$\overline{e_\theta}$ in °	std(e_θ) in °	md(e_θ) in °	$\overline{e_\mathrm{2D}}$ in m	std(e_2D) in m	md(e_2D) in m
GHT/QF	7.82	8.59	4.77	0.45	0.39	0.37
OBB	6.29	7.36	3.51	0.42	0.62	0.31

TABLE II
GHT/QF RESULTS FOR THE CHALLENGING SCENARIOS.

$\overline{e_\theta}$ in °	std(e_θ) in °	md(e_θ) in °	$\overline{e_\mathrm{2D}}$ in °	std(e_2D) in °	md(e_2D) in °
11.59	10.85	8.19	0.74	0.69	0.55

For these challenging scenarios, only the GHT/QF algorithm was employed. In 13 % of cases, the target point distribution did not allow for a distinction between the long and the short side of the vehicle. The results of GHT/QF for the remaining 597 measurements can be seen in Tab. II.

While not as precise as in the baseline scenario, the achieved position and orientation estimation is still in a usable range for making a good estimate. Examination of the normalized probability (PDF) and cumulative density functions (CDF) reveals that around 40 % of measurements have an orientation error of under $5.5°$, with the single mode being at $1.3°$. Regarding the position, 90 % of the measurements experience an error of under $1.46\,\mathrm{m}$, with the single mode being at $1.3\,\mathrm{m}$.

VII. CONCLUSION

In standard scenarios, the proposed template based algorithm achieves similar detection performance for orientation and 2D position of vehicles as classic methods like Oriented Bounding Box (OBB). In scenarios, where only one side of the vehicle is fully visible and OBB fails, it still achieves good orientation accuracy with a median error of $8.19°$ and a very good median 2D position error of $0.55\,\mathrm{m}$. It is therefore a highly viable alternative for when basic approaches like OBB fail due to too few target points or high clutter.

REFERENCES

[1] F. Roos *et al.*, "Estimation of the Orientation of Vehicles in High-Resolution Radar Images," in *IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, Apr. 2015, pp. 1–4.

[2] R. O. Duda and P. E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures," *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.

[3] D. H. Ballard, "Readings in Computer Vision: Issues, Problems, Principles, and Paradigms," M. A. Fischler and O. Firschein, Eds. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1987, pp. 714–725.

[4] J. Schlichenmaier *et al.*, "Adaptive Clustering for Contour Estimation of Vehicles for High-Resolution Radar," in *IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, May 2016, pp. 1–4.