Spatiotemporal Alignment for Low-Level Asynchronous Data Fusion with Radar Sensors in Grid-Based Tracking and Mapping

Georg Tanzmeister and Sascha Steyer

Abstract—Fusion of data from multiple sensors is often necessary to achieve an environment model, which meets the requirements of real-world applications, in particular those of autonomous vehicles. Sensor data fusion at a low-level yields potential advantages, as the data is fused before its interpretation with models and assumptions. However, spatiotemporal alignment, required for a precise fusion in dynamic environments, is difficult, as the sensors often cannot be synchronized. In this work, different approaches for spatiotemporal alignment of data from asynchronous sensors for low-level fusion are presented. Focus is given on radar sensors, as they allow measuring radial velocities in addition to range and bearing. The results are used to calculate fused measurement grids for grid-based tracking and mapping.

I. INTRODUCTION

Estimating the static and dynamic objects and obstacles in the local environment based on sensor data is one of the elementary capabilities of autonomous vehicles. Different sensors, such as laser scanners, radar sensors, cameras, and ultrasonic sensors are used. As every sensor has its strengths and weaknesses, it is typically required to use different measuring principles and to fuse the data in particular for safety critical outdoor applications.

Data from diverse sensors can be fused on different levels. In high-level fusion approaches tracked and classified lists of objects are fused, whereas in low-level fusion approaches raw sensor data is fused before classification and temporal filtering [1]. A major difficulty of sensor data fusion in dynamic environments arises due to asynchronously measuring sensors, as sensors often cannot be synchronized. While in high-level track-to-track fusion approaches [2], the fusion of object tracks is performed by predicting the object states onto a common time step, in low-level fusion object states do not exist at the time of the data fusion, and therefore a precise prediction is in general not possible.

One approach is to use one large Bayes filter, which receives all sensor measurements, one at each time. Object creation and association is, however, more difficult, if sensor raw data from sensors with different measurement principles is individually processed. Moreover, a missing observation from an object is typically interpreted as negative evidence for its existence. As one sensor might detect an object, which cannot be observed by another sensor, even if they observe the same area, consistent filtering can be difficult.

Apart from object-based environment estimation, grid-based representations [3], [4] are popular for estimating a map of the static obstacles around the vehicle. Grid-based models are able to represent arbitrary object shapes, in comparison to, e.g., box models, and solve the data association problem, a major problem in standard object tracking algorithms [5], implicitly with the cell discretization.

Recently, approaches extending occupancy grid mapping to dynamic environments by estimating the static and the dynamic world simultaneously have been presented [6]–[11]. They keep the benefits of grid-based representations, but yield a uniform estimation of the environment.

In evidential grid-based tracking and mapping (GTAM) [9], [11], continuous velocity probability distributions are estimated for each grid cell with a particle filter and used to derive evidences for static occupancy, dynamic occupancy, unclassified occupancy, and free space in a Dempster–Shafer model [12], [13]. From sensor data of a single time step, such as a point cloud from a laser scanner from one scan, a measurement grid is derived with a certain sensor model [4], [14] and used as input to the filter. This common low-level representation eases the use of multiple sensors and the fusion process. However, as described above, with asynchronous sensors, measurement grids cannot be directly fused.

In this work, methods to temporally align the measurements, so that the measurement grids can be correctly fused and directly used in dynamic grid approaches, are presented. Focus is given on radar sensors, as they are able to measure radial velocities in addition to range and bearing measurements. From one single measurement it is not possible to derive the 2-D total velocity of an object. Apart from tracking over time to estimate the total velocity, approaches have been presented [15], [16], which estimate the total velocity by clustering measurements from one or several radar sensors. If multiple measurements from the same object with different measurement directions are available and grouped together, the total velocity can be estimated instantaneously. Clustering, and thus object creation, however, counteract the idea of grid representations and can fail.

The rest of this paper is structured as follows. In Section II an approach is presented, which fuses the unmodified asynchronous measurements by minimizing the time differences of the measurement times. Section III discusses how measurements can be synchronized onto the same time instant using measurement prediction. Finally, Section IV shows a point-to-point radar measurement prediction approach with known object moving direction, which does not increase location uncertainty, and gives moving direction heuristics. Results are presented in Section V.
II. INTERVAL ALIGNMENT WITH ASYNCHRONOUS SENSOR DATA

This section starts with briefly discussing the different sources of errors due to assumptions and uncertainties in standard grid mapping in order to make the contribution of this paper clear. Then, an interval-based alignment that leaves the measurements unmodified is presented.

A. Assumptions and Uncertainties in Grid Mapping

Errors in standard occupancy grid mapping, as given as follows, either arise from measurement errors or from violations of assumptions.

1) Sensor model: Measurements always contain uncertainties, such as uncertainties in the measured range or the measured bearing. For every sensor a sensor model, such as [4], [14], has to be designed, which defines how a measurement grid is created from the sensor data.

2) Known ego pose: Occupancy grid mapping assumes known poses of the robot [4]. Often, odometry is used to calculate the ego poses at particular times. As long as a certain grid size around the vehicle is kept in memory, grid mapping, odometry errors, in the form of drifts, can usually be ignored, as done in this work. It is referred to SLAM literature, such as [17], if the primary source of error in the particular application is due to the ego pose.

3) Cell independence: The posterior over all possible maps is usually intractable. Assuming independence between grid cells allows grid maps of practical sizes to be computed recursively in real-time. Errors due to the cell independence can also usually be ignored with precise sensors. Batch algorithms, such as [18], dropping the cell independence assumption have also been proposed, but they cannot be computed incrementally.

4) Static world: The static world assumption is the most restricting assumption of standard occupancy grid mapping for autonomous vehicles and primary topic of this paper. As given above, approaches such as [6]–[11] have been presented, which drop it and estimate both static and dynamic obstacles. While measurement grids can directly be fused in static environments, under the assumption of known ego poses, dynamic environments pose additional challenges as shown next.

B. Interval Alignment with Reference Measurement Time

The first approach that is discussed ignores small differences in the measurement times. If the velocities in the scene are relatively low and the sensor update rates relatively high, errors due to different measurement times are small. Consider a vehicle driving with 30 km/h and two asynchronous sensors making measurements of this vehicle every 50 ms. If one sensor is taken as reference to find the closest measurement of the second sensor, then the maximum measurement time difference, ignoring sensor failures, is 25 ms. The maximum possible offset between the position in the global reference frame, e.g., an odometry-based world coordinate system, is therefore 0.2083 m. This equals 1-2 grid cells with typical cell sizes of 0.1-0.2 m. It might be even lower than the sensor uncertainty in particular with radar sensors. As it only affects dynamic objects and not the static world, it is therefore usually negligible. Measurements of the same part of the same static object will always be at the same position independent of the measurement time.

In general, the following strategy for multiple sensors with different update rates is used: the sensor with the highest measurement update frequency, i.e., with the smallest interval between consecutive measurements, is taken as reference. Let \( t^* \) denote a measurement time of the reference sensor and \( z \) a sensor measurement. All measurements that are fused at time \( t^* \)

\[
Z_{t^*} = \left\{ z | t^* - \frac{\Delta t^*}{2} \leq t < t^* + \frac{\Delta t^*}{2} \right\}
\]  

are within a window of \( \pm \frac{\Delta t^*}{2} \) around the measurement time of the reference sensor. As the sensor with the smallest update interval is used, the window, and thus the error, is minimized. Moreover, no measurement is dropped. Every measurement is included exactly once in one particular set \( Z_{t^*} \). It is assumed that all sensors have a constant update rate and that a measurement is an observation of a single point in time. All time notions represent time instances of the scene at which the measurements \( z \) have been taken, rather than time instance at which some data has been arrived. Queuing of the sensor data is therefore required, which will introduce latency. Rather than the sensor data, the measurement grids can also directly be used in (1).

If the velocities in the scene increase, however, the errors of the absolute position of the measurements become significant. Given a velocity of 130 km/h, the maximum position error in the above example is already 0.9028 m. With three or more sensors and the fastest sensor with an update rate of 20 Hz as above, the maximum position error is already 1.8056 m. Fig. 1 depicts the fusion of two sample measurements grids from different time steps. The measurement grids are derived from sensor data in a two-class Dempster–Shafer [12], [13] model consisting of free space (green) and occupied (purple). The measurement grids are fused with Dempster’s rule of combination. While the road boundaries are correctly fused as they are static, the L-shaped structures, which arise from the same dynamic object, cannot be correctly fused in this scenario with this approach.

III. MEASUREMENT SYNCHRONIZATION USING SENSOR DATA PREDICTION

More accuracy is achieved if the measurements are temporally aligned. Synchronization between different time instances is usually done by prediction onto the same time instance, such as in object fusion [2]. First, Section III-A shows the general case, while in Section III-B the case of radar sensor data is discussed.

A. Uniform Prediction of Measurements

Given a measurement \( z = (z_x, z_y) \) in a Cartesian coordinate system, the time difference \( \Delta t \) between the measurement time and the synchronization time, and an assumed
maximum speed of objects $v_{\text{max}}$, then the measurement

$$\begin{pmatrix} z_x \\ z_y \end{pmatrix} \rightarrow \begin{pmatrix} \tilde{z}_x \\ \tilde{z}_y \end{pmatrix} + f_{\text{cart}} \left( U \left[ 0, 0 \mid \Delta t \frac{v_{\text{max}}}{2\pi} \right] \right)$$  \hspace{1cm} (2)$$

is mapped using a 2-D uniform distribution $U(a, b)$ with a speed and an orientation component and lower bound $a$ and upper bound $b$. The function $f_{\text{cart}}(\cdot)$ transforms polar to Cartesian coordinates.

Although uniform prediction is general and does not require velocity measuring sensors, depending on the prediction time $\Delta t$ and the maximum assumed speed of objects $v_{\text{max}}$, objects are substantially inflated. Note that sensor measurement uncertainties are not yet included in (2) and further increase the resulting location uncertainty. This approach is therefore in general not useful and mostly given for completeness.

**B. Prediction of Measurements from Radar Sensor Data**

A radar sensor measures, in addition to a position component, the relative radial velocity, i.e., the 1-D projection of the true 2-D object velocity onto the measurement direction. The measurement $z = (z_r, z_\phi, z_{vR})$ consists of a range measurement $z_r$, an angular measurement $z_\phi$, and a measurement of the relative radial velocity $z_{vR}$. Throughout this work, it is assumed that the ego velocity is compensated in the radial velocity measurement and the absolute radial velocity $z_{vR}$ is used. The measurement in the global Cartesian reference coordinate system is denoted as $(s_x, s_y, s_{vR})$. The angular measurement $z_\phi$ in the reference coordinate system is denoted as $z_\phi$.

Given the radial velocity measurement, the prediction of the measurements can greatly be enhanced, as one component of the velocity is already measured. Similarly, the Doppler information can be used for initialization and weighting of velocity hypotheses in a particle filter for grid-based tracking [10], [11].

The radial velocity measurement is modeled as Gaussian distribution $\mathcal{N}(0, \sigma_{vR}^2)$ with standard deviation $\sigma_{vR}$. As nothing is known about the tangential velocity, it is modeled as uniform distribution assuming a maximum speed $v_{\text{max}}$ as before.

It is possible to limit the interval of the uniform distribution. The norm of the 2-D velocity $||v||$ is $v_{\text{max}}^2 + v_T^2$ is a combination of the radial velocity $v_R$ and the tangential velocity $v_T$. Since the speed $||v|| \leq v_{\text{max}}$ is assumed to be smaller than or equal to $v_{\text{max}}$, the maximum tangential velocity

$$v_T^{\text{max}}(v_R) = \sqrt{v_{\text{max}}^2 - v_R^2},$$

given one particular radial velocity $v_R$ with $|v_R| \leq v_{\text{max}}$ is used for the interval of the uniform distribution.

Let $(s_x, s_y)$ denote the position of the sensor in the used Cartesian coordinate system and $R_\alpha$ a rotation of $\alpha$ degrees. Assuming constant velocity during $\Delta t$, the measurement

$$\begin{pmatrix} z_x \\ z_y \end{pmatrix} \rightarrow \begin{pmatrix} s_x \\ s_y \end{pmatrix} +
R_\alpha \eta \left( \mathcal{N}(z_r + \Delta t z_{vR}, (\Delta t \sigma_{vR})^2) \right)$$  \hspace{1cm} (4)$$

is mapped using a combination of a Gaussian and a uniform distribution, as described above, where $\eta$ is a normalizer. Note that uncertainties of $z_r$ and $z_\phi$, as used in standard sensor models, are not included in (4) for simplicity.

Fig. 2 shows the 2-D distribution of velocities as shaded region given one particular radial velocity measurement $z_{vR}$. Note that with a uniform prediction, as described in the previous section, the whole circle represents the distribution of the true velocity. Hence, the uncertainty is greatly reduced. Nevertheless, the uncertainty in tangential direction, depending on $v_{\text{max}}$, $z_{vR}$, and the prediction time, may still be high. If the ego vehicle drives on a straight road, the tangential direction is close to perpendicular to the road boundary. This results in a high uncertainty towards the road center and, if modeled as such with a measurement grid, narrows the free space of the road on which the vehicle is driving, potentially even blocking it. Therefore, a heuristic-based approach is proposed in the next section for the prediction of radar measurements.
IV. POINT-TO-POINT RADAR MEASUREMENT PREDICTION WITH KNOWN OBJECT MOVING DIRECTION

The prediction of the previous section usually results in a high location uncertainty. In this section, it is shown how radar sensor data is predicted point-to-point-wise, if the object moving direction is known.

A. Prediction of Measurements in the Measurement Space

In this section, the prediction is performed directly in the polar measurement space \((r, \phi, v_R, v_T)\) rather than the Cartesian reference coordinate system \((x, y, v_x, v_y)\). It has the advantage that existing algorithms for the computation of measurement grids can directly be used. The sensor measurements at different time instances are simulated by assuming that all objects move with constant velocity. The sensor itself, however, and thus the ego vehicle, stays at a constant position.

It is started with the general prediction equation in Cartesian coordinates. Given the true object velocity \((v_x, v_y)\), the predicted measurement after \(\Delta t\) is directly available as

\[
\begin{pmatrix}
\bar{x} \\
\bar{y}
\end{pmatrix}
= \begin{pmatrix}
x \\
y
\end{pmatrix} + \Delta t \begin{pmatrix}
v_x \\
v_y
\end{pmatrix}.
\]

Similarly, in the polar measurement coordinate system, the predicted range and bearing

\[
\bar{r} = \sqrt{(r + \Delta t v_R)^2 + (\Delta t v_T)^2}
\]

\[
\bar{\phi} = \phi - \arctan\left(\frac{\Delta t v_T}{r + \Delta t v_R}\right)
\]

are also directly computable with the true radial and tangential velocity \((v_R, v_T)\). As before, all velocities are absolute.

If the angle of the predicted measurement changes, i.e., \(\bar{\phi} \neq \phi\), the measured radial velocity has to be modified according to the new measurement direction of the predicted measurement. It is derived as follows. Since constant velocity during \(\Delta t\) is assumed, the norm of the velocity vector

\[
\sqrt{v_R^2 + v_T^2} = \sqrt{\bar{v}_R^2 + \bar{v}_T^2}
\]

stays constant and the predicted radial velocity

\[
|\bar{v}_R| = \sqrt{v_R^2 + v_T^2 - \bar{v}_T^2}
\]

(8)

can be computed with the predicted tangential velocity

\[
\bar{v}_T = \frac{r \sin(\bar{\phi} - \phi)}{\Delta t} = \frac{v_T \bar{r}}{\bar{\phi}}.
\]

Putting (9) into (8) finally results in

\[
\bar{v}_R = \text{sgn}(v_R) \sqrt{v_R^2 + v_T^2 \left(1 - \frac{\bar{r}^2}{\bar{\phi}^2}\right)}.
\]

(10)

Geometric proofs of the given formulas are depicted in Fig. 3, which shows the relations.

As the radial velocity \(v_R\) is a measured quantity of radar sensors, the only missing quantity in (6) to perform the prediction is \(v_T\).

B. Prediction of Radar Measurements with an Estimated Moving Direction

If an estimation of the moving direction of the object at the position of the measurement is known, then the tangential velocity \(v_T\) can be calculated. This becomes clear, when looking at Fig. 4a, which shows the uncertainty of the predicted corresponding measurement after \(\Delta t\). The figure also shows the method described in Section III-B.

Geometrically, given an estimated moving direction \(\psi\) of the object, the predicted point is obtained as the intersection of the line of the tangential uncertainty and the line with angle \(\phi\) that goes through the position of the measurement, as shown in Fig. 4b. Given an estimated moving direction \(\psi\) and let

\[
\gamma = \psi - z_\alpha, \quad \gamma \in [-\pi, +\pi]
\]

denote the angle between the angle of the measurement in the reference coordinate system \(z_\alpha\) and the estimated moving direction \(\psi\) in the reference coordinate system. The estimated tangential velocity

\[
\hat{v}_T = z_{v_R} \tan \gamma, \quad |\gamma| = \pi/2
\]

(12)

can be computed for any estimated moving direction \(\psi\). The angle \(\gamma\) is derived to detect implausible moving direction estimations on \(\gamma\) are derived to detect implausible moving direction estimations.
direction. The radial velocities are still taken directly from the measurements.

4) Ego orientation heuristic: In many driving situations, in particular those including high object velocities such as highways, where synchronization in the form of prediction has the most influence, objects move roughly in parallel. With this assumption, \( \psi \) is given as the moving direction of the ego vehicle \( \psi_{ego} \). Oncoming traffic, also moving in parallel, can be modeled this way as well, if the sign of \( z_{vR} \) is taken into account such that \( \gamma \) holds the constraints in (13) for either \( \psi_{ego} \) or \( \psi_{ego} + \pi \).

Finally, it is noted that it is also possible to predict a measurement multiple times using different angles \( \psi \). Even an approximation of the uniform distribution in the tangential velocity of (4) is possible this way.

D. Error Analysis

The estimated moving direction \( \psi \) of an object may differ from its real moving direction \( \psi' \). If the estimated moving direction fulfills (13) and the sign of the measured radial velocity is correct, then the range of the possible orientation difference

\[
\Delta \psi = \psi' - \psi, \quad \Delta \psi \in [-\pi, +\pi]
\]

is reduced to the interval \( \Delta \psi \in [\xi, \xi + \pi] \), where the angle

\[
\xi = -\left(\left(\gamma + \frac{\pi}{2}\right) \mod \pi\right),
\]

is dependent on the angle \( \gamma \) between the measurement angle \( z_\alpha \) and the estimated moving direction \( \psi \) (11).

The absolute position error of the prediction

\[
\varepsilon_\psi = \Delta t \sqrt{v_{R}^2 + (z_{vR} \tan \gamma)^2} + (v_T - \bar{v}_T)^2,
\]

under constant velocity and orientation assumptions during \( \Delta t \), is used for the error analysis, since the motivation is a spatiotemporal alignment for accurately fusing measurement grids consisting of quadratic cells.

If measurement uncertainties of the radial velocity are neglected, i.e., \( z_{vR} = v_R \), and using (12), (17) is rewritten to

\[
\varepsilon_\psi = \Delta t \sqrt{v_{R}^2 \left(\tan(\gamma + \Delta \psi) - \tan \gamma\right)^2} + |\xi| \neq \frac{\pi}{2}
\]

showing that the error is dependent on the angles \( \Delta \psi \) and \( \gamma \).

Contrary, if no prediction is performed, then the absolute position error is given by

\[
\varepsilon_t = \Delta t \sqrt{v_{R}^2 + v_T^2} = \Delta t \sqrt{v_{R}^2 \left(1 + \tan^2(\gamma + \Delta \psi)\right)}.
\]

In this sense, the prediction with an estimated moving direction is better than performing no prediction if the inequation

\[
\varepsilon_\psi^2 \leq \varepsilon_t^2 \Rightarrow \tan^2 \gamma - 2 \tan \gamma \tan(\gamma + \Delta \psi) \leq 1
\]

holds. The geometric relations under an incorrect moving direction estimation are shown in Fig. 5.

If \( \gamma = 0 \) or \( \gamma = \pm \pi \), the inequation (20) is fulfilled for all possible \( \Delta \psi \), i.e., a prediction only in the radial direction is always better than performing no prediction in terms of
the absolute position error. On the other hand, the moving
direction estimation is error-prone for $|\gamma|$ close to $\pi/2$, as
the object velocity is close to the tangential velocity. To
detect and prevent large errors in these cases, the maximum
allowed tangential velocity is limited by a defined maximum
object velocity, as given in (14). If (14) does not hold, $\gamma$ can
be shrunk to the maximum allowed value, or, as described
before, set to 0, to predict at least in the radial direction,
which is still better than performing no prediction.

The absolute position error is illustrated in Fig. 6 for
three different scenarios. It can be seen that $\Delta \psi$ has to
be very large so that the proposed prediction approach is
worse than performing no prediction with respect to (20). For
highway scenario a difference of 55° between the estimated
or assumed object orientation and the true orientation, such
as in Fig. 6a, is not likely to occur. The ego orientation
heuristic may fail in urban scenarios, in particular where
large $\Delta \psi$ are possible. However, the object speeds and
thus the spatiotemporal misalignment are significantly lower.
In these scenarios, where the moving direction is quite
uncertain, a prediction only in the radial direction may be
preferred, which, in terms of the absolute position error, is
always better than performing no prediction.

V. RESULTS

Different approaches for spatiotemporal alignment have
been presented. The proposed choice of data prediction with
radar sensors is the point-to-point prediction described in
Section IV, where results are therefore given from in the
following. The heuristic of an equal moving direction as the
ego vehicle is used. Particularly on highways, this assump-
tion is justified as the difference between the real moving
direction of a vehicle and the ego orientation typically does
not exceed a few degrees.

In addition to the error analysis in Section IV-D, qualitative
results are given. The moving test vehicle was equipped with
4 laser scanners, one at each side of the vehicle, and 4 radar
sensors, one at each corner. The laser scanners measured with
an update rate of 25 Hz and the radar sensors with a slightly
lower update rate of 20 Hz. The radar sensors had a higher
latency than the laser scanners. The measurement time interval
alignment was not used to minimize the total latency.

Fig. 7 shows a highway scenario with 3 vehicles moving in
the same direction as the ego vehicle. The dark green squares
represent the data from the 4-layer laser scanners. The red
circles represent the original data from the radar sensors,
and the arrows denote the radial measurement directions and
their absolute velocities, i.e., the ego velocity is already
compensated as used throughout this work. The green circles
show the predicted radar data. Note that the odometry has
the lowest latency and hence the vehicle position is more
recent than the sensor data. Past vehicle poses at the time of
each measurement are available. The evidential measurement
grids from the radar sensor data and the laser scanner
data are fused using Dempster’s rule of combination. Large
errors are visible without sensor data alignment (Fig. 7c)
compared to the fused measurement grid of the aligned
data (Fig. 7d). Fig. 7e shows the evidential map from grid-
based tracking and mapping (GTAM) [9] using the correctly aligned fused measurement grid. Small differences between the ego moving direction used for the heuristic and those of the moving objects can be well coped with, as given in the figure due a lane change of the ego vehicle.

Fig. 8 shows a second scenario, which also exhibits an oncoming vehicle. As lanes are often parallel, the ego orientation heuristic also works for oncoming traffic, if the sign of the measured radial velocity is considered.

VI. CONCLUSIONS

This paper has presented different approaches for spatiotemporal alignment of asynchronous sensor data. First, an interval alignment approach has been shown, where the sensor data stays unmodified. Although it is insufficient for fast moving objects and sensors with low update rates, as the measurement data represents different time steps, it can still be used in addition to the subsequently presented approaches. Moreover, two general sensor data prediction approaches have been given, one for position measuring sensors and one for radar sensors. However, usually the uncertainty is too high resulting in high inflations of the measurement grids.

Therefore, a point-to-point prediction approach with radar sensors has been proposed, which is able to precisely align the data. Although it relies on a known object moving direction, an error analysis has demonstrated that only for very large deviations between the estimated and the true object orientation, the approach is worse than if no prediction is performed. Particularly for highway scenarios and by assuming an equal orientation as the ego vehicle, the approach has been proven to provide significant improvements.

REFERENCES