Heterogeneous Track-to-Track Fusion Using Equivalent Measurement and Unscented Transform

Christoph Allig
Corporate Research Department
DENSO AUTOMOTIVE Deutschland GmbH
Eching, Germany
c.allig@denso-auto.de

Gerd Wanielik
Professorship of Communication Engineering
Chemnitz University of Technology
Chemnitz, Germany
gerd.wanielik@etit.tu-chemnitz.de

***Abstract*—This document presents a novel track-to-track fusion (T2TF) approach for heterogeneous tracks. T2TF enables a distributed fusion structure, where tracks from local trackers are transmitted to a global tracker, that fuses the local tracks. T2TF offers the opportunity of low communication load almost without loss of information, compared to centralized measurement-to-track fusion (CMF). Heterogeneous tracks are of different state spaces, which are often non-linearly related. Heterogeneous track-to-track fusion (HT2TF) arises two challenges. Firstly, the fusion of tracks in different state spaces. Secondly, the cross-correlation in the state estimation errors. The presented HT2TF approach is based on the equivalent measurement and the unscented transform (UT). Compared to state-of-the-art approaches, the Jacobian is not required. For evaluation our approach is compared with the corresponding CMF.**

***Index Terms*—heterogeneous track-to-track fusion, high-level fusion, track-to-track fusion, unscented transform, equivalent measurement, unscented kalman filter**

I. Introduction

For fusion of data from multiple sensors there are multiple possible architectures. Low-level fusion architecture is characterized by transferring raw data from each sensor to a central fusion module. Since the fusion is performed solely in the central fusion module, a low-level fusion architecture is also called measurement-to-track, centralized, or central-level fusion [10]. Feature-level fusion is characterized by transferring extracted features from each sensor to the central fusion module, where the features are fused after feature-to-track association. High-level fusion is characterized by distributed tracking in each of the sensor and a central fusion combining the already tracked targets from multiple sensors, for which it is also called track-to-track, sensor-level, or distributed fusion [1].

Low-level fusion is optimal from an information-theory point of view, since no data reduction takes place. Therefore, the entire data is available at the central fusion module. When processing raw data, no model assumptions are made and data between sensors and over time are independent [6]. In feature-level fusion, features are extracted out of raw data, why the entire data is not available at the central fusion module. However, if the features are chosen reasonable, there is solely a slight information loss. In High-level fusion, data between sensors and over time is correlated, processed in an unknown manner and model assumptions are made [1]. One reason for data correlation between sensors is common process noise, used by the distributed tracking algorithms [2]. Neglecting data correlation between sensors may lead to excessive confidence for homogeneous tracks, whereas for heterogeneous tracks this may lead to excessive or underestimated confidence [16].

Data reduction is not performed in low-level fusion before transferring data from sensor to the central fusion. So low-level fusion results in high communication load and requires a high bandwidth, which is not feasible in many practical situations due to communication constraints. Feature-level fusion reduces communication load by transferring only features, which have been extracted out of raw data. High-level fusion reduces the communication load further, if the transfer frequency from local sensor to central fusion is reduced at the expense of a slightly increased latency of the central fusion result. In comparison to low-, or feature-level only small information loss takes place, if transfer frequency is reduced, since the measurements are still incorporated at the local sensor using full frequency. However high-level fusion requires more data to represent already tracked objects, due to describing object's state confidence and due to additional information gained through tracking over time, e.g. velocity.

In sensor architectures using wireless communication for transferring data, e.g. in vehicular communication systems, probability for message loss is comparable high, particularly in dense environments. Resending the message is unreasonable, since resending increases latency and communication load. Therefore, in feature-level fusion the information of messages, which are lost during transmitting, is completely lost. Whereas, in high-level fusion there is only small information loss, because in the distributed sensors the information is incorporated anyway. Sending the tracked target in the next message results only in small information loss, respectively latency of the central fusion result.

For several applications, e.g. for automotive cooperative environment perception, high-level fusion is preferred due to communication constraints or demanded modularity. Considering high-level fusion, there are two types of fusion strategies:

This work is part of the ADAS&ME European Project which has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 688900.

978-0-9964527-7-9 ©2018 ISIF

sensor-to-sensor fusion and sensor-to-global fusion.

In a sensor-to-sensor fusion strategy, also commonly referred to as track-to-track fusion (T2TF) *without* memory, sensor-level tracks are fused together at predefined fusion cycles and results in the global track. The previous global track is completely ignored.

In a sensor-to-global fusion strategy, also commonly referred to as T2TF *with* memory, a sensor-level track is fused into a global track as soon as the sensor-level track arrives. This type of strategy causes the problem that additionally to the common process noise, the sensor-level tracks are not independent over time, which results in a correlation due to common information history in the global track. Nevertheless, sensor-to-global fusion strategy has essential advantages. Firstly, the sensor-level tracks can be incorporated asynchronously and secondly the global track is maintained over time, which enables e.g. a smooth transition for objects that leave the field of view of one sensor and enter into another's.

High-level fusion is also distinguishable in homogeneous T2TF and heterogeneous track-to-track fusion (HT2TF). In case of homogeneous T2TF, the local sensors use the same state space. However, in practical applications local sensors may use different state spaces and system models, which is referred to as HT2TF. One reason for using distinct system models in the local trackers is that e.g. due to different sensor characteristics or depending on the application, a specific system model might be preferred. Considering e.g. automotive cooperative environment perception, heterogeneous tracks are inevitable due to various used coordinate systems, which are related non-linearly to each other.

The goal of this paper is to propose a novel approach for T2TF of heterogeneous tracks in a sensor-to-global strategy, which handles the cross-correlation. Section II provides an overview of corresponding state-of-the-art. In section III, we propose our approach for HT2TF with memory. A simulation scenario is described in section IV, which is used in section V to evaluate the proposed approach.

II. Related Work

Most approaches for T2TF assume that the local tracker use the same system model and state space, e.g. [5], [12], [14], [13], [9], [11] or [7]. Whereas, there are few works dealing with HT2TF. In [17] a linear minimum mean square error approach and a maximum likelihood approach for HT2TF in a sensor-to-sensor strategy are presented and compared with a corresponding centralized measurement-to-track fusion (CMF). For estimating the cross-correlation, a functional model is used, but it does not provide any perceivable benefit. So it is recommended to ignore the cross-correlation for HTTF without memory. Another unexpected result highlighted by this paper is that, when an interacting multiple model tracker is applied, the HT2TF approach may yield better performance compared to the CMF, because angle only measurements offer little information about the true mode of the target motion and may counterproductively confuse the interacting multiple model CMF. In [15] an algorithm for HT2TF in a sensor-to-global strategy is proposed based on the generalized information matrix fusion. The generalized information matrix fusion applies the equivalent measurement, defined in [3], in order to eliminate the cross-correlation, which greatly simplifies the fusion algorithm. An alternative for extracting the equivalent measurement is presented in [8]. The aforementioned HT2TF approaches require the computation of the Jacobian, which is difficult to derive in some domains. In contrast, our approach does not require the computation of the Jacobian. Moreover, in [16] an overview of architectures for T2TF is provided and the problem of HT2TF is discussed.

III. Unscented Kalman Filter using Equivalent Measurement

As mentioned above, HT2TF refers to fusion of tracks, where the local trackers may use different state spaces and system models. Consider local track li with mean $\vec{x}_{li}$ and covariance $\tilde{P}_{li}$ and local track lj with mean $\vec{x}_{lj}$ and covariance $\tilde{P}_{lj}$. Generally the system and observation model of local tracker li are

$$\begin{aligned}\vec{x}_{li}(k|k-1) &= f_{li}\left(\vec{x}_{li}(k-1)\right) + \vec{v}_{li}(k),\\ \vec{z}_{li}(k) &= h_{li}\left(\vec{x}_{li}(k|k-1)\right) + \vec{w}_{li}(k),\end{aligned}$$

and the system and observation model of local tracker lj are

$$\begin{aligned}\vec{x}_{lj}(k|k-1) &= f_{lj}\left((\vec{x}_{lj}(k-1)\right) + \vec{v}_{lj}(k),\\ \vec{z}_{lj}(k) &= h_{lj}\left(\vec{x}_{lj}(k|k-1)\right) + \vec{w}_{lj}(k).\end{aligned}$$

The local tracks are fused in a sensor-to-global fusion strategy into global track G with mean $\vec{x}_G$ and covariance $\tilde{P}_G$. Generally the system model of global tracker G can be written as

$$\vec{x}_G(k|k-1) = f_G\left(\vec{x}_G(k-1)\right) + \vec{v}_G(k),$$

and may have a non-linear relationship to local tracker li respectively local tracker lj

$$\begin{aligned}\vec{x}_{li} &= g\left(\vec{x}_G\right),\\ \vec{x}_{lj} &= g\left(\vec{x}_G\right).\end{aligned}$$

The transition functions f_{li}, f_{lj} and f_G as well as the transformation functions h_{li} and h_{lj} differ and can be non-linear. The process noise of the system model is represented by $\vec{v}_{li}(k)$, $\vec{v}_{lj}(k)$ respectively $\vec{v}_G(k)$. The measurement noise of the observation model is given by $\vec{w}_{li}(k)$ respectively $\vec{w}_{lj}(k)$. The local tracks are asynchronously processed as they arrive and the global track is maintained over time.

For instance, let $\vec{x}_G$ and $\vec{x}_{li}$ be identical and the larger dimension state, given by two-dimensional cartesian position, velocity and driving direction $\vec{x} = \begin{bmatrix} x & y & v & \phi \end{bmatrix}^T$. And let $\vec{x}_{lj}$ be the smaller dimension state, given by angular position and velocity $\vec{x} = \begin{bmatrix} \theta & \dot{\theta} \end{bmatrix}^T$.

The issue to be solved is how to incorporate the heterogeneous local tracks in a sensor-to-global fusion strategy into the global track in a reasonable way, particularly how to handle the cross-correlation. Our novel idea for the HT2TF with memory, is to update the global track using the unscented kalman filter (UKF) with an information gain. The idea behind is to decorrelate the information between the local tracks at k_l and k_{l-1} and only fuse the information that is new into the global track. The information gain is denoted by the equivalent measurement and its associated covariance matrix. The extraction of the information gain out of the current local track and the previously received local track as well as the incorporation of the information gain using the UKF is described subsequently.

The intention of the equivalent measurement is to represent all the measurements, processed by the local tracker since the last time a track from that local tracker was transmitted to the global tracker. According to [4], the necessity to perform the complex cross-covariance computations is eliminated using the equivalent measurement for T2TF. There are several approaches for determining the equivalent measurement. In [8], it is proposed to extract the equivalent measurement $\vec{z}_l'(k_G|k_l, k_l-1)$ from local tracker l out of the current local track l with mean $\vec{x}_l(k_G|k_l)$ and covariance $\tilde{P}_l(k_G|k_l)$ and the previously received local track l with mean $\vec{x}_l(k_G|k_l-1)$ and covariance $\tilde{P}_l(k_G|k_l-1)$ according to (1). In the current local track all the measurements received by the local tracker up to time k_l are incorporated. Whereas, in the previously received local track all the measurements received by the local tracker up to time k_l-1 are incorporated. It should be noted that k_l-1 can be an arbitrary time step before k_l. Both local tracks are predicted to the arrival time k_G of the current local track at the global tracker. In [1], it is proven that the prediction using a small prediction step does not effect the ability to properly extract the information gain.

$$\begin{aligned}\vec{z}_l'(k_G|k_l, k_l-1) &= \vec{x}_l(k_G|k_l) \\ &- \tilde{P}_l(k_G|k_l)\tilde{P}_l(k_G|k_l-1)^{-1}\vec{x}_l(k_G|k_l-1)\end{aligned} \tag{1}$$

Given a nonmaneuvering target, respectively no process noise, $\vec{z}_l'(k_G|k_l, k_l-1)$ is independent of the previously received track. The term $\tilde{P}_l(k_G|k_l)\tilde{P}_l(k_G|k_l-1)^{-1}\vec{x}_l(k_G|k_l-1)$ represents the portion of $\vec{x}_l(k_G|k_l)$ that is correlated with the previously transmitted data. Therefore, subtracting this term from $\vec{x}_l(k_G|k_l)$ results in the equivalent measurement $\vec{z}_l'(k_G|k_l, k_l-1)$. In order to update the global track using the equivalent measurement, the measurement matrix, the expected measurement and the measurement noise covariance matrix, defined in [8], are required.

In [3], it is proposed to use the canonical parametrization for determining the equivalent measurement $\vec{z}_l(k_G|k_l, k_l-1)$ that represents the new information obtained by track l since measurement k_l-1 according to

$$\begin{aligned}&\vec{z}_l(k_G|k_l, k_l-1) = \tilde{R}_l(k_G|k_l, k_l-1) \\ &* \left(\tilde{P}_l(k_G|k_l)^{-1}\vec{x}_l(k_G|k_l) - \tilde{P}_l(k_G|k_l-1)^{-1}\vec{x}_l(k_G|k_l-1)\right),\end{aligned} \tag{2}$$

where the measurement noise covariance matrix is

$$\begin{aligned}&\tilde{R}_l(k_G|k_l, k_l-1) \\ &= \left(\tilde{P}_l(k_G|k_l)^{-1} - \tilde{P}_l(k_G|k_l-1)^{-1}\right)^{-1}.\end{aligned} \tag{3}$$

We decided to extract the information gain according to (2) and (3), since unlike the method presented before, the measurement matrix is an identity matrix. In the following, our novel approach to apply the UKF for incorporation of the information gain is described.

The UKF utilizes the unscented transform (UT) in order to transform a Gaussian distribution through a non-linear function. The UT performs stochastic linearisation using the assumption of normal distributions. For updating the predicted global track by an incoming local track, a set of so-called sigma points is extracted from the predicted global track according to (4), with $\vec{\mu}$, $\tilde{\Sigma}$ and γ short for $\vec{x}_G\left(k_G|k_G-1\right)$, $\tilde{P}_G\left(k_G|k_G-1\right)$ and $\sqrt{n+\lambda}$, respectively. The sigma points are located at the mean and symmetrically along the main axes of the covariance. λ is defined as $\lambda = \alpha^2\left(n+\kappa\right) - n$ with α and κ being scaling parameters that determine how far the sigma points are spread from the mean. For an n-dimensional Gaussian, a set of sigma points comprises $2n+1$ sigma points, which are extracted according to

$$\begin{aligned}&\vec{X}_G^{[0]}\left(k_G|k_G-1\right) = \vec{\mu}, \\ &\vec{X}_G^{[i]}\left(k_G|k_G-1\right) = \vec{\mu} + \gamma\sqrt{\tilde{\Sigma}}_i, \qquad 1 \le i \le n, \\ &\vec{X}_G^{[i]}\left(k_G|k_G-1\right) = \vec{\mu} - \gamma\sqrt{\tilde{\Sigma}}_{i-n}, \quad n+1 \le i \le 2n.\end{aligned} \tag{4}$$

Each sigma point has two associated weights

$$\begin{aligned}w_m^{[0]} &= \frac{\lambda}{n+\lambda}, \\ w_c^{[0]} &= \frac{\lambda}{n+\lambda} + \left(1-\alpha^2+\beta\right), \\ w_m^{[i]} &= w_c^{[i]} = \frac{1}{2\left(n+\lambda\right)}, \quad 1 \le i \le 2n.\end{aligned}$$

One weight, $w_m^{[i]}$, is used when computing the mean, the other weight, $w_c^{[i]}$, is used when recovering the covariance of the Gaussian. The parameter β can be chosen to encode additional knowledge about the distribution.

These sigma points are then passed through the non-linear function

$$\vec{Z}_G^{[i]}\left(k_G|k_G-1\right) = g\left(\vec{X}_G^{[i]}\left(k_G|k_G-1\right)\right), \quad 0 \le i \le 2n,$$

which maps the state space of the global tracker to the state space of the local tracker. After that, mean

$$\hat{\vec{z}}_G\left(k_G|k_G-1\right) = \sum_{i=0}^{2n} w_m^{[i]}\vec{Z}_G^{[i]}\left(k_G|k_G-1\right),$$

and covariance

$$\begin{aligned}\tilde{S}' = \sum_{i=0}^{2n} w_c^{[i]} &\left(\vec{Z}_G^{[i]}\left(k_G|k_G-1\right) - \hat{\vec{z}}_G\left(k_G|k_G-1\right)\right) \\ &\left(\vec{Z}_G^{[i]}\left(k_G|k_G-1\right) - \hat{\vec{z}}_G\left(k_G|k_G-1\right)\right)^T,\end{aligned}$$

of the transformed sigma points are calculated, which represent the predicted global track in the state space of the local tracker. In order to represent the probability for measuring the equivalent measurement, it is required to add the measurement noise covariance matrix to $\tilde{S}'$. Using the equivalent measurement, defined in (2), the measurement noise covariance matrix is given by $\tilde{R}_l(k_G|k_l, k_l-1)$ and is considered as follows

$$\tilde{S} = \tilde{S}' + \tilde{R}_l(k_G|k_l, k_l-1).$$

The cross covariance is obtained by

$$\tilde{P}^{x,z} = \sum_{i=0}^{2n} w_c^{[i]} \left(\vec{X}_G^{[i]}(k_G|k_G-1) - \vec{x}_G(k_G|k_G-1)\right) \left(\vec{Z}_G^{[i]}(k_G|k_G-1) - \hat{\vec{z}}_G(k_G|k_G-1)\right)^T.$$

And the Kalman gain $\tilde{K}$ is defined as

$$\tilde{K} = \tilde{P}^{x,z}\tilde{S}^{-1}.$$

The global track is updated by incorporating the equivalent measurement $\vec{z}_l(k_G|k_l, k_l-1)$ according to

$$\vec{x}_G(k_G|k_G) = \vec{x}_G(k_G|k_G-1) + \tilde{K}\left(\vec{z}_l(k_G|k_l, k_l-1) - \hat{\vec{z}}_G(k_G|k_G-1)\right),$$
$$\tilde{P}_G(k_G|k_G) = \tilde{P}_G(k_G|k_G-1) - \tilde{K}\tilde{S}\tilde{K}^T.$$

Note that all of the tracks are predicted to the arrival time of the current local track beforehand, e.g. by using the prediction step of the UKF. As the decorrelation process occurs after the global track and the current and the previously received local tracks have been temporally aligned to the arrival time of the current local track, also out-of-sequence tracks can be properly integrated into the global track. The only disadvantage of the proposed approach is that, for each local tracker, the previously fused local track must be saved in memory to carry out the decorrelation process.

IV. Simulation

The performance of the proposed HT2TF approach is evaluated in a scenario with an active sensor and a passive sensor, tracking an object with constant velocity. The scenario is depicted in Fig. 1. The active sensor, located at $\begin{bmatrix} x_a & y_a \end{bmatrix}^T = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$, measures target range r_a, azimuth angle θ_a and range rate $\dot{r}_a$ with frequency $f_a = 4/3$ Hz

$$r_a = \sqrt{(x-x_a)^2 + (y-y_a)^2} + w_r,$$
$$\theta_a = \tan^{-1}\left(\frac{y-y_a}{x-x_a}\right) + w_a,$$
$$\dot{r}_a = v\cos\left(\phi - \tan^{-1}\left(\frac{y-y_a}{x-x_a}\right)\right) + w_{\dot{r}}.$$

The passive sensor, located at $\begin{bmatrix} x_p & y_p \end{bmatrix}^T = \begin{bmatrix} 10000 & 20000 \end{bmatrix}^T$, measures target azimuth angle θ_p with frequency $f_p = 3$ Hz

$$\theta_p = \tan^{-1}\left(\frac{y-y_p}{x-x_p}\right) + w_p.$$

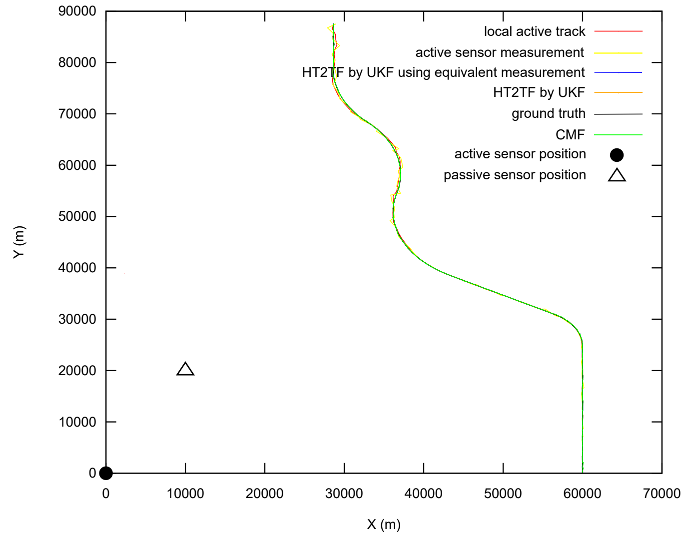

Fig. 1. In the simulation scenario, the target is observed by the active sensor and the passive sensor.

The respective measurement noise is described by w_r, w_a, $w_{\dot{r}}$ respectively w_p, which are zero mean white Gaussian noises with standard deviations $\sigma_r = 20$ m, $\sigma_a = 5$ mrad, $\sigma_{\dot{r}} = 10$ m/s respectively $\sigma_p = 1$ mrad. The target is moving with a constant speed of 750 m/s with initial state

$$\vec{x}(0) = \begin{bmatrix} x(0) \\ y(0) \\ v(0) \\ \phi(0) \end{bmatrix} = \begin{bmatrix} 60000 \\ 0 \\ 750 \\ 90^\circ \end{bmatrix}.$$

At $t = 33$ s the target starts a left turn with 6°/s for 10 s, then continues straight until $t = 63$ s, at which time it turns right with 3°/s for 25 s, then left with 3°/s for 20 s, then right with 3°/s for 15 s. After that it continues straight.

A local active tracker incorporates the measurements from the active sensor in order to track the target. Equally a local passive tracker incorporates the measurements from the passive sensor. The HT2TF by UKF using equivalent measurement is designed as the global tracker and fuses the tracks of both local tracker in a sensor-to-global strategy applying our proposed approach. As a first reference, the centralized measurement-to-track fusion (CMF) directly integrates the measurements from the active sensor and the passive sensor in a centralized measurement-to-track fusion architecture. As a second reference, the HT2TF by UKF simply treats the local tracks as measurements and updates the global track by incorporating the local tracks using the UKF, while assuming independence, which is in fact not the case.

As system model in the local active tracker, in the global tracker and in the first and second reference, a constant turn rate and acceleration (CTRA) model is applied if $v_{\dot{\phi}} \neq 0$. Alternatively, a constant acceleration (CA) model is applied. The state space is given by two-dimensional cartesian position, velocity and driving direction $\vec{x} = \begin{bmatrix} x & y & v & \phi \end{bmatrix}^T$. If $v_{\dot{\phi}} \neq 0$, the previous state $\vec{x}(k-1)$ is predicted according to (5), where the second term describes the estimated change in the state.

$$\vec{x}(k) = \vec{x}(k-1) + \begin{bmatrix} \Delta x \\ \Delta y \\ v_a t \\ v_{\dot{\phi}} t \end{bmatrix} \tag{5}$$

with

$$\Delta x = \frac{1}{v_{\dot{\phi}}^2}\left[v_a\left(cos(\phi + v_{\dot{\phi}}t) - cos(\phi)\right)\right] + \frac{1}{v_{\dot{\phi}}}\left[(v + v_a t)\, sin(\phi + v_{\dot{\phi}}t) - vsin(\phi)\right],$$

and

$$\Delta y = \frac{1}{v_{\dot{\phi}}^2}\left[v_a\left(sin(\phi + v_{\dot{\phi}}t) - sin(\phi)\right)\right] - \frac{1}{v_{\dot{\phi}}}\left[(v + v_a t)\, cos(\phi + v_{\dot{\phi}}t) - vcos(\phi)\right].$$

The parameters v_a and $v_{\dot{\phi}}$ are zero mean white Gaussian noises with standard deviations $\sigma_a = 10$ m/s and $\sigma_{\dot{\phi}} = 0.1$ rad/s and describe the process noise of the acceleration and the process noise of the yaw rate, respectively.

As system model in the local passive tracker, a constant angle rate motion model is applied. The state space is given by angular position and velocity $\vec{x} = \begin{bmatrix}\theta & \dot{\theta}\end{bmatrix}^T$. The previous state $\vec{x}(k-1)$ is predicted according to (6), where the second term describes the estimated change in the state.

$$\vec{x}(k) = \vec{x}(k-1) + \begin{bmatrix} \dot{\theta}t + 0.5 v_{\ddot{\theta}} t^2 \\ v_{\ddot{\theta}} t \end{bmatrix} \tag{6}$$

The parameter $v_{\ddot{\theta}}$ is a zero mean white Gaussian noise with standard deviation $\sigma_{\ddot{\theta}} = 0.1$ rad/s^2 and describe the process noise of the angular acceleration.

For incorporating the local active track into the global track, using a homogeneous T2TF approach is sufficient, since both tracker use the identical state space. Here the information matrix fusion is selected.

The state space and the system model of the global tracker and the local passive tracker are distinct. Therefore a HT2TF approach is required for incorporating the local passive track into the global track and here we apply our proposed approach. The state space of the global tracker is the larger dimension state space, compared to the state space of the local passive tracker, which means that a track represented in the global tracker's state space is convertible to the local passive tracker's state space, applying the following non-linear relationship

$$\theta = \tan^{-1}\left(\frac{y - y_p}{x - x_p}\right),$$
$$\dot{\theta} = \frac{vsin(\phi)}{\sqrt{(x - x_p)^2 + (y - y_p)^2}}.$$

The incorporation of the local active track respectively the local passive track is performed asynchronously at arrival time.

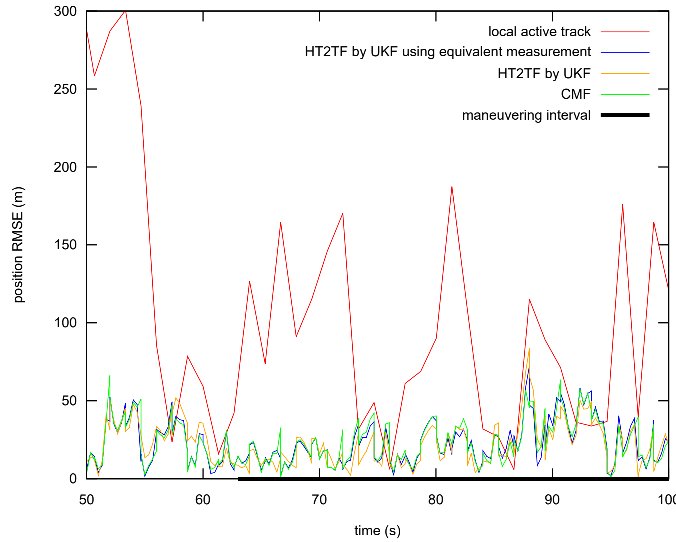

Fig. 2. RMSE for the position for 50 s < t < 100 s of the global tracker, the local active tracker and the two references.

V. Simulation Results

In this section the performance of the algorithm, presented in section III, is evaluated. As first reference the CMF is used, which is widely considered to be the most accurate method of sensor data fusion. As second reference the conventional UKF is applied. The performance is evaluated by the tracker's error and the tracker's consistency. The tracker's error is given by calculating the difference between the true state and the estimated state. A commonly used metric for the tracker's error is the root mean square error (RMSE). The tracker's consistency describes how well the estimated covariance represents the tracker's error. A commonly used metric for the tracker's consistency is the normalized estimation error squared (NEES). The NEES is χ^2-distributed with n_x degrees of freedom, where n_x is the dimension of the state space. An estimate is tested for consistency by determining if the NEES lies within an interval.

In Fig. 2, the RMSE for the global tracker are compared with those for the local active tracker and the two references in position for 50 s < t < 100 s, for an easier comparison. It can be seen that the global tracker and the two references provide a significantly improved estimation performance than the local active tracker. Moreover, the global tracker and the CMF achieve approximately identical estimation performance. Apparently the HT2TF by UKF's error performance differs slightly.

Fig. 3 shows the NEES of the the global tracker and the two references for 50 s < t < 100 s, for an easier comparison. Each of the NEES lies almost always within the 99% confidence region. The NEES of the HT2TF by UKF differs from the other two tracker. The NEES of the global tracker is comparable to the NEES of the CMF. We conclude that our proposed approach achieves comparable consistency to a CMF. As stated previously, a CMF is widely considered

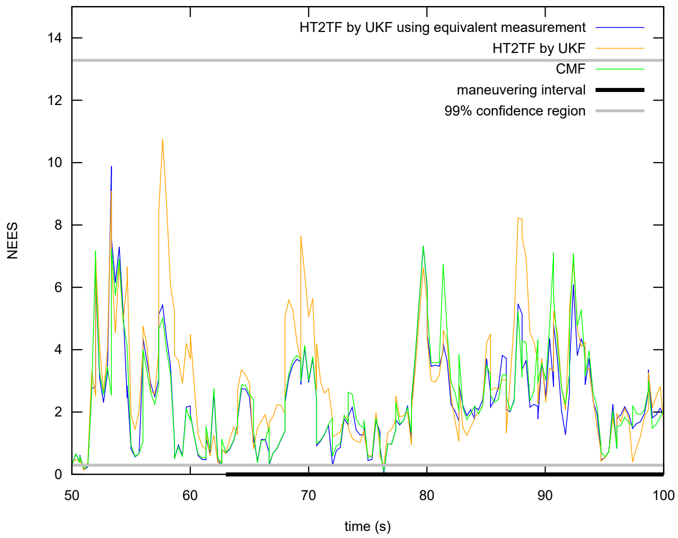

Fig. 3. NEES for 50 s $< t <$ 100 s of the the global tracker and the two references.

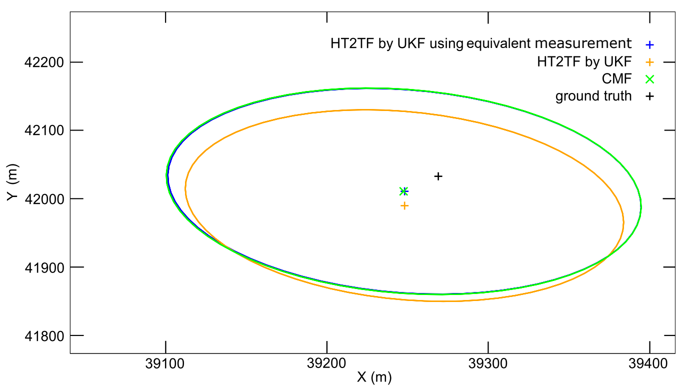

Fig. 4. Position estimation and its associated uncertainty of the global tracker and the two references. The ground truth is displayed additionally.

to be the most accurate method of sensor data fusion. Therefore we also conclude that our approach is capable of handling the cross-correlation by the equivalent measurement correctly, since otherwise the tracker would provide excessive or underestimated confidence due to data correlation, which would result in different consistency performance than the CMF. Moreover, Fig. 4 supports this claim. It shows an arbitrarily chosen position estimate of the global tracker and the two references and its associated position uncertainty. The ground truth is displayed additionally. The HT2TF by UKF's position estimate differs slightly from the other two, which are coincident. The global tracker's position uncertainty is identical to the CMF's. So our claim is supported. The HT2TF by UKF's position uncertainty is smaller, why this tracker affects here an excessive confidence due to ignored data correlation. The conspicuities, as just described, yield mostly for the entire simulation.

VI. Conclusion

A novel approach for HT2TF with memory has been presented, based on the UKF using the equivalent measurement. Similarly to the information matrix fusion for the homogeneous T2TF, our approach does not require the knowledge of the cross covariance between the heterogeneous tracks for the fusion. Instead the information gain from the local track is fused with the global track as if they were independent, which greatly simplifies the fusion algorithm. Compared to state-of-the-art HT2TF approaches, our approach does not require the knowledge of the Jacobian. In a simulation, we have shown that our proposed approaches error and consistency performance is comparable to a CMF, which is widely considered to be the most accurate method of sensor data fusion. Moreover the simulation has shown that the single sensor track's error performance is significantly inferior and a HT2TF, which ignores correlation, affects mostly excessive confidence. Future work will focus on applying the proposed algorithm for automotive cooperative environment perception and on track-to-track association.

References

[1] M. Aeberhard. *Object-Level Fusion for Surround Environment Perception in Automated Driving Applications*. VDI Verlag GmbH, 2017.

[2] Y. Bar-Shalom. On the Track-to-Track Correlation Problem. *IEEE Transactions on Automatic Control*, 26(2):571—-572, 1981.

[3] B. Belkin, S. L. Anderson, and K. M. Sommar. The pseudo-measurement approach to track-to-track data fusion. In *Proceedings 1993 Joint Service Data Fusion Symposium*, pages 519—-538, 1993.

[4] S. Blackman and R. Popoli. *Design and analysis of modern tracking systems*. Norwood, MA: Artech House, 1999.

[5] K. C. Chang, R. K. Saha, and Y. Bar-Shalom. On optimal track-to-track fusion. *IEEE Transactions on Aerospace and Electronic Systems*, 33(4):1271—-1276, 1997.

[6] M. Darms and H. Winner. A modular system architecture for sensor data processing of ADAS applications. In *IEEE Intelligent Vehicles Symposium*, volume 2005, pages 729—-734, 2005.

[7] D. Franken and A. Hüpper. Improved fast covariance intersection for distributed data fusion. In *Proceedings of the 7th International Conference on Information Fusion*, pages 154—-160. IEEE, 2005.

[8] G. Frenkel. Multisensor tracking of ballistic targets. In *Signal and Data Processing of Small Targets 1995*, pages 337—-346. International Society for Optics and Photonics, 1995.

[9] S. J. Julier and J. K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *SPIE Conference on Signal Processing, Sensor Fusion, and Target Recognition*, pages 182—-193, jul 1997.

[10] L. A. Klein. *Sensor and data fusion concepts and applications*. Society of Photo-Optical Instrumentation Engineers, Bellingham, 1993.

[11] W. Niehsen. Information fusion based on fast covariance intersection filtering. In *Proceedings of the Fifth International Conference on Information Fusion.*, volume 2, pages 901—-904, 2002.

[12] R. K. Saha and K. C. Chang. An efficient algorithm for multisensor track fusion. *IEEE Transactions on Aerospace and Electronic Systems*, 34(1):200—-210, 1998.

[13] J. Speyer. Computation and transmission requirements for a decentralized linear-quadratic-Gaussian control problem. *IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes.*, 17:1126—-1131, 1978.

[14] X. Tian. Track-to-Track Fusion Configurations and Association in a Sliding Window. *Journal of advances in information fusion*, 4(1):146—-164, 2009.

[15] X. Tian, Y. Bar-Shalom, T. Yuan, E. Blasch, K. Pham, and G. Chen. A Generalized Information Matrix Fusion Based Heterogeneous Track-to-Track Fusion Algorithm. In *SPIE Conference on Signal Processing, Sensor Fusion, and Target Recognition*, 2011.

[16] X. Tian, T. Yuan, and Y. Bar-shalom. Track-to-Track Fusion in Linear and Nonlinear Systems. In *Advances in Estimation, Navigation*, pages 21—-41. 2015.

[17] T. Yuan, Y. Bar-Shalom, and X. Tian. Heterogeneous Track-To-Track Fusion. In *Proceedings of the 14th International Conference on Information Fusion*, pages 1—-8, 2011.