Track-to-Track Fusion Using Split Covariance Intersection Filter-Information Matrix Filter (SCIF-IMF) for Vehicle Surrounding Environment Perception

Hao Li, Fawzi Nashashibi, Benjamin Lefaudeux, Evangeline Pollard

Track-to-Track Fusion Using Split Covariance Intersection Filter-Information Matrix Filter (SCIF-IMF) for Vehicle Surrounding Environment Perception

Hao LI, Fawzi NASHASHIBI, Benjamin LEFAUDEUX, Evangeline POLLARD

Abstract— Vehicle surrounding environment perception is an important process for many applications. Nowadays, a tendency is to incorporate redundant and complementary sensors into an intelligent vehicle, in order to enhance its perception ability; then an essential issue arises naturally, i.e. what fusion architecture can be used to combine the data from multiple sensors? In this paper, we propose a new track-to-track fusion architecture using the split covariance intersection filter-information matrix filter (SCIF-IMF). The basic idea is to use the IMF (adapted for estimates in split form) to handle the track temporal correlation of each sensor system and to use the SCIF to handle track spatial correlation. The proposed architecture enjoys complete sensor modularity and thus enables flexible self-adjustment. A simulation based comparative study is presented, which shows that the track-to-track fusion architecture using the SCIF-IMF can achieve centralized architecture comparable performance.

I. Introduction

For an intelligent vehicle, the surrounding environment perception is an important process for many applications. For example, in full automated mode, it is a prerequisite for crucial operations such as object avoidance [1-2]. Besides, the perception technologies can be well adapted for driver assistance purpose.

As the price of perceptive sensors decreases, a multi-sensor configuration becomes more and more economically feasible. Naturally, a tendency is to incorporate redundant and complementary sensors into an intelligent vehicle, in order to enhance its perception ability. A vehicle can be equipped with sensors perceiving forward, rearward, and sideward, which form omnidirectional view field [3-4]. It can be equipped with multiple sensors perceiving the same direction [5]. Perceptive sensors of different characteristics, when used together, might well overcome the shortcomings of each other, such as in the case of the laser scanner-camera cooperation [6-7].

A vehicle perception system concerns several essential issues, such as sensor calibration, low-level feature extraction, object detection (and recognition), object tracking etc. When it consists of multiple sensors, one more essential issue arises, i.e. what fusion architecture can be used to combine the data from all these sensors?

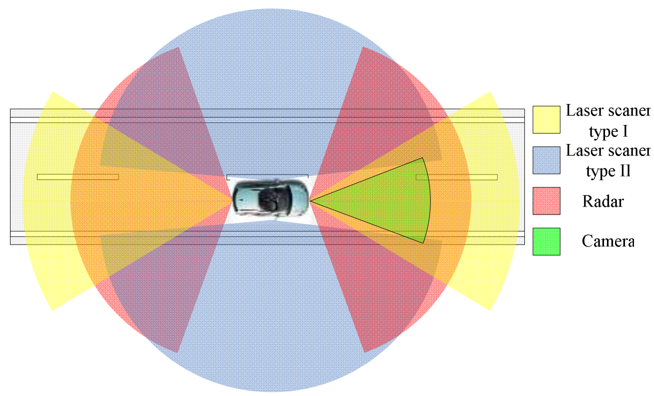

Fig.1. Redundant and complementary perceptive sensor configurations

A typical strategy for fusing several sensors often consists in sensor-level (low-level) fusion that is realized in a centralized way [8-11]: only one fusion center (or estimation center) exists for all the sensors. Sensor data are directly forwarded to the fusion center, which outputs a global estimate for object tracks. A centralized *Kalman filter* (KF) architecture is a typical example for sensor-level fusion.

In contrast, track-level (high-level) fusion or track-to-track fusion tends to be a desirable solution, thanks to its modularity and flexibility [12]. In this architecture, a sensor system independently forms its object track estimate and operates in distributed way. A track-to-track fusion center fuses different sensor tracks into a unified global track.

A strategy for track-to-track fusion is to use a global KF [13-14], where sensor tracks are treated as measurements to this global KF and fused. This strategy, which neglects track temporal correlation and track spatial correlation, suffers from inconsistent estimation— *track temporal correlation* means the correlation among the states of a sensor track at different instants, whereas *track spatial correlation* means the correlation among different sensor tracks, which can be caused by common process noise and common *a priori*.

Another strategy for track-to-track fusion is to generate a global track by fusing sensor tracks of only current fusion cycle [15-16]; previously generated global track is discarded and not used at all. A major drawback of this strategy lies in the global track discontinuity when an object traverses the view boundary of a sensor. Leaving the view field of a

H. Li, B. Lefaudeux, E. Pollard, and F. Nashashibi are with the IMARA team, Institut National de Recherche en Informatique et en Automatique (INRIA), France (e-mail: hao.li@inria.fr, benjamin.lefaudeux@inria.fr, evangeline.pollard@inria.fr, fawzi.nashashibi@inria.fr)

sensor can abruptly degrade the global track estimate.

Recently, the authors in [17] propose a track-to-track fusion architecture using the information matrix filter (IMF), which can well handle track temporal correlation. To handle track spatial correlation, this method adopts traditional compensation algorithms [18-19] that require computing the cross-covariance among different sensor tracks. In this sense, different sensor systems are still coupled with each other, which influences their modularity.

In this paper, we propose a new track-to-track fusion architecture using the split covariance intersection filter-information matrix filter (SCIF-IMF). The basic idea is: the IMF (adapted for estimates in split form) is used to handle track temporal correlation; the SCIF is used to handle track spatial correlation, without computing any cross-covariance among different sensor tracks.

The proposed method has two major merits: first, it can handle both track temporal correlation and track spatial correlation. Second, it enjoys complete sensor modularity; each sensor system is independent and no coupling process exists between different sensor systems.

The SCIF and the IMF are briefly reviewed in Section II; the track-to-track fusion architecture using the SCIF-IMF is described in Section III; Simulation tests are presented in Section IV, followed by a conclusion in Section V.

II. SPLIT COVARIANCE INTERSECTION FILTER AND INFORMATION MATRIX FILTER

A. *Split Covariance Intersection Filter (SCIF)*

Given an estimate $\{\mathbf{X}, \mathbf{P}\}$ where $\mathbf{X}$ denotes the estimated state vector and $\mathbf{P}$ denotes the estimated covariance matrix. Let $\mathbf{P}^*$ denote the true covariance of $\mathbf{X}$. The estimate is called *consistent* if $\mathbf{P} \geq \mathbf{P}^*$; in simple words, an estimate is consistent if it is not over-confident.

Given two consistent source estimates $\{\mathbf{X}_i,\mathbf{P}_i\}(i=1,2)$; if the fusion estimate is also consistent, the fusion is consistent. We hope that the fusion consistency can always be guaranteed, because we do not want to establish any extra confidence on the fusion estimate than what the source estimates can convey. Consider the Kalman Filter [20], which can be equivalently given as:

$$\begin{aligned}
\mathbf{P}^{-1} &= \mathbf{P}_1^{-1} + \mathbf{P}_2^{-1} \\
\mathbf{X} &= \mathbf{P}(\mathbf{P}_1^{-1}\mathbf{X}_1 + \mathbf{P}_2^{-1}\mathbf{X}_2)
\end{aligned}$$

The fusion consistency of the Kalman Filter can not be guaranteed, if source estimates are correlated. The authors in [21] propose a data fusion method named Covariance Intersection Filter (CIF), which forms the fusion estimate by taking a convex combination of the source estimates. The CIF is guaranteed to yield consistent fusion estimate even when facing source estimates of unknown correlation. However, the CIF has a drawback of yielding pessimistic estimate, because it treats the source estimates as totally correlated and neglects possible independent information in them.

In [22], the Split Covariance Intersection Filter (SCIF) is introduced, which provides the ability to handle both known independent information and unknown correlated information in the estimates.

For the SCIF, an estimate is always represented in split form $\{\mathbf{X}, \mathbf{P}_d+\mathbf{P}_i\}$, where the covariance component $\mathbf{P}_d$ represents the maximum degree to which the estimate is possibly correlated with others; the covariance component $\mathbf{P}_i$ represents the degree of its absolute independence. Given two source estimates $\{\mathbf{X}_1, \mathbf{P}_{1d}+\mathbf{P}_{1i}\}$ and $\{\mathbf{X}_2, \mathbf{P}_{2d}+\mathbf{P}_{2i}\}$; the fusion estimate $\{\mathbf{X}, \mathbf{P}_d+\mathbf{P}_i\}$ obtained via the SCIF is given as:

$$\begin{aligned}
\mathbf{P}_1 &= \mathbf{P}_{1d} / w + \mathbf{P}_{1i} \\
\mathbf{P}_2 &= \mathbf{P}_{2d} / (1-w) + \mathbf{P}_{2i} \\
\mathbf{P}^{-1} &= \mathbf{P}_1^{-1} + \mathbf{P}_2^{-1} \\
\mathbf{X} &= \mathbf{P}(\mathbf{P}_1^{-1}\mathbf{X}_1 + \mathbf{P}_2^{-1}\mathbf{X}_2) \\
\mathbf{P}_i &= \mathbf{P}(\mathbf{P}_1^{-1}\mathbf{P}_{1i}\mathbf{P}_1^{-1} + \mathbf{P}_2^{-1}\mathbf{P}_{2i}\mathbf{P}_2^{-1})\mathbf{P} \\
\mathbf{P}_d &= \mathbf{P} - \mathbf{P}_i
\end{aligned} \tag{1}$$

The w belongs to the interval [0,1]. In practice, w can be determined by optimizing an objective function in terms of w such as the determinant of the new covariance [22].

Suppose that $\mathbf{X}_1$ is complete observation i.e. $\mathbf{X}_1=\mathbf{X}_{true}$, whereas $\mathbf{X}_2$ is complete or partial observation i.e. $\mathbf{X}_2=\mathbf{H}\mathbf{X}_{true}$ ($\mathbf{H}$ is the observation matrix). Then the SCIF for this general case can be derived as:

$$\begin{aligned}
\mathbf{P}_1 &= \mathbf{P}_{1d} / w + \mathbf{P}_{1i} \\
\mathbf{P}_2 &= \mathbf{P}_{2d} / (1-w) + \mathbf{P}_{2i} \\
\mathbf{K} &= \mathbf{P}_1\mathbf{H}^T(\mathbf{H}\mathbf{P}_1\mathbf{H}^T + \mathbf{P}_2)^{-1} \\
\mathbf{X} &= \mathbf{X}_1 + \mathbf{K}(\mathbf{X}_2 - \mathbf{H}\mathbf{X}_1) \\
\mathbf{P} &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_1 \\
\mathbf{P}_i &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_{1i}(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{P}_{2i}\mathbf{K}^T \\
\mathbf{P}_d &= \mathbf{P} - \mathbf{P}_i
\end{aligned} \tag{2}$$

Notice that the SCIF in (2) can be regarded as a generalization of the Kalman Filter: let $\mathbf{P}_{1d}$ and $\mathbf{P}_{2d}$ be zero and (2) will become the same to the Kalman Filter. In other words, the Kalman Filter can be treated as a special case of the SCIF, where the source estimates are known to be independent.

B. *Information Matrix Filter (IMF)*

Given two source estimates $\{\mathbf{X}_i,\mathbf{P}_i\}(i=1,2)$; suppose they share some common information represented as $\{\mathbf{X}_0,\mathbf{P}_0\}$. The IMF was first proposed in [23] and can be written as:

$$\begin{aligned}
\mathbf{P}^{-1} &= \mathbf{P}_1^{-1} + \mathbf{P}_2^{-1} - \mathbf{P}_0^{-1} \\
\mathbf{X} &= \mathbf{P}(\mathbf{P}_1^{-1}\mathbf{X}_1 + \mathbf{P}_2^{-1}\mathbf{X}_2 - \mathbf{P}_0^{-1}\mathbf{X}_0)
\end{aligned} \tag{3}$$

We adapt the IMF in (3) for estimates in split form, as follows:

$$
\begin{aligned}
\mathbf{P}^{-1} &= \mathbf{P}_1^{-1} + \mathbf{P}_2^{-1} - \mathbf{P}_0^{-1} \\
\mathbf{X} &= \mathbf{P}(\mathbf{P}_1^{-1}\mathbf{X}_1 + \mathbf{P}_2^{-1}\mathbf{X}_2 - \mathbf{P}_0^{-1}\mathbf{X}_0) \\
\mathbf{P}_i &= \mathbf{P}(\mathbf{P}_1^{-1}\mathbf{P}_{1i}\mathbf{P}_1^{-1} + \mathbf{P}_2^{-1}\mathbf{P}_{2i}\mathbf{P}_2^{-1} - \mathbf{P}_0^{-1}\mathbf{P}_{0i}\mathbf{P}_0^{-1})\mathbf{P} \\
\mathbf{P}_d &= \mathbf{P} - \mathbf{P}_i
\end{aligned} \tag{4}
$$

III. Track-to-Track Fusion using the SCIF-IMF

In this section, we describe the track-to-track fusion architecture using the SCIF-IMF, which is illustrated in Fig.2. A sensor system independently processes its own data in the SCIF framework and outputs its track estimate. A high-level process of track-to-track fusion takes in the sensor tracks and fuses them in the SCIF-IMF framework.

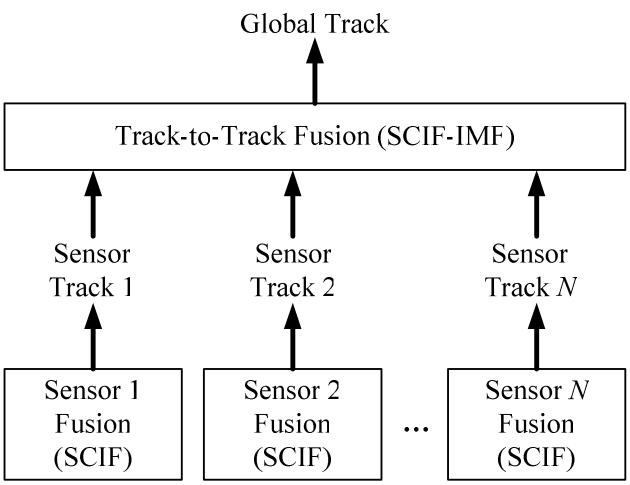

Fig.2. Track-to-track fusion architecture using the SCIF-IMF

Before continuing the description, it is worthy noting that the track-to-track fusion architecture in Fig.2 is only a simplified architecture. This architecture omits some elements such as object detection, temporal and spatial alignment of sensor tracks, track-to-track association, which are important for the whole perception system but are out of the range of this paper. In this paper, we would rather focus on the track-to-track fusion architecture without entangling it with other issues that are not essentially related to the fusion aspect. Therefore, without loss of generality, we assume: 1) object observations (position measurement) can be provided by the preprocessing unit of a sensor system; 2) sensor tracks can be temporally and spatially aligned; 3) track-to-track association can be performed.

A. *The fusion at a sensor system*

For a generic sensor system, it independently estimates the object state in a two-steps recursion: 1) the state is predicted (or evolved) to the measurement time; 2) the state is updated with the measurement, where the predicted state (*a priori*) is fused with the measurement to form the updated state (*a posteriori*). A commonly used method to fuse the predicted state and the measurement is the Kalman filter.

In the proposed architecture, we use the SCIF instead of the Kalman filter to fuse the predicted state and the measurement. In this sub-section, we describe the formalism from the perspective a generic sensor system; therefore, we temporarily use general denotation, without distinguishing different sensor systems. Let the object state be denoted in split form as $\{\mathbf{X}(t), \mathbf{P}_d(t)+\mathbf{P}_i(t)\}$, where $\mathbf{P}_d(t)$ and $\mathbf{P}_i(t)$ represent respectively the covariance correlated part and the covariance independent part; t denotes the period index.

State prediction (evolution):

$$\mathbf{X}(t\mid t-1) = \mathbf{F}(t\mid t-1)\mathbf{X}(t-1) \tag{5a}$$

$$\mathbf{P}_d(t\mid t-1) = \mathbf{F}(t\mid t-1)\mathbf{P}_d(t-1)\mathbf{F}(t\mid t-1)^T + \mathbf{Q}(t\mid t-1) \tag{5b}$$

$$\mathbf{P}_i(t\mid t-1) = \mathbf{F}(t\mid t-1)\mathbf{P}_i(t-1)\mathbf{F}(t\mid t-1)^T \tag{5c}$$

Where $\mathbf{F}(t|t\text{-}1)$ represents a basic constant acceleration kinematic model; $\mathbf{X}(t\text{-}1)$ and $\{\mathbf{P}_d(t\text{-}1)+\mathbf{P}_i(t\text{-}1)\}$ respectively are the previously fused state and covariance; $\mathbf{Q}(t|t\text{-}1)$ is the process noise matrix. As shown in (5b) and (5c), both the covariance correlated component $\mathbf{P}_d$ and the covariance independent component $\mathbf{P}_i$ are evolved, which enables maintaining the known independent information part and the information part that is possibly correlated.

State update (fusion):

The object state is updated with the new measurement. Let the measurement, the measurement noise matrix, and the observation matrix be respectively denoted as $\mathbf{Z}(t)$, $\mathbf{R}(t)$, and $\mathbf{H}$. The SCIF in (2) is used to fuse the predicted state and the measurement, as follows (noting that the measurement can be fairly assumed independent, so the new covariance will be optimized always when w is set to 1):

$$
\begin{aligned}
\mathbf{P}_1 &= \mathbf{P}_d(t\mid t-1) + \mathbf{P}_i(t\mid t-1) \\
\mathbf{K} &= \mathbf{P}_1\mathbf{H}^T(\mathbf{H}\mathbf{P}_1\mathbf{H}^T + \mathbf{R}(t))^{-1} \\
\mathbf{X}(t) &= \mathbf{X}(t\mid t-1) + \mathbf{K}(\mathbf{Z}(t) - \mathbf{H}\mathbf{X}(t\mid t-1)) \\
\mathbf{P}(t) &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_1 \\
\mathbf{P}_i(t) &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_i(t\mid t-1)(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}(t)\mathbf{K}^T \\
\mathbf{P}_d(t) &= \mathbf{P}(t) - \mathbf{P}_i(t)
\end{aligned} \tag{6}
$$

The estimate $\{\mathbf{X}(t), \mathbf{P}(t)\}$ (the entire covariance $\mathbf{P}_d(t)+\mathbf{P}_i(t)$) obtained via the SCIF is the same to that obtained via the Kalman filter. The difference is: the SCIF can provide extra information about the degree of the estimate being independent and the maximum degree of the estimate being possibly correlated.

B. *The track-to-track fusion*

Given N sensor systems which are indexed from 1 to N; let sensor track j (the object state estimated by sensor system j) be denoted as $\{\mathbf{X}_j(t_j), \mathbf{P}_{d,j}(t_j)+\mathbf{P}_{i,j}(t_j)\}$, where t_j denotes the period index for sensor system j ($j=1,\ldots,N$). Let the global track be denoted as $\{\mathbf{X}_G(t_G), \mathbf{P}_{d,G}(t_G)+\mathbf{P}_{i,G}(t_G)\}$.

Generically, suppose a new track update from sensor system j at time t_j is available for global track update at time

t_G. The state and covariance (in split form) of the global track are predicted (evolved) to the new track arrival time t_G:

$$\mathbf{X}_G(t_G \mid t_G-1) = \mathbf{F}(t_G \mid t_G-1)\mathbf{X}_G(t_G-1) \tag{7a}$$

$$\mathbf{P}_{d,G}(t_G \mid t_G-1) = \mathbf{F}(t_G \mid t_G-1)\mathbf{P}_{d,G}(t_G-1)\mathbf{F}(t_G \mid t_G-1)^T + \mathbf{Q}(t_G \mid t_G-1) \tag{7b}$$

$$\mathbf{P}_{i,G}(t_G \mid t_G-1) = \mathbf{F}(t_G \mid t_G-1)\mathbf{P}_{i,G}(t_G-1)\mathbf{F}(t_G \mid t_G-1)^T \tag{7c}$$

Then the global track is updated with the new track from sensor system j at time t_j, according to the following cases.

Case I: The new sensor track is fused into the global track for the first time: there is no track temporal correlation concerning this sensor; the SCIF in (2) is performed to handle possible track spatial correlation:

$$\begin{aligned}
\mathbf{P}_1 &= \mathbf{P}_{d,G}(t_G \mid t_G-1)/w + \mathbf{P}_{i,G}(t_G \mid t_G-1) \\
\mathbf{P}_2 &= \mathbf{P}_{d,j}(t_G \mid t_j)/(1-w) + \mathbf{P}_{i,j}(t_G \mid t_j) \\
\mathbf{K} &= \mathbf{P}_1(\mathbf{P}_1+\mathbf{P}_2)^{-1} \\
\mathbf{X}_G(t_G) &= \mathbf{X}_G(t_G \mid t_G-1) + \\
&\quad \mathbf{K}(\mathbf{X}_j(t_G \mid t_j) - \mathbf{X}_G(t_G \mid t_G-1)) \\
\mathbf{P}_G(t_G) &= (\mathbf{I}-\mathbf{K})\mathbf{P}_1 \\
\mathbf{P}_{i,G}(t_G) &= (\mathbf{I}-\mathbf{K})\mathbf{P}_{i,G}(t_G \mid t_G-1)(\mathbf{I}-\mathbf{K})^T + \\
&\quad \mathbf{K}\mathbf{P}_{i,j}(t_G \mid t_j)\mathbf{K}^T \\
\mathbf{P}_{d,G}(t_G) &= \mathbf{P}_G(t_G) - \mathbf{P}_{i,G}(t_G)
\end{aligned} \tag{8}$$

The w is determined by minimizing the determinant of the new covariance.

Case II: The new sensor track is fused into the global track not for the first time: there is track temporal correlation concerning this sensor; the IMF in (4) is performed to decorrelate the information between the sensor track updates at t_j and at t_j-1, as follows:

$$\begin{aligned}
&\mathbf{P}_G(t_G)^{-1} = \mathbf{P}_G(t_G \mid t_G-1)^{-1} + \\
&\quad (\mathbf{P}_j(t_G \mid t_j)^{-1} - \mathbf{P}_j(t_G \mid t_j-1)^{-1}) \\
&\mathbf{X}_G(t_G) = \\
&\mathbf{P}_G(t_G)\{\mathbf{P}_G(t_G \mid t_G-1)^{-1}\mathbf{X}_G(t_G \mid t_G-1) + \\
&\mathbf{P}_j(t_G \mid t_j)^{-1}\mathbf{X}_j(t_G \mid t_j) - \mathbf{P}_j(t_G \mid t_j-1)^{-1}\mathbf{X}_j(t_G \mid t_j-1)\} \\
&\mathbf{P}_{i,G}(t_G) = \\
&\mathbf{P}_G(t_G)\{\mathbf{P}_G(t_G \mid t_G-1)^{-1}\mathbf{P}_{i,G}(t_G \mid t_G-1)\mathbf{P}_G(t_G \mid t_G-1)^{-1} \\
&+ \mathbf{P}_j(t_G \mid t_j)^{-1}\mathbf{P}_{i,j}(t_G \mid t_j)\mathbf{P}_j(t_G \mid t_j)^{-1} - \\
&\mathbf{P}_j(t_G \mid t_j-1)^{-1}\mathbf{P}_{i,j}(t_G \mid t_j-1)\mathbf{P}_j(t_G \mid t_j-1)^{-1}\}\mathbf{P}_G(t_G) \\
&\mathbf{P}_{d,G}(t_G) = \mathbf{P}_G(t_G) - \mathbf{P}_{i,G}(t_G)
\end{aligned} \tag{9}$$

C. Discussion

In above introduced track-to-track fusion architecture, a sensor system can be completely modularized, because of two features of the architecture: first, a sensor track is *generated independently* by the corresponding sensor system; second, a sensor track is also *used independently* by the track-to-track fusion component, i.e. no coupling information between the sensor track and other sensor tracks is needed for the sensor track to be fused into the global track the second feature is as important as the first feature when we examine the modularity of a sensor system. Thanks to this complete sensor modularity, the proposed architecture enables flexible self-adjustment (adding or removing a sensor, modifying the inner function of a sensor system, etc).

Besides, also thanks to this complete modularity of sensors, there is no need for sensor synchronization, and the out-of-sequence measurement problem can be naturally handled.

IV. Simulation

A. Comparative study

In this section, we evaluate the proposed track-to-track fusion architecture using the SCIF-IMF. Here, we do not intend focusing on the absolute performance of the presented architecture, which in reality depends on *ad hoc* vehicle sensor configurations. Instead, we present a simulation based comparative study. Despite the gap always existing between the simulation performance and the performance in reality, yet a simulation based comparative study can well demonstrate the potential of a method.

The proposed track-to-track fusion method and several other methods were executed using the same synthetic data; their respective performances were compared. These methods under tests are as follows:

Centralized Kalman Filter Architecture (CKF):

The fusion is carried out at sensor-level; one fusion center directly takes in raw measurements (object observations) of all the sensor systems, and fuses the measurements using the Kalman filter.

Track-to-Track Fusion Architecture using the Kalman Filter (TTF_KF):

Each sensor system fuses its own measurements using the Kalman filter and generates a sensor track. The track-to-track fusion component fuses the sensor tracks also using the Kalman filter (global).

Track-to-Track Fusion Architecture using the Information Matrix Filter (TTF_IMF)

Each sensor system fuses its own measurements using the Kalman filter and generates a sensor track. The track-to-track fusion component fuses the sensor tracks using the IMF (without cross-covariance compensation).

Track-to-Track Fusion Architecture using the SCIF-IMF (TTF_SCIF_IMF):

Details are described in previous sections.

B. Simulation configuration and scenario

An overtaking scenario similar to that in [17] was simulated. Suppose there are two vehicles: the observing vehicle (OV) and the target vehicle (TV). The TV was overtaking the OV during totally 15 s: the initial (relative) position and (relative) velocity of the TV was x=-55m, y=0m, v_x=5m/s, and v_y=0m/s, with no initial accelerations; the TV accelerated longitudinally during t=[2, 5]s, with maximum a_x=1.5m/s^2; it changed to the overtaking lane during t=[2, 6]s and changed back to the cruising lane during t=[10, 14]s; it decelerated during t=[11, 14]s, with minimum a_x=-1.5m/s^2.

The OV was equipped with five sensors, namely two rearward sensors, a sideward sensor (facing the overtaking lane side), and two forward sensors. Here, we neglected the *ad hoc* types or features of the sensors; instead, we made an abstraction on their functions: a sensor was supposed to be able to periodically provide TV position observation with some uncertainty, and observe the TV during only a specific time duration (while the TV was traversing the sensor view field). The abstracted sensor configurations are listed in Table I.

TABLE I. SENSOR CONFIGURATIONS

Sensor	Rear1	Rear2	Side	Front1	Front2
Meas. period (s)	0.08	0.06	0.07	0.06	0.08
Meas. σ_x (m)	1.0	1.5	1.0	1.5	1.0
Meas. σ_y (m)	1.5	1.0	1.0	1.0	1.5
Meas. duration (s)	[0, 6]	[2, 7]	[6, 9]	[8, 13]	[9, 15]

Here for simulation, we did not simulate communication delay which can cause the out-of-sequence measurement (OOSM) problem. This was to idealize the CKF architecture, the performance of which then was optimal and could be regarded as the ideal for evaluating the other fusion architectures□ yet it is worthy noting that in reality, no matter what methods are used, the communication delay always exists and should be handled in real implementation (for example, by timestamping the data or estimates, and temporally evolving and aligning them).

C. Simulation Results

Synthetic data were generated according to the simulation configurations described in the previous sub-section. The CKF, the TTF_KF, the TTF_IMF, and the TTF_SCIF_IMF were executed using the same synthetic data. Totally 100 monte carlo simulation trials were carried out.

The performance of each method was evaluated from two aspects: the fused state accuracy and the fused covariance consistency. The fused state accuracy was directly evaluated by the RMS errors of all the trials, as shown in Fig.3. Since the performance of the CKF could be regarded as the ideal, we adopted a simple practice to evaluate the fused covariance consistency of the TTF_KF, the TTF_IMF, and the TTF_SCIF_IMF: let their fused covariance concerning the position component and the velocity component be normalized by that of the CKF, as shown in Fig.4. The closer the normalized covariance is to 1, the more consistent the covariance is.

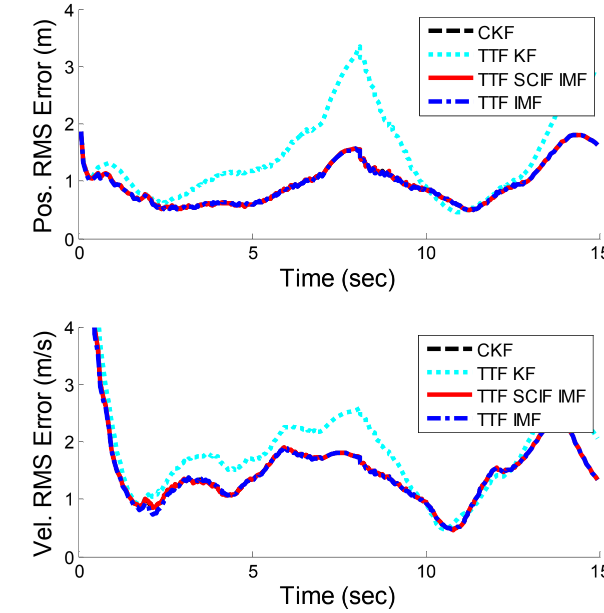

Fig.3. (top) position RMS errors; (bottom) velocity RMS errors

In Fig.3, the horizontal axis represents the time index of the overtaking duration; the vertical axis represents the position RMS errors (Fig.3-top) and the velocity RMS errors (Fig.3-bottom) associated with the methods under tests. Concerning the fused state accuracy, the TTF_SCIF_IMF and the TTF_IMF almost had the same performance to that of the CKF, whereas they had noticeably better performance than the TTF_KF did.

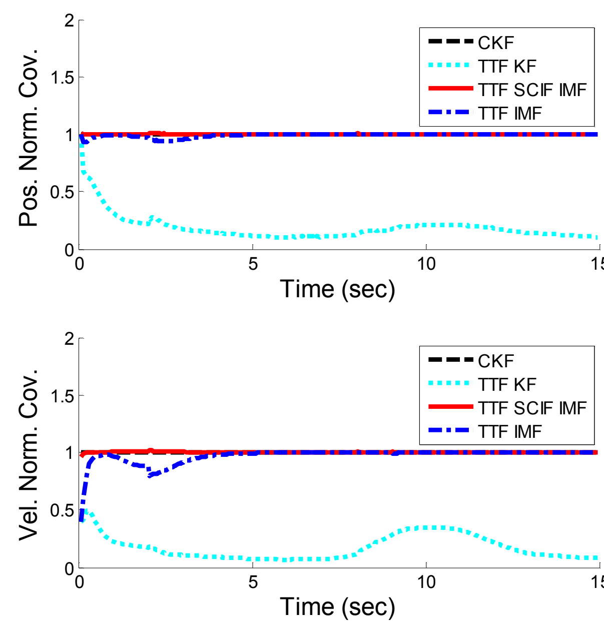

Fig.4. Normalized position covariance and velocity covariance

In Fig.4, the vertical axis represents the normalized position covariance (Fig.4-top) and the normalized velocity covariance (Fig.4-bottom) associated with TTF_KF, the TTF_IMF, and the TTF_SCIF_IMF. As show in Fig.4, the fused covariance of the TTF_KF quickly became highly inconsistent, and would be useless in practical application. The fused covariance of the TTF_IMF was noticeably inconsistent during the first several seconds. In contrast, the fused covariance of the TTF_SCIF_IMF was much more consistent through the entire overtaking process.

D. Discussion

As shown in Fig.3 and Fig.4, the track-to-track fusion architecture using the SCIF-IMF (the TTF_SCIF_IMF) had almost the same performance as the centralized Kalman filter architecture (the CKF) did. As reported in [17], the fusion architecture using the cross-covariance compensation and the IMF, denoted as TTF_CCC_IMF, demonstrates similar performance to that of the CKF. While the TTF_SCIF_IMF and the TTF_CCC_IMF are both track-to-track fusion architectures that have the potential to achieve centralized architecture comparable performance, the TTF_SCIF_IMF enjoys a further advantage, i.e. complete sensor modularity (as previously discussed in Section III-C).

V. CONCLUSION

In this paper, we have proposed a new track-to-track fusion architecture using the SCIF-IMF (split covariance intersection filter-information matrix filter) for vehicle surrounding environment perception. We have introduced how to use the SCIF to maintain estimates in split form for a generic sensor system. Concerning track-level fusion, we have introduced how to use the SCIF to handle possible track spatial correlation, and how to use the IMF (adapted for estimates in split form) to handle the track temporal correlation. The proposed architecture enjoys complete sensor modularity, and thus enables flexible self-adjustment.

We have presented a simulation based comparative study to evaluate the proposed architecture. The simulation tests have shown that the track-to-track fusion architecture using the SCIF-IMF can achieve centralized architecture comparable performance.

As simulation tests have shown promising potential of the track-to-track fusion architecture using the SCIF-IMF, transferring this potential into practical application is expected and will be the focus of future works.

REFERENCES

[1] C.C. Wang, “Simultaneous Localization, Mapping and Moving Object Tracking”, *PhD Thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh PA*, 2004

[2] T.D. Vu, “Vehicle perception: Localization, mapping with detection, classification and tracking of moving objects”, *PhD Thesis, Institut National Polytechnique de Grenoble*, 2009

[3] M. Darms, P. Rybski, C. Baker, C. Urmson, “Obstacle detection and tracking for the urban challenge”, *IEEE Trans. Intell. Transp. Syst.*, vol.10, no.3, pp.475-485, 2009

[4] M. Montemerlo, J. Becker, S. Bhat, H. Dolgov, S. Ettinger, D. Haehnel, T. Hilden, G. Hoffmann, B. Huhnke, D. Johnston, S. Langer, A. Levinson, J. Marcil, D. Orenstein, J. Paefgen, I. Penny, A. Petrovskaya, “Junior: The Stanford entry in the urban challenge”, *Journal of Field Robotics*, vol.25, no.9, pp.569-597, 2008

[5] K. Schueler, T. Weiherer, E. Bouzouraa, U. Hofmann, “360 degree multi sensor fusion for static and dynamic obstacles”, *IEEE Intell. Veh. Symp.*, 2012, pp.692-697

[6] R. Labayrade, C. Royere, D. Gruyer, D. Aubert, “Cooperative fusion for multi-obstacle detection with use of stereovision and laser scanner”, *Autonomous Robot*, vol.19, no.2, pp.117-140, 2005

[7] G. Gate, “Reliable perception of highly changing environments: implementations for car-to-pedestrian collision avoidance systems”, *PhD Thesis, Ecole Nationale Superieure des Mines de Paris*, 2009

[8] S. Pietzsch, T.D. Vu, J. Burlet, O. Aycard, T. Hackbarth, N. Appenrodt, J. Dickmann, B. Radig, “Results of a precrash application based on laser scanner and short-range radars”, *IEEE Trans. Intell. Transp. Syst.*, vol.10, no.4, pp.584-593, 2009

[9] M. Munz, M. Mahlisch, K. Dietmayer, “Generic centralized multi sensor data fusion based on probabilistic sensor and environment models for driver assistance systems”, *IEEE Intell. Transp. Syst. Mag.*, vol.2, no.1, pp.6-17, 2010

[10] Q. Baig, O. Aycard, T.D. Vu, T. Fraichard, “Fusion between laser and stereo vision data for moving objects tracking in intersection like scenario”, *IEEE Intell. Veh. Symp.*, 2011, pp.362-367

[11] D. Gohring, M. Wang, M. Schnurmacher, T. Ganjineh, “Radar/lidar sensor fusion for car-following on highways”, *IEEE Int. Conf. Autom. Robot. Appl.*, 2011, pp.407-412

[12] M. Aeberhard, N. Kaempchen, “High-level sensor data fusion architecture for vehicle surround environment perception”, *Proc. 8^{th} Int. Workshop Intell. Transp.*, Germany, 2011

[13] R. Mobus, U. Kolbe, “Multi-target multi-object tracking, sensor fusion of radar and infrared”, *IEEE Intell. Veh. Symp.*, 2004, pp.732-737

[14] S. Matzka, R. Altendorfer, “A comparison of track-to-track fusion algorithms for automotive sensor fusion”, *IEEE Int. Conf. Multi-Sens. Fusion Integr. Intell. Syst.*, 2008, pp.189-194

[15] Y. Bar-Shalom, X. Li, “Multitarget-multisensor tracking: principles and techniques”, *3^{rd} ed. Storrs, CT: YBS*, 1995

[16] G. Thomaidis, K. Vassilis, P. Lytrivis, M. Tsogas, G. Karaseitanidis, A. Amditis, “Target tracking and fusion in vehicular networks”, *IEEE Intell. Veh. Symp.*, 2011, pp.1080-1085

[17] M. Aeberhard, S. Schlichtharle, N. Kaempchen, T. Bertram, “Track-to-track fusion with asynchronous sensors using information matrix fusion for surround environment perception”, *IEEE Trans. Intell. Transp. Syst.*, vol.13, no.4, pp.1717-1726, 2012

[18] Y. Bar-Shalom, “On the track-to-track correlation problem”, *IEEE Trans. Autom. Control*, vol.AC-26, no.2, pp.571-572, 1981

[19] Y. Bar-Shalom, L. Campo, “The effect of the common process noise on the two-sensor fused-track covariance”, *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no.6, pp.803-805, 1986

[20] R.E. Kalman, “A new approach to linear filtering and prediction problem”, *ASME Trans, Series D: Journal of Basic Engineering*, 82, pp.35-45, 1960

[21] S.J. Julier, J.K. Uhlmann, “A non-divergent estimation algorithm in the presence of unknown correlations”, *Proc. American Control Conf.*, 1997, pp.2369-2373

[22] S.J. Julier, J.K. Uhlmann, “General decentralized data fusion with covariance intersection (CI)”, *Chapter 12, D. Hall, J. Llinas (Eds.), Handbook of Data Fusion, CRC Press, Boca Raton FL USA*, 2001

[23] K. Chang, R. Saha, Y. Bar-Shalom, “On optimal track-to-track fusion”, *IEEE Trans. Aerosp. Electron. Syst.*, vol.33, no.4, pp.1271-1276, 1997