Distributed Object Tracking based on Cubature Kalman Filter

Venkata Pathuri Bhuvana, Melanie Schranz, Mario Huemer and Bernhard Rinner
Klagenfurt University, Institute of Networked and Embedded Systems.
{venkata.pathuri, melanie.schranz, mario.huemer, bernhard.rinner }@aau.at

***Abstract*—In this work, we propose the cubature Kalman filter (CKF) based distributed object tracking algorithm in a visual sensor network (VSN). A VSN consists of several distributed smart cameras having the ability to process and analyze the retrieved data locally. The first objective is to optimize the tracking process within the VSN through the CKF. Under the considered conditions, the CKF-based method shows two times better tracking accuracy than the extended Kalman filter (EKF) based method in terms of the average root mean square error (RMSE). A particle filter (PF) shows even better performance than the CKF, however, it is computationally very complex. The second proposal is to optimize the object tracking by aggregating the tracking results from multiple cameras. Assuming the VSN is a multi-camera network with overlapping field of views (FOVs), cameras having the same object in their FOV exchange their local estimates of the object's position and velocity. During the estimation process, each of the participating cameras aggregates the received states via a consensus algorithm. Thus, the resulting joint state has a much higher probability to be closer to the object's real state, than a single camera's observation would be.**

I. Introduction and Motivation

Distributed estimation of an object's state is an extensively studied topic in the field of object tracking in camera networks. The main advantage is the independence of a central unit responsible for processing raw data streams from the individual cameras. In a completely distributed approach, the tracking is done locally on each camera having a specific object in its field of view (FOV). A consensus algorithm is used to aggregate the locally estimated states to a joint state as proposed by [1]. However, the tracking algorithm in [1] uses the Kalman filter (KF) assuming a linear state model for the tracked objects.

Contrary, the motion of real objects such as humans can only be accurately modeled by a non-linear state-space model. In such cases, the extended Kalman filter (EKF) based object tracking can be used, as in [2]. However, it has limited accuracy due to the inherent linearization errors. More efficient non-linear filters, such as the particle filter (PF) [3] can be used for distributed object tracking but it is computationally complex in sensor networks, generally. Recently, the cubature Kalman filter (CKF) [4] has been developed as an computationally efficient solution for non-linear state estimation.

The objective of this paper is to propose a distributed CKF-based object tracking approach in a VSN with overlapping FOVs. The CKF-based tracking algorithm runs locally on each camera to track multiple objects in the observed scene. Only cameras having the same object in their FOV exchange the object's state determined via the local CKF-based tracking algorithm among themselves. Each of the participating cameras performs the state estimation by aggregating the received states from the neighboring cameras tracking the same object.

II. Extended Summary of the Proposed Method

In this work, we consider a VSN consisting of a fixed set of calibrated smart cameras c_i, where $i = 1, 2, \cdots, C$, with overlapping FOVs as illustrated in Fig. 1. The smart cameras have the ability to analyze and process the retrieved data locally [5]. The task of the VSN is to observe the given

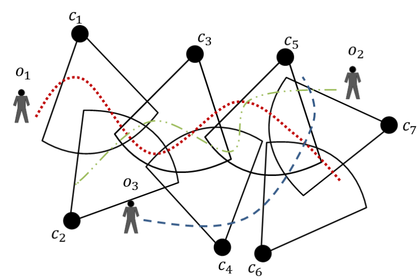

Fig. 1: Visual sensor network consisting of spatially distributed smart cameras.

environment and to identify and track a specific object o_k, where $k = 1, 2 \cdots, K$. This is achieved by a distributed tracking algorithm performed by each of the cameras c_i in the network. As these cameras are calibrated, there exists a homography to calculate the object's position on the ground plane. The re-identification of an object at any camera is a typical feature of the distributed tracking. Re-identification can be achieved with the help of the position on the ground plane in case of overlapping FOVs, or appearance features calculated already in the tracking process. In this work, each object can be (re-)identified by the cameras in the VSN.

The state of an object o_k comprises of its position and the velocity in the world coordinates. Thus, the state at time t is $\mathbf{x}_t^{ik} = [x\ y\ \dot{x}\ \dot{y}]$ where k and i represents the identity of the object and the camera, respectively. The non-linear state transition is given by,

$$\mathbf{x}_{t+1}^{ik} = \mathbf{f}_t^{ik}\left(\mathbf{x}_t^{ik}\right) + \mathbf{w}^{ik}, \tag{1}$$

where $\mathbf{w}^{ik}$ is an independent and identically distributed (IID) Gaussian process noise vector with covariance $\mathbf{Q}^{ik}$. The state of the object is generally estimated from a set of the measurements taken at each time step t. The system's measurement equation is given by

$$\mathbf{y}_t^{ik} = \mathbf{h}_t^{ik}\left(\mathbf{x}_t^{ik}\right) + \mathbf{v}^{ik}, \tag{2}$$

where $\mathbf{v}^{ik}$ is an IID measurement noise vector with covariance $\mathbf{R}^{ik}$. The measurement function $\mathbf{h}_t^{ik}$ converts the object's coordinates from the image to the ground plane.

A. CKF-based Tracking of an Object in the VSN

Under the assumption of the non-linear state and the measurement equations with the additive Gaussian noise, the state estimation boils down to intractable multi-dimensional Gaussian integrals of the form

$$I(f) = \int_{R^n} f(\mathbf{x}) \exp\left(-\mathbf{x}^T\mathbf{x}\right) d\mathbf{x}. \tag{3}$$

The third degree cubature rule can be used to numerically solve these types of integrals given by

$$I(f) \approx \frac{\sqrt{\pi^n}}{2n} \sum_{i=1}^{2n} f\left(\sqrt{\frac{n}{2}}\xi_i\right), \tag{4}$$

where n is the dimension of the vector $\mathbf{x}$, ξ_i is the i-th cubature point located at the intersection of the n-dimensional unit sphere and its axes. Hence, the CKF based on the third degree cubature rule provides an approximation to the states of the object at time t.

B. State Aggregation in a Multi-Camera Network

In our approach, each camera observes the object's position and estimates its corresponding state using CKF, locally. The local estimates are exchanged among themselves and aggregated to a joint state. Thus, consensus describes an agreement over the state $\mathbf{x}_t^{ik}$ from different cameras c_i on the same object o_k. The global state of an object o_k at time t at each camera c_i is then updated with

$$\mathbf{x}_t^{k} = \mathbf{x}_t^{ik} + \sum_{j=1}^{N} w_{ij}(\mathbf{x}_t^{jk} - \mathbf{x}_t^{ik}), \tag{5}$$

where $w_{ij} \in \{0,1\}$ stands for the weights of the communication link. Thus, the weights correspond to the adjacency matrix. Comparing the individual states to the joint state, the latter is much more optimized in terms of closeness to the real object's state.

III. Preliminary Results and Conclusions

In the preliminary simulation run, a VSN with two cameras having an overlapping FOV tracking a single object, is considered. The object's movement is modeled in a non-linear way, moves and stops according to a certain probability. The corresponding results show that the CKF-based distributed object estimation has significant improvement over the EKF-based method. Fig. 2 illustrates a time snapshot of the estimated x and y coordinates for a single object of the CKF-, EKF-, and PF-based distributed tracking methods compared to the ground truth. In addition, this figure shows the corresponding RMSEs. Although, the PF-based method achieves a slight improvement over the CKF-based method in terms of the RMSE, it is computationally more complex. Table I compares the average RMSE of the three methods and their complexity. The average RMSE is calculated over a fifty simulation runs as each simulation run tracks the object for 200 time steps. The complexity is illustrated in terms of the execution time for a single simulation run in Matlab (number of particles for the PF is considered as 1000). Since the accuracy and the complexity are two important measures for the object tracking in a VSN, the CKF-based distributed tracking is the best trade-off.

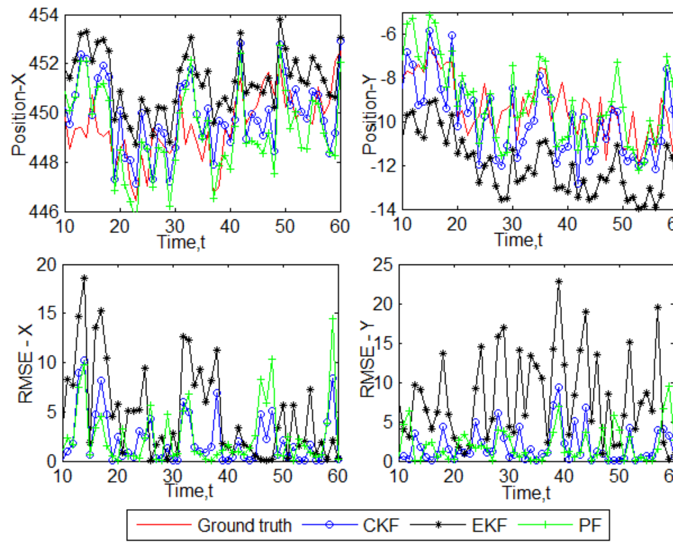

Fig. 2: Comparison of the distributed object tracking methods.

TABLE I: Efficiency and complexity of the object tracking methods.

Method	%RMSE	Execution time (sec)
EKF	2.76	0.0114
CKF	1.31	0.0899
PF (N=1000)	1.14	4.3534

In the full paper, the algorithm of the proposed CKF-based object tracking and the consensus algorithm in a distributed VSN will be provided in detail. In addition, we will compare the efficiency of the CKF-based method to that of the EKF and PF-based tracking methods using real time measurement data. The complexity and the communication overhead of the consensus algorithm will also be discussed in detail.

References

[1] B Song et al. Distributed camera networks. *IEEE Signal Proc. Mag.*, 28(3), 2011.

[2] C. Ding et al. Collaborative sensing in a distributed ptz camera network. *IEEE Trans. Image Process.*, 21(7), 2012.

[3] O. Hlinka et al. Distributed gaussian particle filtering using likelihood consensus. In *Proc. IEEE ICASSP*, 2011.

[4] I. Arasaratnam and Simon Haykin. Cubature kalman filters. *IEEE Trans. Automat. Contr.*, 54(6), 2009.

[5] B. Rinner et al. The evolution from single to pervasive smart cameras. In *Proc. ACM/IEEE ICDSC*, 2008.