Square-Root Sigma-Point Information Consensus Filters for Distributed Nonlinear Estimation

Guoliang Liu * and Guohui Tian

School of Control Science and Engineering, Shandong University, Jinan 250061, China; g.h.tian@sdu.edu.cn
* Correspondence: liuguoliang@sdu.edu.cn; Tel.: +86-531-88392115

Abstract: This paper focuses on the convergence rate and numerical characteristics of the nonlinear information consensus filter for object tracking using a distributed sensor network. To avoid the Jacobian calculation, improve the numerical characteristic and achieve more accurate estimation results for nonlinear distributed estimation, we introduce square-root extensions of derivative-free information weighted consensus filters (IWCFs), which employ square-root versions of unscented transform, Stirling's interpolation and cubature rules to linearize nonlinear models, respectively. In addition, to improve the convergence rate, we introduce the square-root dynamic hybrid consensus filters (DHCFs), which use an estimated factor to weight the information contributions and shows a faster convergence rate when the number of consensus iterations is limited. Finally, compared to the state of the art, the simulation shows that the proposed methods can improve the estimation results in the scenario of distributed camera networks.

Keywords: target tracking; sensor network; information filter; distributed estimation

1. Introduction

Consensus based distributed estimation has attracted a lot of attention in the field, due to its outstanding performance in many applications, e.g., distributed camera networks [1], mobile sensor networks [2,3] and multi-agent systems [4,5]. Compared to the centralized estimation, the distributed estimation has many advantages, such as good scalability, low computation cost and robustness to the sensor failure [6]. In many applications, sensor nodes in the distributed network may have multiple measurements of the target state. It is important to fuse all the measurement information from sensor nodes to achieve a robust estimation result. In distributed sensor networks, there are no central fusion nodes. Instead, a fusion result or a common estimation goal can be achieved by using the novel consensus method. Consensus means reaching an agreement regarding a certain quantity of interest which depends on the state of all sensor nodes [7]. In the consensus algorithm, the node communicates to its neighbor nodes, and converges to a global mean result after a number of iterations, e.g., arithmetic mean or geometric mean [8]. Due to the limited bandwidth of the real network, only a limited number of iterations can be applied, so that the true convergence may not be always reached [8]. Therefore, the convergence rate of the consensus algorithm is very important, which is the main research aspect of this note.

In order to estimate the state of the node in the consensus architecture, the Kalman filter or information filter and their extensions are often employed [2,8–13]. In [9], a Kalman filter is used to work with the consensus algorithm, which is called Kalman consensus filter (KCF). The original KCF algorithm works well when all nodes can observe the target, but has decreased performance when the sensor node becomes naive, e.g., the node has limited observability [8]. As an alternative, the information filter is introduced to replace the Kalman filter, which is the information consensus filter (ICF) [2,10]. The information filter uses information matrix and information vector instead of

the moments (mean and covariance) used in the Kalman filter to represent the Gaussian distribution. In this way, the information filter has advantages to handle sensor fusion tasks and unknown prior covariance conditions [14]. However, the ICF as well as the KCF did not address the naivety problem and the redundancy problem. The redundancy problem is caused by the iterative information exchange in consensus methods, which correlates the node estimation and delays the convergence to the optimal result [12].

To overcome the naivety and redundancy problem, an information weighted consensus filter (IWCF) was proposed in [1,15]. The IWCF solves the naivety and redundancy problem by giving less weight to the prior information when the new information contribution is fused, since the redundancy information is present only in the prior information. In addition, the IWCF can converge to the centralized solution when the number of the consensus iterations to the infinity by setting the weight as $1/N$, where N is the number of sensor nodes. However, the sensor nodes can be overweighted in the IWCF if only a limited number of consensus iterations is performed, and the consistence of the local filter can be destroyed [11]. As an alternative weighting scheme, the inverse of an estimated factor S/N can be used to weight the new information contributions, where S means the number of valid sensor nodes that can observe the target currently. This choice has the desirable property of preserving the consistence of local filters and the novel information is never overestimated. Since the estimated value S/N is changing all the time due to the limited filed of view (FOV) of the sensor, we call this new filter a dynamic hybrid consensus filter (DHCF).

To handle the nonlinear system, the extended information weighted consensus filter (EIWCF) was introduced in [12], where the extended information filter is used to handle the nonlinearity. However, the IWCF and the EIWCF use a fixed consensus rate in the consensus algorithm, which is not optimal for achieving a fast convergence rate. As an alternative, the Metropolis weight and Maximum-degree weight have been proposed to work with the consensus approaches [2,11,16,17]. The Metropolis weight can achieve a faster convergence rate than the Maximum-degree weight proved in [16], which means it can achieve more accurate results when a limited number of consensus iterations is performed. Therefore, it is interesting to see how well the IWCF and the EIWCF work with the Metropolis weight. In addition, the EIWCF employs the first order of Taylor series to linearize the nonlinear models, which has low accuracy compared to other modern optimized linearization techniques, i.e., Stirling's interpolation [18,19], unscented transform [20,21], spherical cubature rules [22,23] and their square-root extensions [14,24–26]. These optimized linearization methods can be called sigma-point filters since they use a number of sampled sigma-points to approximate the distribution of the state variables, and then propagate these sigma-points through the nonlinear functions to get predicted state and observations. The posterior statistics are then calculated by a weighted summation of all sigma-points. The main difference between these sigma-point filters is how to set the weights in the summation. Furthermore, these weights can affect the positive definite property of the covariance, which can make the filter numerically unstable. For instance, the parameter used in classical unscented transform can lead to the negative weights and further destroy the positive definite property of the covariance as shown in [14,27,28]. However, the cubature rules as a special kind of unscented transform is more stable by setting the parameter $\alpha = 1$, $\beta = 0$, $\kappa = 0$ [27]. On the other hand, Stirling's interpolation also shows improved numerical performance compared to the classical unscented transform by making the weights positive [14]. Recently, a square-root cubature information weighted consensus filter (SRCIWCF) based on cubature rules has been developed for distributed object tracking, which shows that the SRCIWCF is more numerically accurate and stable than the EIWCF [13,29]. Motivated by the development of the nonlinear IWCF and its square-root extension, we developed a class of square-root sigma-point information consensus filters here, which employ Stirling's interpolation, unscented transform and cubature rules for linearization in the framework of IWCF and DHCF, respectively. The comprehensive performance comparison between the state-of-the-art and proposed square-root sigma-point information consensus filters is also demonstrated using a simulated camera network.

The structure of this note is the following: first, the nonlinear information weighted consensus filters are introduced in Section 2, which are based on the first order of Taylor expansion, Stirling's interpolation, unscented transform and cubature rules, respectively. Second, the nonlinear dynamic hybrid consensus filters are proposed in Section 3, which use an alternative consensus weighting scheme and can keep the consistency of local filters. Finally, a sparse of camera network is simulated to illustrate the performance of proposed square-root sigma-point information consensus methods in Section 4.

2. Nonlinear Information Weighted Consensus Filters

In the following sections, we consider that the sensor network has N nodes, which construct an undirected graph $G = (C, E)$ where $C = \{1, 2, 3, \cdots, N\}$ denotes vertex set and $E \subset \{\{i, j\} | i, j \subset C\}$ means the edge set. The neighbor nodes of the i_{th} node can be defined as $\mathcal{N}_i = \{j \in C | \{i, j\} \in E\}$, which has N_i nodes.

Algorithm 1 Extended Information Weighted Consensus Filter

- Initialization: Consensus rate ϵ, number of consensus iteration L, process noise Q and measurement noise R.
- For $k = 1, \cdots, \infty$:

  1. Prediction for the next time step:
  $$\hat{x}_{i,k} = f(x_{i,k-1}), \tag{1}$$
  $$\hat{Y}_{i,k} = (J_k Y_{i,k-1}^{-1} J_k^T + W_k Q_k W_k^T)^{-1}, \tag{2}$$
  $$\hat{y}_{i,k} = \hat{Y}_{i,k} \hat{x}_{i,k}. \tag{3}$$

  2. Compute consensus proposals
  $$v_{i,k}^0 = \frac{1}{N} \hat{y}_{i,k} + \phi_{i,k}, \tag{4}$$
  $$V_{i,k}^0 = \frac{1}{N} \hat{Y}_{i,k} + \Phi_{i,k}. \tag{5}$$

  3. Perform consensus on $v_{i,k}^0$ and $V_{i,k}^0$
     for $l = 1$ to L **do**
        (a) Send $v_{i,k}^{l-1}$ and $V_{i,k}^{l-1}$ to all neighbors $j \in \mathcal{N}_i$
        (b) Receive $v_{j,k}^{l-1}$ and $V_{j,k}^{l-1}$ from all neighbors $j \in \mathcal{N}_i$
        (c) Update consensus terms:
  $$v_{i,k}^l = v_{i,k}^{l-1} + \epsilon \sum_{j \in \mathcal{N}_i} (v_{j,k}^{l-1} - v_{i,k}^{l-1}), \tag{6}$$
  $$V_{i,k}^l = V_{i,k}^{l-1} + \epsilon \sum_{j \in \mathcal{N}_i} (V_{j,k}^{l-1} - V_{i,k}^{l-1}). \tag{7}$$

     end for
  4. Compute the posterior at k time step
  $$y_{i,k} = N v_{i,k}^L, \tag{8}$$
  $$Y_{i,k} = N V_{i,k}^L, \tag{9}$$
  $$x_{i,k} = Y_{i,k}^{-1} y_{i,k}. \tag{10}$$

2.1. Extended Information Weighted Consensus Filter

For the nonlinear system, the motion model of the target and the measurement model of the sensor can be described as

$$x_k = f(x_{k-1}, q_k), \tag{11}$$

$$z_{i,k} = h_i(x_k, v_{i,k}), \tag{12}$$

where f and h_i are the state transition function and measurement function of the i_{th} sensor node, respectively, and q_k and $v_{i,k}$ are zero mean white Gaussian noises with covariance matrix Q_k and $R_{i,k}$, respectively. In the case of the nonlinear function f and/or h_i, the extended information filter (EIF) can be used for the linearization. According to [12], the EIF based consensus algorithm EIWCF can be summarized as Algorithm 1. In the prediction step, the J_k and W_k in (2) are Jacobians of function f with respect to x_k and w_k, respectively. To calculate the consensus quantity $v_{i,k}^0$ and $V_{i,k}^0$ in (4) and (5), the information contributions are derived as

$$\phi_{i,k} = J_{h,k}^T R^{-1}(z_{i,k} - h_i(\hat{x}_{i,k}) + J_{h,k}\hat{x}_{i,k}), \tag{13}$$

$$\Phi_{i,k} = J_{h,k}^T R^{-1} J_{h,k}, \tag{14}$$

where $J_{h,k}$ is Jacobian of function h_i with respect to x_k.

The iterations of consensus are performed in a loop through (6) and (7), where ϵ is the consensus weight which determines the convergence rate of the algorithm. Normally, the ϵ is between 0 and $1/\Delta_{max}$, where Δ_{max} is the maximum degree of the graph G. The original IWCF algorithm in [8] uses a deterministic value $\epsilon = 0.65/\Delta_{max}$, which is not optimal for convergence as we show in the following sections. In addition, the fixed ϵ used here requires the knowledge of the global maximum degree of the graph, which is not robust for handling the network topology changing problem. As suggested in [16], the Metropolis weights can offer a faster convergence rate without the knowledge the number of sensor nodes N, which is defined as

$$\epsilon_{i,j,k} = \begin{cases} \frac{1}{1+max\{d_{i,k}, d_{j,k}\}} & if\, j \in \mathcal{N}_i, \\ 1 - \sum_{j \in \mathcal{N}_i} \epsilon_{i,j,k} & if\, i = j, \\ 0 & otherwise, \end{cases} \tag{15}$$

where $d_{i,k}$ and $d_{j,k}$ are the degrees of the node i and node j, respectively. The Metropolis weight only needs to know the local degree of the neighbor nodes, whereas the fixed consensus weight used in the original EIWCF algorithm requires the global knowledge of the maximum degree of the graph. Therefore, the Metropolis weight is more robust to handle the network topology changing problem. By using Metropolis weight (15) in the consensus steps (6) and (7), the new version of EIWCF algorithm can be derived which is called EIWCFM in the following part of this note.

The advantage of the EIWCF and EIWCFM would guarantee convergence to the optimal centralized estimation when the number of consensus iterations $L \to \infty$. However, due to the limited communication resources, the number of consensus iterations is limited. Therefore, it is important for the EIWCF and EIWCFM to have a faster convergence rate when a finite number of consensus iterations is performed.

2.2. Square-Root Central Difference Information Weighted Consensus Filter

---

Algorithm 2 Square-Root Central Difference Information Weighted Consensus Filter (SRCDIWCF)

- Initialization: Number of consensus iteration L, process noise $Q = S_q S_q^T$ and measurement noise $R = S_r S_r^T$.
- For $k = 1, \cdots, \infty$:

  1. Prediction for the next time step:

$$\hat{x}_{i,k} = \sum_{\tau=0}^{2m} w_\tau^x \mathcal{X}_{i,\tau,k|k-1}^x, \tag{16}$$

$$\hat{S}_{i,k}^x = qr\{[A; B]\}, \tag{17}$$

$$\hat{S}_{i,k}^y = qr\{(\hat{S}_{i,k}^x)^{-1} I\}, \tag{18}$$

$$\hat{y}_{i,k} = (\hat{S}_{i,k}^x)^{-T}(\hat{S}_{i,k}^x)^{-1}\hat{x}_{i,k}. \tag{19}$$

  2. Compute consensus proposals

$$v_{i,k}^0 = \frac{1}{N}\hat{y}_{i,k} + \phi_{i,k}, \tag{20}$$

$$V_{i,k}^0 = qr\{[\frac{1}{\sqrt{N}}\hat{S}_{i,k}; \Phi_{i,k}]\}. \tag{21}$$

  3. Perform consensus on $v_{i,k}^0$ and $V_{i,k}^0$
     for $l = 1$ to L **do**
       (a) Send $v_{i,k}^{l-1}$ and $V_{i,k}^{l-1}$ to all neighbors $j \in \mathcal{N}_i$
       (b) Receive $v_{j,k}^{l-1}$ and $V_{j,k}^{l-1}$ from all neighbors $j \in \mathcal{N}_i$
       (c) Update consensus terms:

$$v_{i,k}^l = v_{i,k}^{l-1} + \epsilon \sum_{j\in\mathcal{N}_i}(v_{j,k}^{l-1} - v_{i,k}^{l-1}), \tag{22}$$

$$V_{i,k}^l = qr\{[\sqrt{1-\epsilon N_i}V_{i,k}^{l-1}; \sqrt{\epsilon}V_{j=1,k}^{l-1}; \cdots; \sqrt{\epsilon}V_{j=N_i,k}^{l-1}]\}. \tag{23}$$

     end for

  4. Compute the posterior at k time step

$$y_{i,k} = Nv_{i,k}^L, \tag{24}$$

$$S_{i,k}^y = \sqrt{N}V_{i,k}^L, \tag{25}$$

$$x_{i,k} = (S_{i,k}^y)^{-T}(S_{i,k}^y)^{-1}y_{i,k}. \tag{26}$$

The core of the proposed square-root central difference information weighted consensus filter (SRCDIWCF) is the Stirling's interpolation for linearization. It first generates a number of sample sigma-points according to the current augmented state and covariance. The sampled state $\mathcal{X}_{i,\tau,k-1}^x$ together with its sampled process noise $\mathcal{X}_{i,\tau,k-1}^q$ construct the sampled augmented state $\mathcal{X}_{i,\tau,k-1}^{aq} = [\mathcal{X}_{i,\tau,k-1}^x \ \mathcal{X}_{i,\tau,k-1}^q]^T$ generated by

$$\mathcal{X}_{i,\tau,k-1}^{aq} = \begin{cases} x_{i,k-1}^{aq}, & \tau = 0, \\ x_{i,k-1}^{aq} + (hS_{i,k-1}^{aq})_\tau, & \tau = 1, \cdots, m, \\ x_{i,k-1}^{aq} - (hS_{i,k-1}^{aq})_\tau, & \tau = m+1, \cdots, 2m, \end{cases} \tag{27}$$

where $x_{i,k-1}^{aq} = [x_{i,k-1} \ \bar{q}_{i,k-1}]^T$ and $S_{i,k-1}^{aq} = \sqrt{P_{i,k-1}^{aq}} = \sqrt{diag\{P_{i,k-1}, Q_{i,k-1}\}}$ are the augmented state and square root of augmented covariance, τ indicates the τ_{th} column of the matrix, and m is the dimension of the augmented state. The parameter $h \geq 1$ is the scalar central difference step size. If the random variables obey a Gaussian distribution, the optimal value of h is $\sqrt{3}$ [19]. We can see that the square root calculation of the covariance in (27) requires that the covariance matrix $P_{i,k-1}^{aq}$ must be symmetric and positive definite. However, due to the errors introduced by arithmetic operations performed on finite word-length digital computers, or ill-conditioned nonlinear filtering problems, the positive definite property of the covariance can be destroyed. In the literature, the square-root representation of the covariance is preferred to handle such an issue. Therefore, we here present the SRCDIWCF summarized in Algorithm 2, which can avoid square-root operation, improve numerical accuracy, have double order precision and preserve symmetry of the covariance.

After the generation of the sigma-points in (27), we can propagate them through the nonlinear state transition function of (11), and sum them up to derive the predicted state as in (16). Therefore, no Jacobian matrix calculation is required here. The predicted square-root of covariance $\hat{S}_{i,k}^x$ can be calculated using QR decomposition as in (17), where

$$A = \sqrt{w_1^{p_1}} \left(\mathcal{X}_{i,1:m,k|k-1}^x - \mathcal{X}_{i,m+1:2m,k|k-1}^x \right), \tag{28}$$

$$B = \sqrt{w_1^{p_2}} (\mathcal{X}_{i,1:m,k|k-1}^x + \mathcal{X}_{i,m+1:2m,k|k-1}^x - 2\mathcal{X}_{i,0,k|k-1}^x). \tag{29}$$

The corresponding weights for the predicted mean and square-root of covariance are defined as

$$\begin{aligned} w_0^x &= \frac{h^2-m}{h^2}, \\ w_\tau^x &= \frac{1}{2h^2}, \\ w_\tau^{p_1} &= \frac{1}{4h^2}, \\ w_\tau^{p_2} &= \frac{h^2-1}{4h^4}, \quad \tau = 1, \cdots, 2m, \end{aligned} \tag{30}$$

where we can see that weights $w_\tau^{p_1}$ and $w_\tau^{p_1}$ for the covariance updating are all non-negative values since $h \geq 1$, which is an important property of the proposed SRCDIWCF, since it can protect the positive property of the covariance matrix and further improve the numerical characteristics of the proposed method [14]. Because we are interested in the information form representation, the predicted information vector and information matrix can be computed from predicted mean and covariance as shown in (19) and (18), respectively.

The second step is to compute the consensus quantities $v_{i,k}^0$ and $v_{i,k}^0$ defined as (20) and (21), where $\phi_{i,k}$ and $\Phi_{i,k}$ are information contributions calculated as

$$\Phi_{i,k} = \hat{S}_{i,k}^y (\hat{S}_{i,k}^y)^T \hat{P}_{i,xz} S_{r,i,k}^{-T}, \tag{31}$$

$$\phi_{i,k} = \Phi_{i,k} (S_{r,i,k})^{-1} (z_{i,k} - \hat{z}_{i,k} + \hat{P}_{i,xz}^T \hat{y}_{i,k}), \tag{32}$$

where $z_{i,k}$ and $\hat{z}_{i,k}$ are real and predicted sensor measurements of the target object, respectively. The predicted $\hat{z}_{i,k}$ can be calculated using Stirling's interpolation according to the predicted state of the target as

$$\hat{z}_{i,k} = \sum_{\tau=0}^{2m} w_\tau^x \mathcal{Z}_{i,\tau,k|k-1}, \tag{33}$$

$$\mathcal{Z}_{i,\tau,k|k-1} = h(\mathcal{X}_{i,\tau,k|k-1}), \tag{34}$$

$$\mathcal{X}_{i,\tau,k|k-1} = \begin{cases} \hat{x}_{i,k}, & \tau = 0, \\ \hat{x}_{i,k} + (h\hat{S}_{i,k}^x)_\tau, & \tau = 1, \cdots, m, \\ \hat{x}_{i,k} - (h\hat{S}_{i,k}^x)_\tau, & \tau = m+1, \cdots, 2m, \end{cases} \tag{35}$$

where $\mathcal{X}_{i,\tau,k|k-1}$ is the generated sigma point according to the predicted state and covariance, and $\mathcal{Z}_{i,\tau,k|k-1}$ is the predicted sigma point of the measurement. The cross-covariance $\hat{P}_{i,xz}$ between the state and the measurement calculated as

$$\hat{P}_{i,xz} = \sqrt{w_1^{p_1}}\hat{S}_{i,k}^x(\mathcal{Z}_{i,1:m} - \mathcal{Z}_{i,m+1:2m})^T. \tag{36}$$

From (20) and (21), we can see that both the prior information vector $\hat{y}_{i,k}$ and prior information matrix $\hat{Y}_{i,k}$ are weighted by $1/N$. The reason to do this weighting is to remove the information redundancy during information sharing between sensor nodes.

The third step is to perform consensus iteratively. The sensor node exchanges the consensus quantities $v_{i,k}^l$ and $V_{i,k}^l$ with neighbor nodes for L steps, and then update its local estimation using (22) and (23), where $\epsilon \in (0, 1/\Delta_{max})$ is defined as the same constant parameter as the EIWCF presented in [12]. If the Metropolis weight is used for ϵ instead of the fixed one, the SRCDIWCF is called SRCDIWCFM for distinguishment. Therefore, the basic idea of the consensus is to get the weighted summation of information quantities, so the information from neighbor nodes can be fused. To output the final results, the fourth step is to derive the estimated information vector $y_{i,k}$, information matrix $S_{i,k}^y$ and state $x_{i,k}$ from the final information quantities using (24)–(26).

2.3. Square-Root Unscented Information Weighted Consensus Filter

Here, we present the unscented transform based square-root unscented information weighted consensus filter (SRUIWCF). The main difference between the SRUIWCF and SRCDIWCF is the weighting scheme for the sigma points to calculate the predicted mean, covariance and information contributions. For the SRUIWCF, the sigma-points are generated by

$$\mathcal{X}_{i,\tau,k-1}^{aq} = \begin{cases} x_{i,k-1}^{aq}, & \tau = 0, \\ x_{i,k-1}^{aq} + (\gamma S_{i,k-1}^{aq})_\tau, & \tau = 1, \cdots, m, \\ x_{i,k-1}^{aq} - (\gamma S_{i,k-1}^{aq})_\tau, & \tau = m+1, \cdots, 2m, \end{cases} \tag{37}$$

where $\gamma = \sqrt{(\lambda + m)}$ is the composite scaling parameter, m is the dimension of the state, and $\lambda = \alpha^2(m + \kappa) - m$. α and κ are scaling parameters that determine how far the sigma points spread from the mean value. The predicted mean and the square-root of covariance can be derived by

$$\hat{x}_{i,k} = \sum_{\tau=0}^{2m} w_\tau^x \mathcal{X}_{i,\tau,k|k-1}^x, \tag{38}$$

$$\hat{S}_{i,k}^x = cholupdate\{C, D, sign\{w_0^p\}\}, \tag{39}$$

$$C = qr\{\sqrt{w_1^p}(\mathcal{X}_{i,1:2m,k|k-1}^x - \hat{x}_{i,k})\}, \tag{40}$$

$$D = \sqrt{w_0^p}(\mathcal{X}_{i,0,k|k-1}^x - \hat{x}_{i,k}), \tag{41}$$

where the weights defined by

$$\begin{aligned} w_0^x &= \frac{\lambda}{m+\lambda}, \\ w_0^p &= \frac{\lambda}{m+\lambda} + (1 - \alpha^2 + \beta), \\ w_\tau^x &= w_\tau^p = \frac{1}{2(m+\lambda)}, & \tau = 1, \cdots, 2m. \end{aligned} \tag{42}$$

Since the weight w_0^p might be negative, we need an additional *cholupdate* to update the *Cholesky factor* $\hat{S}_{i,k}^x$ in (39), whereas the SRCDIWCF does not need this step since all weights used for the covariance update are positive. The negative update might destroy the positive definite property

of the Cholesky factor, such that the SRCDIWCF is preferable to the SRUIWCF concerning the numerical stability.

For calculating information contributions, the predicted measurement $\mathcal{Z}_{i,\tau,k|k-1}$ and the cross-covariance $\hat{P}_{i,xz}$ are given by

$$\hat{z}_{i,k} = \sum_{\tau=0}^{2m} w_\tau^x \mathcal{Z}_{i,\tau,k|k-1}, \tag{43}$$

$$\hat{P}_{i,xz} = \sum_{\tau=0}^{2m} w_\tau^p (\mathcal{X}_{i,\tau,k|k-1} - \hat{x}_{i,k})(\mathcal{Z}_{i,\tau,k|k-1} - \hat{z}_{i,k})^T. \tag{44}$$

The consensus quantities and iterations steps of the SRUIWCF are same as the SRCDIWCF.

2.4. Square-Root Cubature Information Weighted Consensus Filter

Basically, the cubature rule is a special case of the unscented transform defined by

$$\mathcal{X}_{i,\tau,k-1}^{aq} = \begin{cases} x_{i,k-1}^{aq} + (\sqrt{m} S_{i,k-1}^{aq})_\tau, & \tau = 1, \cdots, m, \\ x_{i,k-1}^{aq} - (\sqrt{m} S_{i,k-1}^{aq})_\tau, & \tau = m+1, \cdots, 2m, \end{cases} \tag{45}$$

and the weights used for calculating posterior mean and covariance given by

$$w_\tau^x = w_\tau^p = \frac{1}{2m}, \quad \tau = 1, \cdots, 2m. \tag{46}$$

By setting the parameter of the unscented transform as $\alpha = 1$, $\beta = 0$ and $\kappa = 0$, the unscented transform becomes the cubature rule [27,28]. Therefore, the square-root cubature information weighted consensus filter (SRCIWCF) can be derived from SRUIWCF by using this specific parameter. Since the weights used in SRCIWCF are positive, the SRCIWCF is more numerically stable than the SRUIWCF.

3. Nonlinear Dynamic Hybrid Consensus Filters

The IWCFs can converge to the centralized solution when the number of consensus L is infinity. However, for the real-time requirements of the real applications, L is usually small. In such a case, the weight $1/N$ used in IWCFs can overweight the prior information for some sensor nodes, which can affect the convergence rate of the algorithm. Here, we present an alternative way to weight the information, in order to ensure that the sensor node never get overweighted, and has a faster convergence rate if a small L is used.

Algorithm 3 Square-Root Central Difference Dynamic Hybrid Consensus Filter (SRCDDHCF)

- Initialization: Number of consensus iteration L, process noise $Q = S_q S_q^T$ and measurement noise $R = S_r S_r^T$.
- For $k = 1, \cdots, \infty$:

1. Prediction for the next time step:

$$\hat{x}_{i,k} = \sum_{\tau=0}^{2m} w_\tau^x \mathcal{X}_{i,\tau,k|k-1}^x, \tag{47}$$

$$\hat{S}_{i,k}^x = qr\{[A;\ B]\}, \tag{48}$$

$$\hat{S}_{i,k}^y = qr\{(\hat{S}_{i,k}^x)^{-1}\ I\}, \tag{49}$$

$$\hat{y}_{i,k} = (\hat{S}_{i,k}^x)^{-T}(\hat{S}_{i,k}^x)^{-1}\hat{x}_{i,k}. \tag{50}$$

2. Compute consensus proposals
   if $i \in \mathcal{S}$ **then**

$$\Phi_{i,k} = \hat{S}_{i,k}^y(\hat{S}_{i,k}^y)^T \hat{P}_{i,xz} S_{r,i,k}^{-T}, \tag{51}$$

$$\phi_{i,k} = \Phi_{i,k}(S_{r,i,k})^{-1}(z_{i,k} - \hat{z}_{i,k} + \hat{P}_{i,xz}^T \hat{y}_{i,k}), \tag{52}$$

$$b_{i,k} = 1, \tag{53}$$

   else

$$\Phi_{i,k} = 0,\ \phi_{i,k} = 0,\ b_{i,k} = 0. \tag{54}$$

   end if

3. Perform hybrid consensus iterations
   Initialization: $b_{i,k}^0 = b_{i,k}$, $(\hat{y}_{i,k}^0 = \hat{y}_{i,k}, \hat{S}_{i,k}^0 = \hat{S}_{i,k}^y)$ and $(\phi_{i,k}^0 = \phi_{i,k}, \Phi_{i,k}^0 = \Phi_{i,k})$
   for $l = 1$ to L **do**
   (a) Send $b_{i,k}^{l-1}$, $(\hat{y}_{i,k}^{l-1}, \hat{S}_{i,k}^{l-1})$ and $(\phi_{i,k}^{l-1}, \Phi_{i,k}^{l-1})$ to all neighbors $j \in \mathcal{N}_i$
   (b) Receive $b_{i,k}^{l-1}$, $(\hat{y}_{i,k}^{l-1}, \hat{S}_{i,k}^{l-1})$ and $(\phi_{i,k}^{l-1}, \Phi_{i,k}^{l-1})$ from all neighbors $j \in \mathcal{N}_i$
   (c) Update consensus terms:

$$y_{i,k}^l = \sum_{j \in \mathcal{N}_i} \epsilon_{i,j,k} y_{j,k}^{l-1}, \tag{55}$$

$$S_{i,k}^l = qr\{[\sqrt{\epsilon_{i,j,k}} S_{j=1,k}^{l-1};\ \cdots;\ \sqrt{\epsilon_{i,j,k}} S_{j=N_i,k}^{l-1}]\}, \tag{56}$$

$$\phi_{i,k}^l = \sum_{j \in \mathcal{N}_i} \epsilon_{i,j,k} \phi_{j,k}^{l-1}, \tag{57}$$

$$\Phi_{i,k}^l = qr\{[\sqrt{\epsilon_{i,j,k}} \Phi_{j=1,k}^{l-1};\ \cdots;\ \sqrt{\epsilon_{i,j,k}} \Phi_{j=N_i,k}^{l-1}]\}, \tag{58}$$

$$b_{i,k}^l = \sum_{j \in \mathcal{N}_i} \epsilon_{i,j,k} b_{j,k}^{l-1}. \tag{59}$$

   end for

4. Compute the posterior at k time step

$$\omega_{i,k}^L = \begin{cases} 1/b_{i,k}^L & if\ b_{i,k}^L \neq 0, \\ 1 & otherwise, \end{cases} \tag{60}$$

$$y_{i,k} = \hat{y}_{i,k}^L + \omega_{i,k}^L \phi_{i,k}^L, \tag{61}$$

$$S_{i,k}^y = qr\{[S_{i,k}^L;\ \sqrt{\omega_{i,k}^L} \Phi_{i,k}^L]\}, \tag{62}$$

$$x_{i,k} = (S_{i,k}^y)^{-T}(S_{i,k}^y)^{-1} y_{i,k}. \tag{63}$$

To keep the consistence of local filters, the inverse of an estimated factor S/N can be used for weighting information contributions derived by

$$\omega_{i,k}^l = \begin{cases} 1/b_{i,k}^l & if\ b_{i,k}^l \neq 0, \\ 1 & otherwise, \end{cases} \tag{64}$$

where $b_{i,k}^l$ is an estimation of the fraction S/N at l_{th} iterative step via the consensus algorithm

$$b_{i,k}^l = \sum_{j \in \mathcal{N}_i} \epsilon_{i,j,k} b_{j,k}^{l-1}. \tag{65}$$

The initialization of $b_{i,k}^0$ is set as 1 if the i_{th} node can observe the target, i.e., $i \in \mathcal{S}$, where $\mathcal{S}$ means the set of these effective nodes, whereas it is set as 0 for other nodes as shown in (54). Here, the consensus weight $\epsilon_{i,j,k}$ employs the Metropolis weight. In such a way, the quantity $\omega_{i,k}^L \epsilon_{i,j,k} \leq 1$ for any pair of (i,j) since $b_{i,k}^L = \sum_{j \in \mathcal{N}_i} \epsilon_{i,j,k}$. Therefore, the information contributions will not get overweighted and the consistence of the local filter has been kept [11]. In addition, compared to the IWCFs, no prior knowledge of network connections is required here. To summarize, based on this new weighting schemes, the square-root central difference dynamic hybrid consensus filter (SRCDHCF) can be derived as in Algorithm 3. In a similar way, the square-root unscented dynamic hybrid consensus filter (SRUDHCF) and the cubature dynamic hybrid consensus filter (SRCDHCF) can be further developed. However, the DHCFs require more consensus quantities to be shared with other neighbor nodes than IWCFs, i.e., prior information $(\hat{y}_{i,k}^{l-1}, \hat{S}_{i,k}^{l-1})$, information contributions $(\phi_{i,k}^{l-1}, \Phi_{i,k}^{l-1})$ and $b_{i,k}^{l-1}$, which means the algorithms of DHCFs require more data bandwidth.

4. Simulation

To show the performance of the proposed square-root sigma-point information consensus filters, a sparse network is simulated using nine cameras (c_1 to c_9) as shown in [12], which has a maximum connection degree $\Delta_{max} = 2$. Each camera has a fixed field of view (FOV), i.e., 200×200. When the target moves, some of the camera nodes may lose the measurement information. Therefore, the estimated value S/N for the algorithm DHCFs is time varying,

In this simulation, the state transition model and measurement model of the object are nonlinear functions. The state of the target is defined as $\mathbf{x}_k = [x_k, y_k, v_{x,k}, v_{y,k}, \delta_k]^T$, and the motion model of the target is given as

$$\mathbf{x}_{k+1} = \begin{bmatrix} x_k + v_{x,k}\delta_k + a_x\delta_k^2/2 \\ y_k + v_{y,k}\delta_k + a_y\delta_k^2/2 \\ v_{x,k} + a_x\delta_k \\ v_{y,k} + a_y\delta_k \\ \delta_k + e \end{bmatrix}, \tag{66}$$

where (x_k, y_k), $(v_{x,k}, v_{y,k})$ and (a_x, a_y) are the position, velocity and acceleration of the target, respectively. The acceleration is modeled as Gaussian noise. δ_k is the time step between two consecutive measurements. The synchronization error among cameras is also considered as a Gaussian variable e. The vector $w = [a_x, a_y, e]^T$ is considered as the Gaussian noise vector with zero mean and covariance $Q = diag\{1, 1, 0.01\}$. We consider a nonlinear measurement model of the camera node i as

$$\mathbf{z}_{i,k} = \begin{bmatrix} u_{i,k} \\ v_{i,k} \end{bmatrix} = \begin{bmatrix} \frac{H_{11}x_k + H_{12}y_k + H_{13}}{H_{31}x_k + H_{32}y_k + H_{33}} \\ \frac{H_{21}x_k + H_{22}y_k + H_{23}}{H_{31}x_k + H_{32}y_k + H_{33}} \end{bmatrix} + \mathbf{v}_{i,k}, \tag{67}$$

where $(u_{i,k}, v_{i,k})$ is the pixel coordinates of the target in the image, and $(H_{11}, \cdots, H_{33})$ are elements of the homography matrix defined as

$$H = \begin{bmatrix} 397.2508 & 95.2020 & 287280 \\ 51.7437 & 396.9189 & 139100 \\ 0.0927 & 0.1118 & 605.2481 \end{bmatrix}, \tag{68}$$

which are taken from one of the cameras of the APIDIS dataset [12].

4.1. Normal Measurement

The measurement noise $v_{i,k}$ is defined as Gaussian noise with zero mean and covariance $R = diag\{15, 15\}$. The initial covariance matrix of the state for each camera node is set to be a diagonal matrix $P = diag\{0.1, 0.1, 0.1, 0.1, 0.0001\}$.
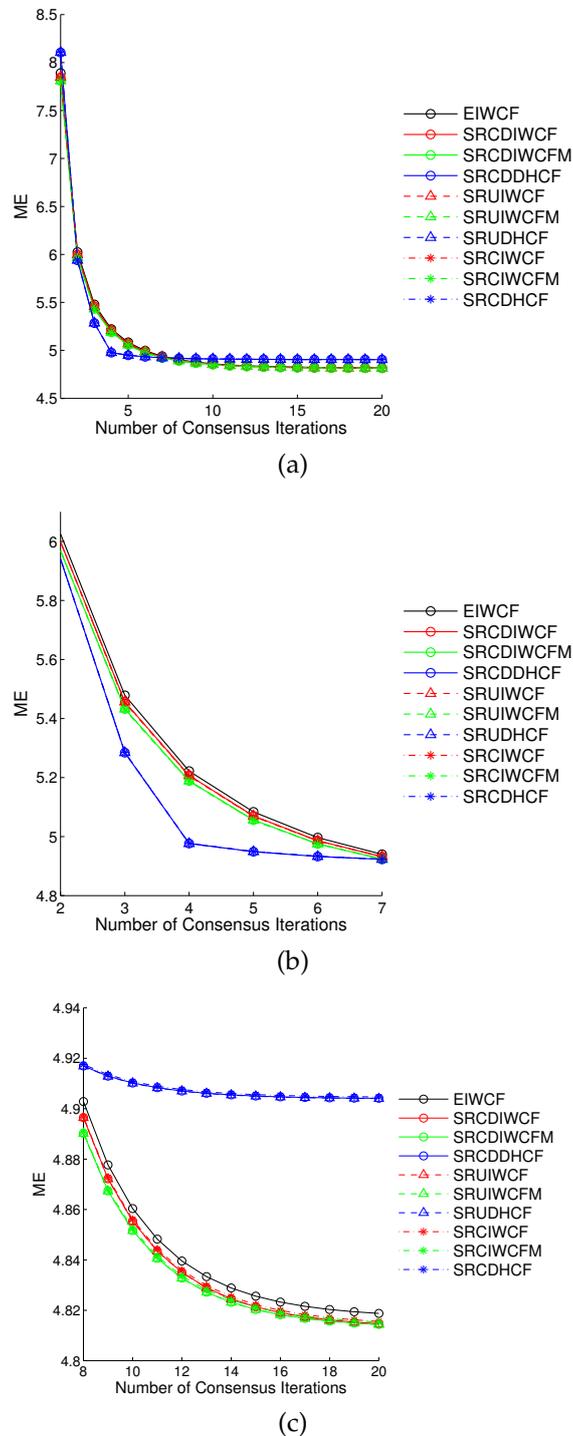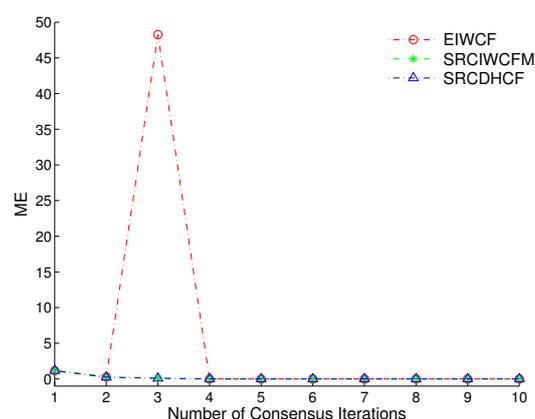

Figure 1. The comparison of mean errors (ME) of proposed square-root sigma-point information consensus filters and the state-of-the-art extended information weighted consensus filter (EIWCF) for 50 Monte Carlo simulations. The details of the (**a**) is shown in (**b**) for iterations 2–7 and (**c**) for iterations 8–20.

We here demonstrate the simulation for 50 Monte Carlo runs, and the result is shown in Figure 1. Figure 1a shows the overall performance comparison for the number of consensus iterations L from 1 to 20, whereas the Figure 1b shows the first half part of Figure 1a for L from 2 to 7, and the Figure 1c shows the second half part of Figure 1a for L from 8 to 20. By analyzing the result data, we can derive following conclusions: first, the square-root sigma-point information consensus filters outperform the EIWCF as shown in Figure 1b,c, since they can capture higher order terms of Taylor expansion. Because the dimension of the state is small, the differences between the Stirling's interpolation, unscented transform and cubature rule based methods are minor, e.g., the SRCDDHCF, SRUDHCF and SRCDHCF are almost overlapped in Figure 1 (blue lines). Second, the DHCFs have a faster convergence rate than IWCFs, i.e., the DHCFs are close to a convergence state at the iteration $L = 4$ as shown in Figure 1b. The reason is that IWCFs based methods can overweight some sensor nodes, which can decrease the convergence rate for smaller consensus iteration value L. Third, the IWCFs can achieve more accurate results than DHCFs by running more iterations, e.g., $L > 7$. Fourth, the Metropolis weight indeed can improve the performance of IWCFs compared with the fixed one, e.g., SRCDIWCFM, SRUIWCFM and SRCIWCFM using Metropolis weight have better performance than SRCDIWCF, SRUIWCF and SRCIWCF using $0.65/\Delta_{max}$. Therefore, the Metropolis weight is more preferred for square-root sigma-point information consensus filters.

4.2. Ill Condition: Near Perfect Measurement

The ill condition can lead the filter system to be unstable due to the accumulated round-off errors in the computational system. The square-root filters have shown their advantages with handling such problems in previous publications [13,14,25,26]. Here, we use a similar idea to compare the performance of SRCIWCFM, SRCDHCF and original EIWCF by setting the measurement covariance to be a very small value $R = diag\{2^{-52}, 2^{-52}\}$ for a normal 32-bit operational system. The simulations are demonstrated for 20 Monte Carlo runs with the iterations $L = 1, \cdots, 10$. The experimental results are shown in Figure 2. We can see that the performance of SRCIWCFM and SRCDHCF are very close to each other, whereas the EIWCF becomes unstable when the iteration number is equal to 3. In addition, the simulation software MATLAB (R2010a, The MathWorks Inc, Natick, MA, USA) always shows the warning message "Matrix is close to singular or badly scaled" when the EIWCF algorithm runs with the near perfect measurements. In contrast, the SRCIWCFM and SRCDHCF are robust to such ill conditions and converge to a very stable result.

Figure 2. The comparison of mean errors (ME) of square-root cubature information weighted consensus filter with Metropolis weights (SRCIWCFM), square-root cubature dynamic hybrid consensus filter (SRCDHCF) and the original EIWCF for 20 Monte Carlo simulations.

5. Conclusions

In this paper, we proposed two kinds of square-root sigma-point information consensus filters, which are based on the IWCF and DHCF methods, respectively. By comparing to the state-of-the-art method, the proposed methods can achieve more accurate estimation results than the EIWCF by using Stirling's interpolation, unscented transform and cubature rules for linearization of nonlinear models. Furthermore, we also show that the DHCF based methods are preferred for a small number of consensus iterations, since they can achieve faster convergence rate. However, the IWCF based methods can offer more accurate results when the number of consensus iterations is large enough, especially when the redundancy information between the sensor nodes is equally distributed. Currently, the topology of sensor networks in our experiment has a fixed connection, which is a limitation if the sensors are dynamic, e.g., robots. An interesting and possible future extension of current work is using an M-matrix approach to handle such issues [30,31].

Author Contributions: Guoliang Liu conceived and designed the experiments; Guoliang Liu performed the experiments; Guoliang Liu analyzed the data; Guohui Tian contributed analysis tools; and Guoliang Liu wrote the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kamal, A.T.; Farrell, J.A.; Roy-Chowdhury, A.K. Information weighted consensus filters and their application in distributed camera networks. *IEEE Trans. Autom. Control* **2013**, *58*, 3112–3125.

2. Casbeer, D.W.; Beard, R.W. Multi-static radar target tracking using information consensus filters. In Proceedings of the AIAA Guidance, Navigation, and Control Conference, Chicago, IL, USA, 10–13 August 2009; pp. 1–9.

3. Wang, L.; Zhang, Q.; Zhu, H.; Shen, L. Adaptive consensus fusion estimation for MSN with communication delays and switching network topologies. In Proceedings of the 2010 49th IEEE Conference on Decision and Control (CDC), Atlanta, GA, USA, 15–17 December 2010; pp. 2087–2092.

4. Olfati-Saber, R.; Murray, R.M. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control* **2004**, *49*, 1520–1533.

5. Aragues, R.; Cortes, J.; Sagues, C. Distributed consensus on robot networks for dynamically merging feature-based maps. *IEEE Trans. Robot.* **2012**, *28*, 840–854.

6. Li, W.; Jia, Y. Consensus-based distributed multiple model ukf for jump markov nonlinear systems. *IEEE Trans. Autom. Control* **2012**, *57*, 230–236.

7. Olfati-Saber, R.; Fax, J.A.; Murray, R.M. Consensus and cooperation in networked multi-agent systems. *Proc. IEEE* **2007**, *95*, 215–233.

8. Kamal, A.T.; Farrell, J.A.; Roy-Chowdhury, A.K. Information weighted consensus. In Proceedings of the IEEE Conference on Decision and Control, Maui, HI, USA, 10–13 December 2012; pp. 2732–2737.

9. Olfati-Saber, R. Distributed Kalman filter with embedded consensus filters. In Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference (CDC-ECC '05), Seville, Spain, 12–15 December 2005; pp. 8179–8184.

10. Casbeer, D.W.; Beard, R. Distributed information filtering using consensus filters. In Proceedings of the American Control Conference, Hyatt Regency Riverfront, St. Louis, MO, USA, 10–12 June 2009; pp. 1882–1887.

11. Battistelli, G.; Chisci, L.; Mugnai, G.; Farina, A.; Graziano, A. Consensus-based linear and nonlinear filtering. *IEEE Trans. Autom. Control* **2015**, *60*, 1410–1415.

12. Katragadda, S.; Sanmiguel, J.C.; Cavallaro, A. Consensus protocols for distributed tracking in wireless camera networks. In Proceedings of the 2014 17th International Conference on Information Fusion (FUSION), Salamanca, Spain, 7–10 July 2014; pp. 1–8.

13. Chen, Y.; Zhao, Q. A Novel Square-Root Cubature Information Weighted Consensus Filter Algorithm for Multi-Target Tracking in Distributed Camera Networks. *Sensors* **2015**, *15*, 10526–10546.

14. Liu, G.; Woergoetter, F.; Markelic, I. Square-Root Sigma-Point Information Filtering. *IEEE Trans. Autom. Control* **2012**, *57*, 2945–2950.

15. Guo, Z.; Li, S.; Wang, X.; Heng, W. Distributed Point-Based Gaussian Approximation Filtering for Forecasting-Aided State Estimation in Power Systems. *IEEE Trans. Power Syst.* **2016**, *31*, 2597–2608.

16. Xiao, L.; Boyd, S.; Lall, S. A scheme for robust distributed sensor fusion based on average consensus. In Proceedings of the International Conference on Information Processing in Sensor Networks, Los Angeles, CA, USA, 25–27 April 2005; pp. 63–70.

17. Ilic, N.; Stankovic, M.S.; Stankovic, S.S. Adaptive consensus-based distributed target tracking in sensor networks with limited sensing range. *IEEE Trans. Control Syst. Technol.* **2014**, *22*, 778–785.

18. Ito, K.; Xiong, K. Gaussian filters for nonlinear filtering problems. *IEEE Trans. Autom. Control* **2000**, *45*, 910–927.

19. Merwe, R.V.D. Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models. Ph.D. Thesis, Oregen Health & Science University, Portland, OR, USA, 2004.

20. Julier, S.J.; Uhlmann, J.K. Unscented filtering and nonlinear estimation. *Proc. IEEE* **2004**, *92*, 401–422.

21. Lee, D.j. Nonlinear Estimation and Multiple Sensor Fusion Using Unscented Information Filtering. *IEEE Signal Process. Lett.* **2008**, *15*, 861–864.

22. Arasaratnam, I.; Haykin, S. Cubature Kalman Filters. *IEEE Trans. Autom. Control* **2009**, *54*, 1254–1269.

23. Liu, Y.; He, Y.; Wang, H. Squared-root cubature information consensus filter for non-linear decentralised state estimation in sensor networks. *IET Radar Sonar Navig.* **2014**, *8*, 931–938.

24. Van der Merwe, R.; Wan, E.A. The square-root unscented Kalman filter for state and parameter-estimation. In Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, USA, 7–11 May 2001; Volume 6, pp. 3461–3464.

25. Bharani Chandra, K.P.; Chandra, B.; Gu, D.W.; Postlethwaite, I. Square root cubature information filter. *IEEE Sens. J.* **2013**, *13*, 750–758.

26. Arasaratnam, I. Sensor Fusion with Square-Root Cubature Information Filtering. *Intell. Control Autom.* **2013**, *2013*, 11–17.

27. Särkkä, S.; Solin, A. On continuous-discrete cubature Kalman filtering. In Proceedings of the 2012 16th IFAC Symposium on System Identification (SYSID 2012), Brussels, Belgium, 11–13 July 2012; Volume 16, pp. 1210–1215.

28. Wang, S.; Feng, J.; Tse, C.K. A class of stable square-root nonlinear information filters. *IEEE Trans. Autom. Control* **2014**, *59*, 1893–1898.

29. Chen, Y.; Zhao, Q.; An, Z.; Lv, P.; Zhao, L. Distributed Multi-Target Tracking Based on the K-MTSCF Algorithm in Camera Networks. *IEEE Sens. J.* **2016**, *16*, 5481–5490.

30. Wen, G.; Duan, Z.; Chen, G.; Yu, W. Consensus Tracking of Multi-Agent Systems With Lipschitz-Type Node Dynamics and Switching Topologies. *IEEE Trans. Circuits Syst. I* **2013**, *61*, 1–13.

31. Wen, G.; Yu, W.; Hu, G.; Cao, J.; Yu, X. Pinning synchronization of directed networks with switching topologies: A multiple lyapunov functions approach. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 3239–3250.