SURVEY OF ESTIMATE FUSION APPROACHES

Jiří Ajgl and Miroslav Šimandl
Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia in Pilsen
Univerzitní 8, Plzeň, 306 14, Czech Republic

Keywords: Dynamic systems, State estimation, Optimal estimation, Sensor fusion, Filtering problems.

Abstract: The paper deals with fusion of state estimates of stochastic dynamic systems. The goal of the contribution is to present main approaches to the estimate fusion which were developed during the last four decades. The hierarchical and decentralised estimation are presented and main special cases are discussed. Namely the following approaches, the distributed Kalman filter, maximum likelihood, channel filters, and the information measure, are introduced. The approaches are illustrated in numerical examples.

1 INTRODUCTION

The classical estimation theory deals with estimating the value of some attribute by using measured data. (Simon, 2006) reviews the optimal state estimation techniques for linear systems and their extension to non-linear systems. However, there are other dimensions of the estimation problem. The direction discussed here is the multisensor problem that assumes the system state to be estimated by multiple estimators. Each estimator uses different data sets and it can communicate its estimate to the other estimators. The question is how to combine multiple estimates to obtain optimal results.

The key issues in the multisensor fusion are communication and dependences. In practise, it is possible to communicate raw measurements among estimators. In such a case, each estimator can process the measurements only and no estimate fusion is required. But in the case of the on-line state estimation of dynamic systems the out-of-sequence problems occurs. Updating the estimate by an old measurement is complicated, see (Bar-Shalom, 2002) or (Challa et al., 2003). Moreover, in general network of estimators, the estimators must log a list of all measurements they have processed or the measurement must be passed with a list of estimators that have processed it. Otherwise the multiple processing of the same data is inevitable.

If two estimators use measurements with dependent errors, their estimates will be dependent. A non-zero state noise causes dependence of the estimates as well as the communication of the estimates with the consequent fusion. The fused estimate and the estimates before fusion are obviously dependent. In a rooted tree estimator network, some restarts of the estimators can be applied to solve the communication dependence problem, see (Chong et al., 1999).

In the fusion point of view, the classical estimation is named as centralised. A central estimator processes raw measurements only. If the estimators are organised in a rooted tree, the root is called a fusion centre and the fusion is denoted as hierarchical or distributed. If there is not a fusion centre, the fusion is decentralised. Only a local knowledge of the network is usually assumed in these cases. The above mentioned approaches have been introduced in the literature by different ways during last decades. However, a unique survey of the approaches is missing.

Therefore, the aim of the paper is to give a survey of main results in estimate fusion and to show numerical illustrations. Both hierarchical and decentralised estimation are presented and discussed. In the hierarchical framework, namely the distributed Kalman filter and the fusion based on the maximum likelihood estimation are considered. In the decentralised framework, the stress is laid on the channel filters and the information measure approach.

The paper is organised as follows. Section 2 defines the fusion problem, section 3 and 4 discuss the hierarchical and decentralised approaches, respectively. A numerical example is given in section 5 and finally section 6 summarises the fusion problems.

Ajgl J. and Šimandl M. (2010).
SURVEY OF ESTIMATE FUSION APPROACHES.
In *Proceedings of the 7th International Conference on Informatics in Control, Automation and Robotics*, pages 191-196
DOI: 10.5220/0002947201910196

2 PROBLEM STATEMENT

Let the discrete-time stochastic system be described by state transition and measurement conditional probability density functions

$$p(\mathbf{x}_{k+1}|\mathbf{x}_k), \quad (1)$$

$$p(\mathbf{z}_k^{(1)}, \mathbf{z}_k^{(2)}, \ldots, \mathbf{z}_k^{(N)}|\mathbf{x}_k). \quad (2)$$

where $\mathbf{z}_k^{(j)}$, $j = 1, \ldots, N$, are local measurements at time k, $k = 0, 1, \ldots$ and the initial condition $p(\mathbf{x}_0)$ is known. Let the system be linear gaussian. In such case, analytical solutions to estimation problems exist. The linear gaussian system can be described by state and measurement equations

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{G}\mathbf{w}_k, \quad (3)$$

$$\mathbf{z}_k^{(j)} = \mathbf{H}^{(j)}\mathbf{x}_k + \mathbf{v}_k^{(j)}, j = 1, \ldots, N, \quad (4)$$

where $\mathbf{F} \in \mathbb{R}^{n_x \times n_x}$, $\mathbf{H}^{(j)} \in \mathbb{R}^{n_z^{(j)} \times n_x}$, and $\mathbf{G} \in \mathbb{R}^{n_x \times n_w}$ are known matrices, $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is the immeasurable system state and $\mathbf{z}_k^{(j)} \in \mathbb{R}^{n_z^{(j)}}$ is the local measurement coming from j-th sensor. The variables $\mathbf{w}_k \in \mathbb{R}^{n_w}$ and $\mathbf{v}_k^{(j)} \in \mathbb{R}^{n_z^{(j)}}$ represent the state and measurement white Gaussian noises with zero mean and with known covariance matrices $\mathbf{Q}$, $\mathbf{R}^{(jj)}$, respectively. The processes $\{\mathbf{v}_k^{(j)}\}$ are independent of the process $\{\mathbf{w}_k\}$ and all of them are independent on the system initial state described by the Gaussian pdf $p(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0 : \bar{\mathbf{x}}_0, \mathbf{P}_0)$. The measurement error processes $\{\mathbf{v}_k^{(j)}\}$ can be generally mutually dependent, with cross-correlations ,$\mathbf{R}_k^{(ij)} = \mathrm{E}(\mathbf{v}_k^{(i)}\mathbf{v}_k^{(j)\mathrm{T}})$, but there are often assumed to be independent, $\mathbf{R}_k^{(ij)} = 0$ for $i \neq j$.

Let each sensor have its estimator, i.e. there exist N state estimates $\hat{\mathbf{x}}^{(j)}$, $j = 1, \ldots, N$, with corresponding error covariance matrices $P^{(j)}$. The estimators are connected with some others by data link. The communication network can be described by a directed graph with nodes in each sensor and with edges representing the oriented data links. It is assumed that measurements coming from other sensor nodes can not be processed directly, e.g. due to the unknown measurement equation of the respective sensors, or the communication of the measurements would be ineffective. So it is assumed that only the estimates are communicated. The goal of the fusion is to combine local estimates.

3 HIERARCHICAL FUSION

In the hierarchical fusion, the local estimates are communicated to a fusion centre. The method are based on the classical one-sensor estimation, which is described in subsection 3.1. The distributed Kalman filter extracts independent information from the estimates and is discussed in subsection 3.2. In the maximum likelihood approach, the estimates are regarded as dependent measurements. The respective fusion is shown in subsection 3.3.

3.1 Optimal Centralised Estimate

In the case of one sensor system, there is no fusion of estimates. The classical Kalman filter solution is the exact Bayessian solution to the filtering problem for a linear Gaussian system. You can see (Simon, 2006) for many numerical approximations to the exact solution for non-linear systems. The Kalman filter estimate is a standard against which other methods can be compared. The filtering (measurement update) equations

$$\mathbf{P}_{k|k}^{-1}\hat{\mathbf{x}}_{k|k} = \mathbf{P}_{k|k-1}^{-1}\hat{\mathbf{x}}_{k|k-1} + \mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{z}_k, \quad (5)$$

$$\mathbf{P}_{k|k}^{-1} = \mathbf{P}_{k|k-1}^{-1} + \mathbf{H}_k^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{H}_k, \quad (6)$$

can be interpreted as a fusion of the predictive estimate with the information based on the last measurement only. The prediction (time update) equations

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}_k\hat{\mathbf{x}}_{k|k}, \quad (7)$$

$$\mathbf{P}_{k+1|k} = \mathbf{F}_k\mathbf{P}_{k|k}\mathbf{F}_k^{\mathrm{T}} + \mathbf{Q}_k. \quad (8)$$

correspond to the dynamics of the system. If more explicit notation is required further in this article, the general conditional pdf notation will be used. The exact Bayessian solution is given by

$$p(\mathbf{x}_k|\mathbf{z}_k, \mathbf{Z}_{k-1}) \propto p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Z}_{k-1}) \quad (9)$$

$$p(\mathbf{x}_{k+1}|\mathbf{Z}_k) = \int_R p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Z}_k)\mathrm{d}\mathbf{x}_k \quad (10)$$

where $\propto$ means proportional to and $\mathbf{Z}_k \triangleq \{\mathbf{z}_k, \mathbf{Z}_{k-1}\}$ denotes the set of all measurements up to the time k.

The centralised estimator is a hypothetical estimator which assumes that all measurements are immediately available to the estimator and that the correspondent measurement equations are known at the centre. The local measurement equations (4) can be merged to one equation with

$$\mathbf{z}_k = \begin{bmatrix} \mathbf{z}_k^{(1)} \\ \vdots \\ \mathbf{z}_k^{(N)} \end{bmatrix}, \quad \mathbf{H}_k = \begin{bmatrix} \mathbf{H}_k^{(1)} \\ \vdots \\ \mathbf{H}_k^{(N)} \end{bmatrix}, \quad \mathbf{v}_k = \begin{bmatrix} \mathbf{v}_k^{(1)} \\ \vdots \\ \mathbf{v}_k^{(N)} \end{bmatrix}, \quad (11)$$

$\mathbf{R}_k = [\mathbf{R}_k^{(ij)}]_{i,j=1}^N$. The centralised Kalman filter is given by (5)-(8) and (11).

3.2 Distributed Kalman Filter

The distributed Kalman filter consists of N local Kalman filters which send their estimates to one fusion centre. It is also possible to distribute the local filters recursively. The name hierarchical Kalman filter is also used. Note that the term decentralised is misused in the literature to express that this filter is not the centralised one.

The main assumption is the independence of the local measurement errors,

$$\mathbf{R}_k^{(ij)} = 0, \quad i \neq j. \tag{12}$$

Then the pieces of information gained from the same time measurements are independent and can be simply summed up. The fusion centre filtering equation can be derived from (5), (6) with the use of (11) as

$$\mathbf{P}_{k|k}^{-1}\hat{\mathbf{x}}_{k|k} = \mathbf{P}_{k|k-1}^{-1}\hat{\mathbf{x}}_{k|k-1} + \\ + \sum_{j=1}^{N}\left(\mathbf{P}_{k|k}^{(j)\,-1}\hat{\mathbf{x}}_{k|k}^{(j)} - \mathbf{P}_{k|k-1}^{(j)\,-1}\hat{\mathbf{x}}_{k|k-1}^{(j)}\right), \tag{13}$$

$$\mathbf{P}_{k|k}^{-1} = \mathbf{P}_{k|k-1}^{-1} + \sum_{j=1}^{N}\left(\mathbf{P}_{k|k}^{(j)\,-1} - \mathbf{P}_{k|k-1}^{(j)\,-1}\right), \tag{14}$$

where indexes $^{(j)}$ denotes the local estimates. The fusion centre predictive equations are identical to (7), (8). It is possible to compute the predictive estimates at each local estimator, but it requires to send predictive estimate to the fusion centre. Instead of that, the fusion centre predictive can be send to each local estimator where it replaces the local estimate

$$\hat{\mathbf{x}}_{k+1|k}^{(j)} \leftarrow \hat{\mathbf{x}}_{k+1|k}, \quad \mathbf{P}_{k+1|k}^{(j)} \leftarrow \mathbf{P}_{k+1|k}, \tag{15}$$

$j = 1, \ldots, N$. This feedback brings the globally optimal estimate to each local estimator and the estimation is expected to be better if the extension to nonlinear systems approximated by linearisation is considered.

The distributed Kalman filter for the system with dependent noises is discussed in (Hashemipour et al., 1988). (Berg and Durrant-Whyte, 1992) minimise the communication by reducing the dimension of the estimated state at each local estimator and using internodal transformations; there is no communication of the state components that are not influenced by the measurement.

The fusion centre filtering equations (13), (14) can be written by the conditional densities as

$$p(\mathbf{x}_k|\mathbf{z}_k, \mathbf{Z}_{k-1}) \propto p(\mathbf{x}_k|\mathbf{Z}_{k-1}) \prod_{j=1}^{N} \frac{p(\mathbf{x}_k|\mathbf{z}_k^{(j)}, \mathbf{Z}_{k-1})}{p(\mathbf{x}_k|\mathbf{Z}_{k-1})}, \tag{16}$$

where the feedback is given by

$$p(\mathbf{x}_k|\mathbf{z}_k^{(j)}, \mathbf{Z}_{k-1}) \leftarrow p(\mathbf{x}_k|\mathbf{Z}_k), \tag{17}$$

$j = 1, \ldots, N$, and is analogous to (15). Note that the division by the predictive density $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ can not be easily extended to general non-Gaussian densities.

3.3 Fusion by the Maximum Likelihood

This subsection discusses the fusion of dependent estimates at a fusion centre. The cornerstone idea is to treat the local estimates as if they were measurements. It arises from the identity, see (Li et al., 2003),

$$\hat{\mathbf{x}}_{k|k}^{(j)} = \mathbf{x}_k + (\hat{\mathbf{x}}_{k|k}^{(j)} - \mathbf{x}_k) = \mathbf{x}_k + (-\tilde{\mathbf{x}}_{k|k}^{(j)}) \tag{18}$$

where $\tilde{\mathbf{x}}_{k|k}^{(j)}$ is the error of the estimate at the j-th estimator. The covariance matrices of these measurements are the error covariance matrices $\mathbf{P}_{k|k}^{(jj)} = \mathbf{P}_{k|k}^{(j)}$. Assuming the local estimates are obtained by Kalman filters with Kalman gains $\mathbf{K}_k^{(j)} = \mathbf{P}_{k|k}^{(jj)}\mathbf{H}_k^{(j)\mathrm{T}}\mathbf{R}_k^{(j)-1}$, the cross-covariances $\mathbf{P}_{k|k}^{(ij)} = \mathrm{E}(\tilde{\mathbf{x}}_{k|k}^{(i)}\tilde{\mathbf{x}}_{k|k}^{(j)\mathrm{T}})$ are given by

$$\mathbf{P}_{k|k}^{(ij)} = (\mathbf{I}_{n_x} - \mathbf{K}_k^{(i)}\mathbf{H}_k^{(i)})\mathbf{P}_{k|k-1}^{(ij)}(\mathbf{I}_{n_x} - \mathbf{K}_k^{(j)}\mathbf{H}_k^{(j)})^{\mathrm{T}} + \\ + \mathbf{K}_k^{(i)}\mathbf{R}_k^{(ij)}\mathbf{K}_k^{(j)\mathrm{T}}, \tag{19}$$

where $\mathbf{I}_{n_x}$ is the identity matrix of the size n_x, with the initial condition $\mathbf{P}_{0|-1}^{(ij)} = \mathbf{P}_0$. The predictive covariance $\mathbf{P}_{k|k-1}^{(ij)}$ is computed by (8).

Then the fusion centre measurement equation is given by

$$\mathbf{z}_k^{FC} = \mathbb{I}_N \mathbf{x}_k + \xi_k \tag{20}$$

where $\mathrm{cov}(\xi_k) = \mathbf{P}_k = [\mathbf{P}_{k|k}^{(ij)}]_{i,j=1}^{N}$ and

$$\mathbf{z}_k^{FC} = \begin{bmatrix} \hat{\mathbf{x}}_{k|k}^{(1)} \\ \vdots \\ \hat{\mathbf{x}}_{k|k}^{(N)} \end{bmatrix}, \; \mathbb{I}_N = \begin{bmatrix} \mathbf{I}_{n_x} \\ \vdots \\ \mathbf{I}_{n_x} \end{bmatrix}, \; \xi_k = \begin{bmatrix} -\tilde{\mathbf{x}}_{k|k}^{(1)} \\ \vdots \\ -\tilde{\mathbf{x}}_{k|k}^{(N)} \end{bmatrix}. \tag{21}$$

Unfortunately, the process $\{\xi_k\}$ is correlated with $\mathbf{x}_k$ and it is coloured, so it is not possible to use a Kalman filter in the fusion centre. But the central estimate can be obtained, see (Chang et al., 1997), by the maximum likelihood method

$$\hat{\mathbf{x}}_{k|k} = (\mathbb{I}_N^{\mathrm{T}}\mathbf{P}_k^{-1}\mathbb{I}_N)^{-1}\mathbb{I}_N^{\mathrm{T}}\mathbf{P}_k^{-1}\mathbf{z}_k^{FC}, \tag{22}$$

$$\mathbf{P}_{k|k} = (\mathbb{I}_N^{\mathrm{T}}\mathbf{P}_k^{-1}\mathbb{I}_N)^{-1}. \tag{23}$$

Note that the above fusion requires to send the Kalman filter gains $\mathbf{K}_k^{(j)}$, $j = 1, \ldots, N$ to the fusion centre to compute the cross-correlations of the estimates (19). The measurement matrices $\mathbf{H}_k^{(j)}$ must be known at or sent to the fusion centre also.

4 DECENTRALISED FUSION

In the decentralised fusion, information is processed localy. The channel filters enable to obtain a globally optimal solution in a tree network and they are described in subsection 4.1. The information measure approach discussed in subsection 4.2 sacrifices the Bayessian optimality for the the possibility to be easily used in an arbitrary network.

4.1 Channel Filters

The principle of the channel filter approach, that was introduced in (Grime and Durrant-Whyte, 1994), is the same as that of the distributed Kalman filter in fact. The new information is extracted and summed up. The necessary condition is that there is one and only one way of the information propagation, i.e. the network structure is a tree. The density notation will be used to explicitly denote the set of the measurements that were exploited by each estimator.

The essential rule of the estimate fusion is

$$p(\mathbf{x}_k|\mathcal{Z}_A \cup \mathcal{Z}_B) = \frac{p(\mathbf{x}_k|\mathcal{Z}_A)p(\mathbf{x}_k|\mathcal{Z}_B)}{p(\mathbf{x}_k|\mathcal{Z}_A \cap \mathcal{Z}_B)}. \quad (24)$$

The posterior probability density function of the state conditioned on the union of two measurement sets is equal to the product of the densities conditioned on each measurement set divided by the density conditioned on the intersection of the measurement sets.

The equation (24) is the core of the channel filters. It is assumed that all local measurement errors are independent, (12). Thus, the measurement density can be factorised, $p(\mathbf{z}_k^{(1)}, \mathbf{z}_k^{(2)}, \ldots, \mathbf{z}_k^{(N)}|\mathbf{x}_k) = \prod_{j=1}^{N} p(\mathbf{z}_k^{(j)}|\mathbf{x}_k)$.

First, all local estimators filter their predictive estimates according to (9). Then the filtering estimates are communicated to the neighbouring estimators. The fusion is given by a repeated use of the fusion rule (24) as

$$p(\mathbf{x}_k|\mathcal{Z}_k^j) = p(\mathbf{x}_k|\mathbf{z}_k^{(j)}, \mathcal{Z}_{k-1}^j) \prod_{i \in \mathcal{N}_j} \frac{p(\mathbf{x}_k|\mathbf{z}_k^{(i)}, \mathcal{Z}_{k-1}^i)}{p(\mathbf{x}_k|\mathcal{Z}_{k-1}^j \cap \mathcal{Z}_{k-1}^i)}, \quad (25)$$

where $\mathcal{Z}_k^j = (\mathcal{Z}_{k-1}^j \cup \mathbf{z}_k^{(j)}) \bigcup_{i \in \mathcal{N}_j} (\mathcal{Z}_{k-1}^i \cup \mathbf{z}_k^{(i)})$ is the set of the measurements that were exploited by the j-th estimator at the time k after the fusion with the incoming estimates $p(\mathbf{x}_k|\mathbf{z}_k^{(i)}, \mathcal{Z}_{k-1}^i)$, $\mathcal{N}_j$ is the set of the neighbours of the j-th estimator that have sent their estimates to it, and $p(\mathbf{x}_k|\mathcal{Z}_{k-1}^j \cap \mathcal{Z}_{k-1}^i)$ is the estimate of the channel filter ij. The fusion (25) uses the fact that the measurement errors are independent and thus

$$(\mathcal{Z}_{k-1}^i \cup \mathbf{z}_k^{(i)}) \cap (\mathcal{Z}_{k-1}^j \cup \mathbf{z}_k^{(j)}) = \mathcal{Z}_{k-1}^i \cap \mathcal{Z}_{k-1}^j. \quad (26)$$

The predictive estimates are computed according to (10) and the channel filter estimate is given by

$$p(\mathbf{x}_k|\mathcal{Z}_k^j \cap \mathcal{Z}_k^i) = \frac{p(\mathbf{x}_k|\mathbf{z}_k^{(j)}, \mathcal{Z}_{k-1}^j)p(\mathbf{x}_k|\mathbf{z}_k^{(i)}, \mathcal{Z}_{k-1}^i)}{p(\mathbf{x}_k|\mathcal{Z}_{k-1}^j \cap \mathcal{Z}_{k-1}^i)} \quad (27)$$

where the equations (24), (26) and the relation

$$(\mathcal{Z}_{k-1}^i \cup \mathbf{z}_k^{(i)}) \cup (\mathcal{Z}_{k-1}^j \cup \mathbf{z}_k^{(j)}) = \mathcal{Z}_k^i \cap \mathcal{Z}_k^j \quad (28)$$

were used.

The local estimates equal to centralised estimates with delayed measurements. The delays are given by the length of the path between the respective sensors decreased by one. Note that the division by the channel filter density in the equations (25) and (27) is easily tractable for Gaussian densities only.

4.2 Information Measure Approach

In general networks, the optimality cannot be reached without inadequate effort. It can be impossible to decide which measurements have been used to compute the estimates. And even if this is possible, the common information in the denominator of (24) is too complicated to find and to compute with. Multiple processing of the same measurements, with the illusion that the errors are independent, is inevitable. Therefore to not underestimate the estimate error, some bounds must be used.

The idea of the Covariance Intersection method, see (Julier, 2009) for example, arises from the geometrical interpretation of the estimates. The fused estimate $\{\hat{\mathbf{x}}, \mathbf{P}\}$ is required to be consistent, i.e. the error covariance must not be underestimated, $\mathbf{P} - \mathrm{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T] \geq 0$, where $\mathbf{x}$ denotes the true state. Assuming the local estimates $\{\hat{\mathbf{x}}_1, \mathbf{P}_1\}$, $\{\hat{\mathbf{x}}_2, \mathbf{P}_2\}$ are consistent, the convex combination of them

$$\mathbf{P}^{-1}\hat{\mathbf{x}} = \omega \mathbf{P}_1^{-1}\hat{\mathbf{x}}_1 + (1-\omega)\mathbf{P}_2^{-1}\hat{\mathbf{x}}_2, \quad (29)$$

$$\mathbf{P}^{-1} = \omega \mathbf{P}_1^{-1} + (1-\omega)\mathbf{P}_2^{-1}, \quad (30)$$

where $\omega \in [0,1]$, leads to consistent estimate $\{\hat{\mathbf{x}}, \mathbf{P}\}$ for arbitrary cross-covariance $\mathbf{P}_{12} = \mathrm{E}[(\mathbf{x} - \hat{\mathbf{x}}_1)(\mathbf{x} - \hat{\mathbf{x}}_2)^{\mathrm{T}}]$, i.e. for arbitrary common information.

The weight ω can be chosen in order to minimise various criteria. The usual criterion is the determinant of the fused error covariance matrix,

$$\omega^* = \arg\min_{\omega \in [0,1]} (\det \mathbf{P}), \quad (31)$$

but the trace $\mathrm{tr}(\mathbf{P})$ is also used. The optimal weight ω^* can be approximated by the use of fast algorithms, see (Fränken and Hüpper, 2005). Special covariance consistency methods can be found in (Uhlmann, 2003).

(Hurley, 2002) generalise the Covariance Intersection method to the combination of probability density functions. The geometrical combination

$$p_\omega(\mathbf{x}) = \frac{p_1^\omega(\mathbf{x})p_2^{1-\omega}(\mathbf{x})}{\int_R p_1^\omega(\mathbf{x})p_2^{1-\omega}(\mathbf{x})\mathrm{d}\mathbf{x}} \qquad (32)$$

is used and the criterion of entropy, i.e. the Shannon information, of the fused density

$$\mathcal{H}(p_\omega) = -\int_R p_\omega(\mathbf{x})\ln p_\omega(\mathbf{x})\mathrm{d}\mathbf{x}, \qquad (33)$$

that corresponds to the determinant criterion of the fused estimate of Gaussian density, can be applied. Other proposed criterion is the Chernoff information $\mathcal{C}(p_1, p_2) = -\min_{0\leq\omega\leq 1}\left(\ln\int_R p_1^\omega p_2^{1-\omega}(\mathbf{x})\mathrm{d}\mathbf{x}\right)$. The optimal density is equally distant from the local densities in the Kullback-Leibler divergence sense, $\mathcal{D}(p_{\omega^*} \parallel p_1) = \mathcal{D}(p_{\omega^*} \parallel p_2)$, where the Kullback-Leibler divergence is defined as $\mathcal{D}(p_1 \parallel p_2) = \int_R p_1(\mathbf{x})\ln\left(\frac{p_1(\mathbf{x})}{p_2(\mathbf{x})}\right)\mathrm{d}\mathbf{x}$. (Julier, 2006) studies the Chernoff fusion approximation for Gausian-mixture models, (Farrell and Ganesh, 2009) and (Wang and Li, 2009) consider fast convex combination methods.

5 NUMERICAL ILLUSTRATION

In this section, the fusion approaches will be illustrated by a numerical example. Let the system (3), (4) with three sensors be t-invariant and given by

$$\mathbf{F} = \mathbf{I}_2, \quad \mathbf{G} = \mathbf{I}_2, \quad \mathbf{Q} = \begin{bmatrix} 1.44 & -1.2 \\ -1.2 & 1 \end{bmatrix}, \qquad (34)$$

$$\left\{ \begin{array}{lllll} \mathbf{H}^{(1)} & = & \begin{bmatrix} 1 & 0 \end{bmatrix}, & \mathbf{R}^{(11)} & = & 1, \\ \mathbf{H}^{(2)} & = & \begin{bmatrix} 1 & -1 \end{bmatrix}, & \mathbf{R}^{(22)} & = & 2, \\ \mathbf{H}^{(3)} & = & \begin{bmatrix} 0 & 1 \end{bmatrix}, & \mathbf{R}^{(33)} & = & 1, \end{array} \right. \qquad (35)$$

where the measurement errors are independent, $\mathbf{R}^{(12)} = \mathbf{R}^{(13)} = \mathbf{R}^{(23)} = 0$, and the initial condition is given by $p(\mathbf{x}_0) = \mathcal{N}([0,0]^\mathrm{T}, \mathbf{I}_2)$.

The used hierarchical and decentralised networks are shown on the Fig. 1, the numbers denote the respective estimators. The data links $1 \leftrightarrow 2$ and $2 \leftrightarrow 3$ are considered in the decentralised network.

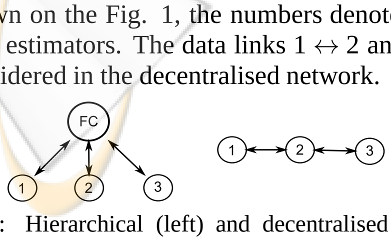

Figure 1: Hierarchical (left) and decentralised network (right), FC = fusion centre.

The centralised fusion (11), will be compared with the maximum likelihood (21), (22), (23), distributed Kalman filter (13)-(15), channel filters (25), (27) and the information measure approaches (29)-(31). The 1-σ bounds, i.e. the multidimensional parallels of the standard deviation, will show the uncertainty of the fused estimates. The bounds will be centred to zero to allow a better graphical comparison and are given by $\{\mathbf{x} : \mathbf{x}^\mathrm{T}\mathbf{P}^{-1}\mathbf{x} = 1\}$, where $\mathbf{P}$ is the estimate covariance and the $\mathbf{x} = [x_1, x_2]^\mathrm{T}$.

All estimators, including the fusion centre of the distributed Kalman filter and the channel filters, have the same initial condition $p(\mathbf{x}_0)$. The system is simulated and the 1-σ bounds at the times $k = 1$, $k = 5$, and $k = 20$ are shown in the Fig. 2 for the hierarchical and decentralised estimators.

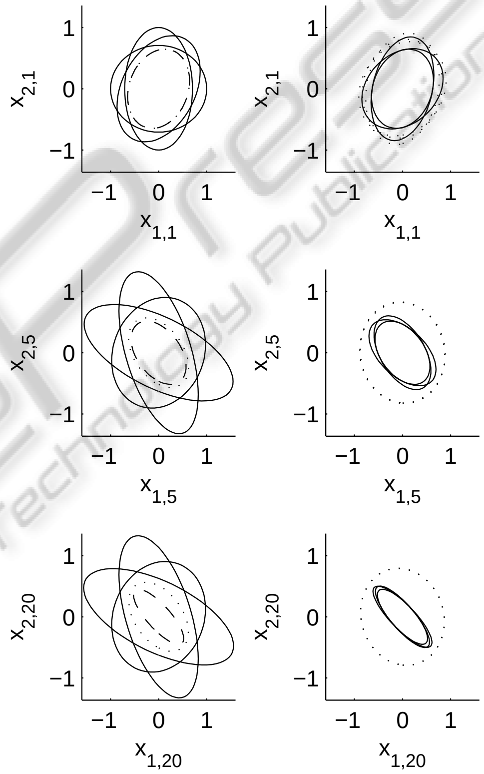

Figure 2: A comparison of the 1-σ bounds of the hierarchical and decentralised estimates at times $k = 1, 5, 20$.

The left half of the Fig. 2 shows the optimal centralised estimator (dashed line), the distributed Kalman filter (with the same estimate - dashed line), local Kalman filters (solid lines), and the fusion by the maximum likelihood at the fusion centre (dotted line). At the time $k = 1$ (top), the maximum likelihood estimate and the centralised estimate have equal covariances, the lines seem to be dash-dotted. At the times $k = 5$ and $k = 20$ (middle and bottom), the influence of not incorporating the prior information is evident, the covariance of the maximum likelihood estimate is greater than that of the centralised estimate. The local filters are the least accurate.

The right half of the Fig. 2 shows the local estimates with the channel filter fusion (solid lines) and the Covariance Intersection fusion (dotted lines). The estimate of the estimator 2 with the channel filter fusion is equal to the centralised estimate in this case. The one-step delay of the measurement exploitation in the estimators 1 (which measures x_1) and 3 (which measures x_2) is visible, there is greater uncertainty in the x_2 and x_1 axis, respectively. The the local estimates which use the Covariance Intersection get close to each other after a few steps. In this example, the estimates 2 and 1 are fused first and the result is fused with the estimate 3. The estimates overestimate the error covariance, but at least they are not worse than the estimates that use local measurements only without any fusion (compare with the solid lines on the left half of the figure). The information measure approaches are useful for more complex networks.

6 SUMMARY

Main approaches to the state estimate fusion for the linear stochastic systems were introduced. The principles and algorithms of hierarchical and decentralised fusion were presented and discussed. Contrary to the standard estimation problem, which is based on using all measurements simultaneously, the estimate fusion allows to respect an alternative technical specification concerning the measurement location and to prefer local information processing. The hierarchical fusion is more suitable for systems with a small number of sensors. In the case of general network with many sensors, the decentralised fusion based on information measures should be preferred due to its simplicity and modest assumptions.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic, project no. 1M0572, and by the Czech Science Foundation, project no. 102/08/0442.

REFERENCES

Bar-Shalom, Y. (2002). Update with out-of-sequence measurements in tracking: Exact solution. *IEEE Transactions on AES*, 38(3):760–778.

Berg, T. M. and Durrant-Whyte, H. F. (1992). Distributed and decentralized estimation. In *SICICI '92 proceedings*, volume 2, pages 1118–1123.

Challa, S., Evans, R. J., and Wang, X. (2003). A bayesian solution and its approximations to out-of-sequence measurement problems. *Information Fusion*, 4(3):185–199.

Chang, K.-C., Sahat, R. K., and Bar-Shalom, Y. (1997). On optimal track-to-track fusion. *IEEE Transactions on AES*, 33(4):1271–1276.

Chong, C.-Y., Mori, S., Chang, K.-C., and H., B. W. (1999). Architectures and algorithms for track association and fusion. In *Proceedings of the 2nd International Conf. on Information Fusion*.

Farrell, W. J. and Ganesh, C. (2009). Generalized chernoff fusion approximation for practical distributed data fusion. In *Proceedings of the 12th International Conf. on Information Fusion*.

Fränken, D. and Hüpper, A. (2005). Improved fast covariance intersection for distributed data fusion. In *Proc. of the 8th Int. Conf. on Inf. Fusion*.

Grime, S. and Durrant-Whyte, H. F. (1994). Data fusion in decentralized sensor networks. *Control Engineering Practice*, 3(5):849–863.

Hashemipour, H. R., Roy, S., and Laub, A. J. (1988). Decentralized structures for parallel kalman filtering. *IEEE Transactions on AuC*, 33(1):88–94.

Hurley, M. B. (2002). An information theoretic justification for covariance intersection and its generalization. In *Proceedings of the 5th International Conference on Information Fusion*.

Julier, S. J. (2006). An empirical study into the use of chernoff information for robust, distributed fusion of gaussian mixture models. In *Proceedings of the 9th Int. Conf. on Information Fusion*.

Julier, S. J. (2009). Estimating and exploiting the degree of independent information in distributed data fusion. In *Proceedings of the 12th International Conference on Information Fusion*.

Li, X.-R., Zhu, Y., Wang, J., and Han, C. (2003). Optimal linear estimation fusion—part i: Unified fusion rules. *IEEE Transactions on Information Theory*, 49(9):2192–2208.

Simon, D. (2006). *Optimal State Estimation*. Wiley.

Uhlmann, J. K. (2003). Covariance consistency methods for fault-tolerant distributed data fusion. *Information Fusion*, 4(3):201–215.

Wang, Y. and Li, X.-R. (2009). A fast and fault-tolerant convex combination fusion algorithm under unknown cross-correlation. In *Proceedings of the 12th Int. Conf. on Information Fusion*.