Fusion of Time Delayed Measurements With Uncertain Time Delays

Simon J. Julier and Jeffrey K. Uhlmann

***Abstract*— In this paper we consider the problem of estimating the state of a dynamic system from a sequence of observations that are imprecisely timestamped. We argue that this problem can be addressed using the Covariance Union (CU) technique, and we demonstrate its application in a particular example.**

I. INTRODUCTION

Many applications require the maintenance of a high-fidelity estimate of the state of a dynamic system based on a sequence of noisy observations. Such applications demand the use of a filtering mechanism, such as the Kalman filter, to fit the observation sequence to a given model of the system dynamics. One assumption that is frequently made is that the measurement is propagated from the sensing device to the filter occurs without delay. However, in practice time delays can occur between when an observation is taken by a sensor and when it becomes available to the filtering algorithm leading to fusion with time delayed measurements. When measurements with different time delays are interleaved with one another, this is known as the Out-of-Sequence Measurement problem (OOSM) [1].

When the time delays are known, the Kalman filter can be readily extended to account for the delays. Probably the simplest approach is to buffer all incoming measurements and run the filter behind real-time, the length of the delay sufficient to guarantee that few measurements appear out of order. Nettleton derived an elegant implementation of this scheme using the inverse covariance (or information) form of the Kalman filter [2] when all the delayed measurements were taken at the same time. However, because the filter has to run with the latency equal to the maximum latency of any given sensor, the filter might not be able to run in real time. To overcome this problem, Larsen developed a scheme for extrapolating a measurement forward through time [3]. Using this method, a time delayed measurement could be incorporated into the filter at a later time. Challa performed a detailed Bayesian analysis of the problem of fused measurements and proved that the filtering problem requires the consideration of the joint density of the current target state and the target state corresponding to the delayed measurement [1]. Van der Merwe et al. used this approach to fuse delayed GPS measurements with inertial measurements for the navigation of a UAV [4]. The equations derived by this approach are the same as those derived by Larsen showing that Larsen's approach and Challa's approach are identical. Zhang has recently developed a number of sub-optimal algorithms that attempt to reduce the computational and storage costs by approximating the cross correlations [5]

Time delays can also be used to fuse measurements that are taken over time. Both Leonard [6] and Fitzgibbons [7] calculate the positions of targets in three dimensions using bearings only measurements. When a measurement of a target is made, the position of the sensor is recorded. From a history of measurements and sensor positions, the position of the target can be reconstructed.

However, all of the methods assume that, although the measurements are delayed (by a potentially random amount), the amount of delay is known. However, situations can arise in which the time delay is not known perfectly. Possible reasons include:

1) Observations are time stamped when they become available to the filter, not when the observation was taken. One example of this is image acquisition using consumer-grade hardware such as a webcamera. Each image acquired by the camera is not time stamped when the image was taken. Rather, the image is time stamped when the data becomes available to an image processing application. Because of non deterministic OS delays (due to buffering, context switches and the like), the latency between acquiring an image and making it available is known imperfectly.
2) Observations are time stamped from a local clock without centralized clock synchronization. In this case, there will always be deviations between clocks at different nodes and, as a result, latency cannot be calculated by subtracting the observation time from the current filter time.

Even though the latency is unknown, it might be possible to statistically characterize it through experimentation[1]. In this paper we assume that the latency is quantized and is bounded by known values. As a result, there are a finite number of possible values of latency for the sensor.

Aron [9] considered the problem of using an inertial system to assist a computer-vision based tracking system when the camera tracker became lost. The inertial system contained unmodelled latency and so a RANSAC algorithm was applied to fuse the inertial and camera data to find the

Simon J. Julier is with ITT AES / Naval Research Laboratory, 4555 Overlook Avenue SW, Washington, DC 20375, USA. Email: simon.julier@nrl.navy.mil

Jeffrey K. Uhlmann is with the Department of Computer Science, University of Missouri–Columbia, Columbia, MO 65211 USA. Email: uhlmannj@missouri.edu

[1]Liang [8], for example, measured the latency in an orientation system by affi xing a sensor to a pendulum.

"best fit" when RANSAC gives best results. This resembles the least squares errors approaches discussed in the survey paper by Fernandes [10]. However, the difficulty of these approaches is that, if the time step is incorrectly identified, catastrophic failures can result.

A more principled way to handle this problem is to extend Challa's Bayesian formalism [1] to include the uncertainties in the time delays. This is analogous to a problem that arises in multiple target tracking (MTT) [11]. MTT occurs when a tracking system receives an observation of one of several different targets, but the exact identity of the observed target is not known. In the delayed fusion problem, the tracking system receives an observation which can arise from one of several different time steps. However, many MTT algorithms are either fragile or have computational costs that scale exponentially and cannot be implemented in real time.

In this paper, we consider the problem of developing a mechanism for accommodating timestep uncertainty directly into the observation covariance so that filter consistency is always maintained using the Covariance Union (CU) algorithm. CU derives consistent estimates when the association between the observation and the target state is unknown [12]. Given a set of plausible assocations, CU calculates the minimum estimate that is guaranteed to be consistent with respect to all of those assocations.

The structure of this paper is as follows. Section II describes the problem statement. Section III describes the approach using the CU algorithm. An example is presented in Section IV and conclusions are given in Section V.

II. Problem Statement

A. System Description

Consider the system described by the discrete-time linear equation

$$\mathbf{x}(k) = \mathbf{F}(k-1)\,\mathbf{x}(k-1) + \mathbf{v}(k-1)$$

where $\mathbf{x}(k)$ is the state vector at time k, $\mathbf{v}(k-1)$ is the process noise, and $\mathbf{F}(k-1)$ is the state transition matrix.

The observation model is

$$\mathbf{z}(k) = \mathbf{H}(k)\,\mathbf{x}(k) + \mathbf{w}(k) \tag{1}$$

where $\mathbf{z}(k)$ is the observation vector, $\mathbf{w}(k)$ is the observation noise vector and $\mathbf{H}(k)$ is the observation matrix. The noise vectors $\mathbf{v}(k-1)$ and $\mathbf{w}(k)$ are assumed to be zero-mean and uncorrelated with covariances $\mathbf{Q}(k)$ and $\mathbf{R}(k)$ respectively.

Let the estimate of $\mathbf{x}(i)$ using observations up to time step j be $\hat{\mathbf{x}}(i \mid j)$ with covariance $\mathbf{P}(i \mid j)$. Defining the estimation error to be

$$\tilde{\mathbf{x}}(i \mid j) = \mathbf{x}(i) - \hat{\mathbf{x}}(i \mid j), \tag{2}$$

the estimate is said to be *consistent* if

$$\mathbf{P}(i \mid j) - \mathrm{E}\left[\tilde{\mathbf{x}}(i \mid j)\,\tilde{\mathbf{x}}^T(i \mid j)\right] \geq \mathbf{0}. \tag{3}$$

It is normally assumed that the measurement is made available at the time the measurement is made. However, this is not always the case. As explained in the introduction, latencies in sensing and transmission mean that observations can be delayed by a number of time steps before they are received and processed by the filter.

B. Fusion of Time Delayed Measurements

Suppose that a given measurement may be delayed by n timesteps. Defining $k_n = k - n$, the update equation is

$$\hat{\mathbf{x}}(k \mid k) = \hat{\mathbf{x}}(k \mid k-1) + \mathbf{W}(k)\,\nu(k_n) \tag{4}$$

where

$$\nu(k_n) = \mathbf{z}(k_n) - \hat{\mathbf{z}}(k_n \mid k_{n-1})$$

is the innovation calculated from the observation *at the time the observation was taken*. The estimation error committed by (4) is

$$\tilde{\mathbf{x}}(k \mid k) = \tilde{\mathbf{x}}(k \mid k-1) - \mathbf{W}(k)\,\nu(k_n).$$

Taking outer products and expectations,

$$\begin{aligned}\mathbf{P}(k \mid k) = \mathbf{P}(k \mid k-1) &- \mathbf{X}(k)\,\mathbf{W}^T(k) \\ &- \mathbf{W}(k)\,\mathbf{X}^T(k) \\ &+ \mathbf{W}(k)\,\mathbf{P}_{\nu\nu}(k_n \mid k_{n-1})\,\mathbf{W}^T(k)\end{aligned} \tag{5}$$

where

$$\mathbf{X}(k) = \mathrm{E}\left[\tilde{\mathbf{x}}(k \mid k-1)\,\nu^T(k_n)\right]$$

is the cross correlation between the prediction at the current time and the innovation at the time the observation was taken. Substituting from (1) and (2) and using the fact that $\mathbf{w}(k)$ is independent,

$$\mathbf{X}(k) = \mathrm{E}\left[\tilde{\mathbf{x}}(k \mid k-1)\,\tilde{\mathbf{x}}(k_n \mid k_n - 1)\right]\mathbf{H}^T. \tag{6}$$

Taking derivatives, it can be shown that the value of $\mathbf{W}(k)$ which minimises the trace of $\mathbf{P}(k \mid k)$ is

$$\mathbf{W}(k) = \mathbf{X}(k)\,\mathbf{P}_{\nu\nu}^{-1}(k_n \mid k_n - 1). \tag{7}$$

This equation provides a simple, intuitive interpretation of the weight in the time delayed Kalman filter. The weight assigned to a measurement is a function of the degree to which the measurement is correlated with the current state of the system[2].

Therefore, the difficulty in implementing the time-delayed Kalman filter is in calculating $\mathbf{X}(k)$. In this paper we exploit the linearity assumptions to contruct a linear closed form solution [3].

By combining the prediction and update steps together, the prediction from one time step to the next can be written as

$$\begin{aligned}\hat{\mathbf{x}}(k+1 \mid k) = \mathbf{\Phi}(k+1, k)\,\hat{\mathbf{x}}(k \mid k-1) &+ \mathbf{F}(k)\,\mathbf{w}(k) \\ &+ \mathbf{v}(k)\end{aligned}$$

[2]The result here is algebraically the same as derived by Larsen [3]. However, the interpretation is very different. Larsen considered taking an observation and extrapolating its value *forwards* to the current time step in the filter. We consider calculating the correlation *backwards* from the current time to the time when the observation was made.

where

$$\boldsymbol{\Phi}(k+1,k)=\begin{cases}\Big(\mathbf{I}-\mathbf{W}(k)\,\mathbf{H}(k)\Big)\,\mathbf{F}(k) & \text{observation}\\ \mathbf{F}(k) & \text{no observation}\end{cases}$$

combines the effect of the update at time k together with the prediction at time $k+1$. Noting that $\mathbf{v}(\cdot)$ and $\mathbf{w}(\cdot)$ are independent,

$$\begin{aligned}&\mathrm{E}\left[\tilde{\mathbf{x}}(k \mid k-1)\,\tilde{\mathbf{x}}(k_n \mid k_n-1)\right]\\ &\quad \mathrm{E}\left[\boldsymbol{\Phi}(k,k_n)\,\tilde{\mathbf{x}}(k_n \mid k_n-1)\,\tilde{\mathbf{x}}^T(k_n \mid k_n-1)\right]\\ &\quad =\boldsymbol{\Phi}(k,k_n)\,\mathbf{P}(k_n \mid k_n-1)\end{aligned} \tag{8}$$

where

$$\boldsymbol{\Phi}(k,k_n)=\prod_{i=k_n}^{k-1}\boldsymbol{\Phi}(i+1,i) \tag{9}$$

These results assume that k_n is known precisely. However, as explained in the introduction, various practical limitations mean that n is not always known perfectly.

C. The Effect of Unknown Time Delays

Suppose the true timestep is k_N whereas the assumed time step is k_n.

Now,

$$\mathbf{x}(k_N)=\boldsymbol{\Psi}(k_N,k_n)\,\mathbf{x}(k_n)+\mathbf{v}(k_N,k_n)$$

where

$$\boldsymbol{\Psi}(k_N,k_n)=\prod_{i=k_n}^{k_N}\mathbf{F}(i)\,.$$

The innovation is

$$\begin{aligned}\nu(k_n)&=\mathbf{z}(k_N)-\hat{\mathbf{z}}(k_n \mid k_n-1)\\ &=\mathbf{H}\mathbf{x}(k_N)+\mathbf{w}(k_N)\\ &\quad-\mathbf{H}\hat{\mathbf{x}}(k_n \mid k_n-1)\\ &=\mathbf{H}\boldsymbol{\Psi}(k_N,k_n)\,\mathbf{x}(k_n)+\mathbf{H}\mathbf{v}(k_N,k_n)\\ &\quad+\mathbf{w}(k_N)-\mathbf{H}\hat{\mathbf{x}}(k_n \mid k_n-1)\\ &=\mathbf{H}\boldsymbol{\Psi}(k_N,k_n)\,\tilde{\mathbf{x}}(k_n \mid k_n-1)\\ &\quad+\mathbf{H}\boldsymbol{\Psi}(k_N,k_n)\,\hat{\mathbf{x}}(k_n \mid k_n-1)+\mathbf{H}\mathbf{v}(k_N,k_n)\\ &\quad+\mathbf{w}(k_N)-\mathbf{H}\hat{\mathbf{x}}(k_n \mid k_n-1)\\ &=\mathbf{H}\boldsymbol{\Psi}(k_N,k_n)\,\tilde{\mathbf{x}}(k_n \mid k_n-1)\\ &\quad+\mathbf{H}\left(\boldsymbol{\Psi}(k_N,k_n)-\mathbf{I}\right)\hat{\mathbf{x}}(k_n \mid k_n-1)\\ &\quad+\mathbf{v}(k_N,k_n)+\mathbf{w}(k_N)\end{aligned}$$

Therefore, an error in the assumed time step can be treated as an error in the observation model which includes a mis-specification of the observation matrix and an observation sequence which contains correlated noise.

This structure bares a strong relationship to that which arises with Multiple Target Tracking (MTT). One of the most widely used methods for dealing with MTT is Multiple Hypotheses Tracking (MHT) [11]. MHT creates a distinct updated state estimate for each element of the set of possibilities under the assumption that subsequent observations will tend to be consistent only with the valid updates, thus allowing the others to be discarded (pruned). However, MHT can result in an exponential proliferation of hypotheses and cannot be implemented in real time. Rather, various adhoc pruning strategies must be used to limit the computational growth. The simplest approach, the Probabilistic Data Association Filter (PDAF), assumes that there is a single target [11]. However, the fewer hypotheses an algorithm maintains, the more prone it is to catastrophic failure due to an incorrect identification of the timestep.

In this paper we consider the alternative approach of expanding the observation covariance so that it takes account of the uncertainty in the time stamp using the Covariance Union (CU) algorithm.

III. COVARIANCE UNION

Covariance Union (CU) considers the following problem: suppose a filtering algorithm is provided with two observations with means and covariances $(\mathbf{m}_1,\mathbf{M}_1)$ and $(\mathbf{m}_2,\mathbf{M}_2)$ respectively. It is known that one observation corresponds to a correct association, and the other to an incorrect association. However, the identity of the consistent estimate is unknown and cannot be determined. In this circumstance, the only way the KF can be guaranteed to give a consistent estimate is if it updates with an observation which is consistent with respect to *both* measurements [12][3]. This unioned estimate has a mean and covariance $(\mathbf{u},\mathbf{U})$ and obeys the property

$$\mathbf{U} \;\geq\; \mathbf{M}_1+(\mathbf{u}-\mathbf{m}_1)(\mathbf{u}-\mathbf{m}_1)^T \tag{10}$$

$$\mathbf{U} \;\geq\; \mathbf{M}_2+(\mathbf{u}-\mathbf{m}_2)(\mathbf{u}-\mathbf{m}_2)^T \tag{11}$$

where some measure of the size of $\mathbf{U}$ (e.g., determinant) is minimized. Given this Covariance Union (CU) of the two measurements, the KF can be applied directly to update the prediction $(\hat{\mathbf{x}}(k \mid k-1),\mathbf{P}(k \mid k-1))$ with the CU-derived observation $(\mathbf{u},\mathbf{U})$. In other words, the above equations simply say that if the estimate $(\mathbf{m}_1,\mathbf{M}_1)$ is consistent, then the translation of the vector $\mathbf{m}_1$ to $\mathbf{u}$ will require its covariance to be enlarged by the addition of a matrix at least as large as the outer product of $(\mathbf{u}-\mathbf{m}_1)$ in order to be consistent. The same reasoning applies if the estimate $(\mathbf{m}_2,\mathbf{M}_2)$, is consistent.

One simple strategy would be to choose the unioned estimate to be set the unioned mean to be one of the input estimates, $\mathbf{u}=\mathbf{m}_1$. In this case, $\mathbf{U}$ must be chosen so that the estimate is consistent with the worst case error — namely that the real update should be $\mathbf{m}_2$. However if $\mathbf{u}$ can be placed somewhere between $\mathbf{m}_1$ and $\mathbf{m}_2$ then the worst case error is reduced and a smaller value of $\mathbf{U}$ is needed.

Because the inequality is convex, many algorithms for convex or semidefinite optimization methods can be used.

[3]This is related to a result by Nishimura [13]. Nishimura considered the problem of modelling errors and showed that even if the process and observation models contain errors, a consistent estimate can be achieved by increasing the process and observation noise by a suffi cient amount.

In this paper we assume that $\mathbf{u}$ can be formed from a convex combination of $\mathbf{m}_1$ and $\mathbf{m}_2$[4]:

$$\mathbf{u} = \omega \mathbf{m}_1 + (1-\omega)\mathbf{m}_2$$

where $0 \leq \omega \leq 1$. The appendix summarises an iterative method for calculating $\mathbf{U}$ [14].

This result can be applied directly to the case of filtering measurements with unknown time stamps. Specifically, given that a set of time steps can be identified $[k_{max}, k_{min}]$, the updates can be calculated for k_{max} and k_{min} and CU can be applied to merge the two sets of updates together.

IV. EXAMPLE

Consider the problem of estimating the position and velocity of a 1D particle. The state of the system is

$$\mathbf{x}(k) = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}.$$

The particle moves with piecewise constant velocity:

$$\mathbf{F}(k) = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}.$$

The process noise acts on the acceleration. Therefore,

$$\mathbf{Q}(k) = \begin{bmatrix} \Delta t^3/3 & \Delta t^2/2 \\ \Delta t^2/2 & \Delta t \end{bmatrix} q$$

where $q = 10^{-3}$ is the acceleration covariance.

The position is measured by the sensor,

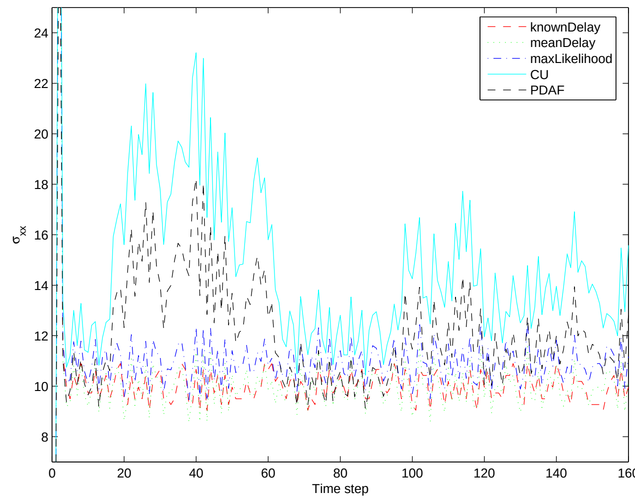
$$\mathbf{H}(k) = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

The observation noise is additive with covariance $\mathbf{R}(k) = 10^{-2}$.

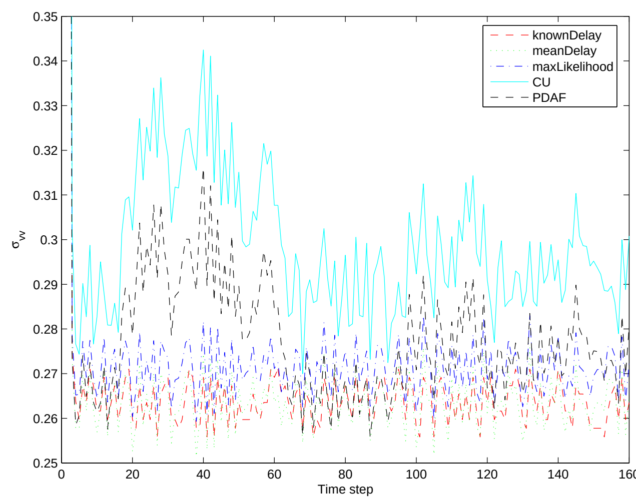
The measurement is taken at a rate of once every 50 timesteps. However, there is a random delay, uniformly distributed between 2 and 10 timesteps, before the measurement can be timestamped and made available to the filter.

Five algorithms were tested[5]:

1) **knownDelay.** This filter is implemented with the delay known perfectly. It is a test of the delayed filter and provides a means of quantifying the effect of unknown time delays.
2) **meanDelay.** This filter assumes the time delay is the average value, 6.
3) **maxLikelihood.** For each time step in the range [2,10] calculate the predicted value and the likelihood of the estimate using the Gaussian likelihood model. Select the time step with the highest likelihood.
4) **CU.** Calculate the estimates k_2 and k_{10} and fuse them together using CU. ω was chosen to minimise the determinant of $\mathbf{P}(k \mid k)$.
5) **PDAF.** Like maxLikelihood, calculate the likelihood of each time step. Take the weighted average of the estimates from the different time steps, weighted by the likelihood.



(a) Covariance history of x.



(b) Covariance history of $\dot{x}$.

Fig. 1. Covariance histories for the filter estimates.

Fig. 1 plots the covariance histories of one complete run for all of the filters. We only plot the estimate immediately following the time step (time step in plot = 50 * filter time step, hence 8000 time steps has 160 updates). These plots suggest that knownDelay, meanDelay and maxLikelihood filters calculate similar covariances whereas the CU and PDAF calculate estimates with significantly larger covari-

[4] This form of $\mathbf{u}$ was chosen for computational convenience and it is possible that other forms can lead to a significantly smaller value of $\mathbf{U}$.

[5] We do not examine MHT in our tests because its performance is almost entirely dependent on the amount of extra memory that is permitted and the strategy that is used for pruning hypotheses. All that can be said for sure is that any amount of pruning will make an MHT filter susceptible to producing inconsistent estimates in the same way as our maximum likelihood filter.

Scheme	σ_{xx}	$\sigma_{\dot{x}\dot{x}}$
knownDelay	5.1176	0.2131
meanDelay	5.1148	0.2131
maxLikelihood	5.8736	0.2226
CU	9.2287	0.2567
PDAF	7.0729	0.2294

TABLE I

THE AVERAGE STANDARD DEVIATIONS OF x AND $\dot{x}$ FOR EACH FILTER.

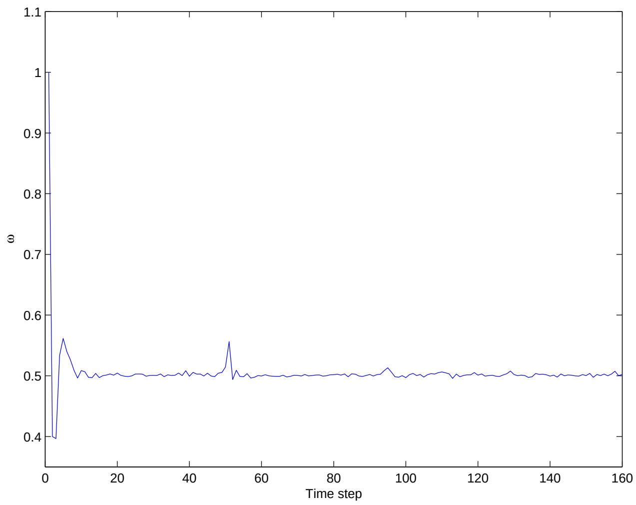


Fig. 2. Time history of ω for one run of the CU filter.

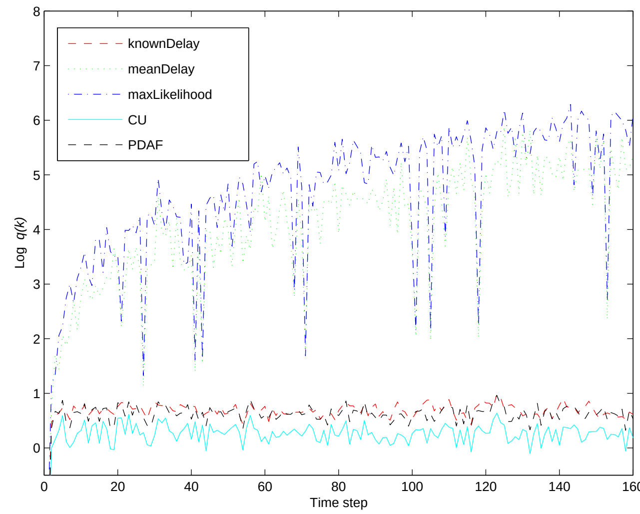


Fig. 3. The time history of $\log[q(k)]$.

Scheme	$\mathrm{E}\left[q(k)\right]$
knownDelay	1.9992
meanDelay	37.6949
maxLikelihood	54.9323
CU	1.3172
PDAF	1.8749

TABLE II

THE AVERAGE NORMALISED MEAN SQUARED ERROR FOR EACH FILTER.

ances. This is quantified in Table I which lists the average standard deviation in each filter. Fig. 2 plots the history of ω for one run of the CU filter. After some initial transients, the filter rapidly settles to a value of $\omega \approx 0.5$. This solution appears to resemble that from the meanDelay filter, but expanding the covariance to cover the worst case errors. Furthermore, the CU filter is different from using the PDAF and setting the probabilities of k_2 and k_{10} to be 0.5. The reason is that the PDAF would take the numerical average of the covariances whereas the CU filter takes the minimum bounding covariance.

However the covariance histories, by themselves, do not prove whether the filter is consistent. Rather, we perform a statistical test on the *normalised state error* of the estimate [11].

Let

$$q(k) = \left(\hat{\mathbf{x}}\left(k \mid k\right) - \mathbf{x}\left(k\right)\right) \mathbf{P}^{-1}\left(k \mid k\right) \\ \times \left(\hat{\mathbf{x}}\left(k \mid k\right) - \mathbf{x}\left(k\right)\right)^{T}.$$

This quantity is a χ^2 distribution with 2 degrees of freedom. Therefore, if the true mean squared error in the estimate is $\mathbf{P}\left(k \mid k\right)$, then $\mathrm{E}\left[q(k)\right] = 2$. If the expected value exceeds this then the filter is not consistent — the filter estimated covariance is smaller than the mean squared error. If the expected value is smaller than this then the filter overestimates the covariance.

Fig. 3 plots the logarithm of the normalised state error for all the filters calculated across all the runs[6] and Table II presents the average normalised results. The results clearly show that the smaller covariances on the averageDelay and maxLikelihood filters are misleading — both yield highly inconsistent results. The result for the knownDelay filter validates the expression for time delayed measurements. The PDAF filter operates consistently and suggests that the effect of the uncertainty in the time stamp is to increase the average standard deviation of x by approximately 40%. The CU estimate is also consistent, but has a larger covariance.

These results clearly indicate that the uncertainty in the time stamp has a significant impact on estimator performance and, in this case, the PDAF gives a smaller covariance. However, two caveats must be noted. First, the PDAF has a significantly higher computational cost in terms of the number of filter evaluations which must be made. Kalman filter updates (and the innovation likelihoods) must be calculated for all the timesteps within which the delay is feasible. In this example, eight updates are required. However, it is clear that the computational load increases as the window of the potential time delay increases. On the other hand, the CU filter only needs to be calculated twice. Second, it is important to note that this example is a simple, linear time invariant system and, as such, the likelihood

[6]The log-normalised plot is necessary because of the large magnitude errors of the maximum likelihood estimator.

assumptions are correct. However, real systems are highly nonlinear and no such guarantees hold true. For example, low quality inertial systems contain errors (due to compass biases) that are highly correlated. As a result, the likelihood calculation will be extremely crude.

V. CONCLUSIONS

This paper has considered the problem of applying the Kalman filter to observation sequences in which the timestamp for each observation is not precisely known. We have argued that this problem has analogies with the identity ambiguity problem that arises in multi-target tracking applications. Such problems can be addressed with multiple hypothesis and Covariance Union (CU) methods. We have described a way in which CU can be applied to the problem and we have compared the performance of the algorithm against a range of approaches. The results show that PDAF yields the most accurate performance. However, it requires significant computational costs and is strongly dependent on the accuracy of the likelihood model. CU, on the other hand, only requires the evaluation of two Kalman filter updates and does not rely on specific assumptions as to the veracity of the likelihood model.

Future work will examine the use of a hybrid of CU and MHT in which we will assume that the true MHT hypothesis is maintained so that the only remaining problem is resolving the ambiguity of which hypothesis it is. The application of CU to the set of maintained hypotheses will yield a consistent result that may be superior to the use of CU alone.

APPENDIX

This derivation was first described in [14].

Given the prior observations $(\mathbf{m}_1, \mathbf{M}_1)$ and $(\mathbf{m}_2, \mathbf{M}_2)$ and a candidate mean $\mathbf{u}$, this appendix describes a method for calculating $\mathbf{U}$ which is guaranteed to be consistent. We assume that the estimate $\mathbf{u}$ is given by the convex combination of the two prior estimates,

$$\mathbf{u} = \omega \mathbf{m}_1 + (1-\omega)\mathbf{m}_2$$

Define

$$\begin{aligned}
\mathbf{U}_1 &= \mathbf{M}_1 + \{\mathbf{u} - \mathbf{m}_1\}\{\mathbf{u} - \mathbf{m}_1\}^T \\
&= \mathbf{M}_1 + (1-\omega)^2(\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T \\
\mathbf{U}_2 &= \mathbf{M}_2 + \{\mathbf{u} - \mathbf{m}_2\}\{\mathbf{u} - \mathbf{m}_2\}^T \\
&= \mathbf{M}_2 + \omega^2(\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T
\end{aligned}$$

Let

$$\mathbf{S} = \sqrt{[\mathbf{U}_2]}$$

where $\mathbf{U}_2 = \mathbf{S}^T\mathbf{S}$. Furthermore, let $\mathbf{V}$ and $\mathbf{D}$ be the matrices of eigenvectors and eigenvalues of $\left(\mathbf{S}^{-1}\right)^T \mathbf{U}_1 \mathbf{S}^{-1}$. A covariance matrix $\mathbf{U}$ which obeys conditions (10) and (11) is

$$\mathbf{U} = \mathbf{S}^T \mathbf{V} \max\left(\mathbf{D}, \mathbf{I}\right) \mathbf{V}^T \mathbf{S}$$

where $\max\left(\mathbf{A}, \mathbf{B}\right)$ calculates the matrix which is the element-wise maximum of the matrices $\mathbf{A}$ and $\mathbf{B}$.

REFERENCES

[1] S. Challa, R. J. Evans and X. Wang, "A bayesian solution and its approximations to out-of-sequence measurement problems," *Information Fusion*, vol. 4, pp. 185–199, 2003.

[2] E. W. Nettleton and H. F. Durrant-Whyte, "Delayed and asequent data in decentralised sensing networks," in *Sensor Fusion and Decentralized Control in Robotic Systems IV*, G. T. McKee and Paul S. Schenker, Ed., vol. 4571, SPIE. Newton, MA, USA: SPIE, Bellingham, WA, USA, 28 Oct–2 Nov 2001, pp. 1–9.

[3] T. D. Larsen, N. A. Andersen, O. Ravn and N. Kjolstad Poulson, "Incorporation of Time Delayed Measurements in a Discrete-time Kalman Filter," in *Proceedings of CDC'98: the 38th Conference on Decision and Control*, 1998.

[4] R. van der Merwe, E. Wan, and S. J. Julier, "Sigma-point kalmanfi lters for nonlinear estimation and sensor-fusion: Applications to integrated navigation," in *Proceedings of the AIAA Guidance, Navigation & Control Conference*, Providence, RI, USA, 16–19 August 2004.

[5] K. Zhang and X. R. Li, "Optimal update with out-of-sequence measurements for distributed fi ltering," in *Proceedings of the 5th International Conference on Information Fusion*, Annapolis, MD, USA, July 2002, pp. 1519–1526.

[6] J. J. Leonard and R. J. Rikoski, "Incorporation of delayed decision making into stochastic mapping," in *Experimental Robotics VII*, ser. Lecture Notes in Control and Information Sciences. Springer Verlag, 2001, vol. 271, pp. 533–542.

[7] T. Fitzgibbons and E. Nebot, "Application of vision in simultaneous localisation and mapping," in *Proceedings of the 2001 Australian Conference on Robotics and Automation*, Sydney, Australia, 14–15 November 2001, pp. 121–127.

[8] J. Liang, C. Shaw and M. Green, "On temporal-spatial realism in the virtual reality environment," in *Proceedings of the Fourth Annual Symposium on User Interface Software and Technology*, Hilton Head, SC, USA, 11–13 November 1991, pp. 19–25.

[9] M. Aron, G. Simon and M. Berger, "Handling uncertain sensor data in vision-based camera tracking," in *Proceedings of the Fourth Annual Symposium on Mixed and Augmented Reality*, Arlington, VA, USA, 2–5 November 2004, pp. 58–67.

[10] J. M. Fernandes and A. R. Ferreira, "A survey on time delay system estimation," in *Proceedings of ECC'97*, Brussels, Belgium, 1–4 July 1997.

[11] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. New York NY, USA: Academic Press, 1988.

[12] J. Uhlmann, "Covariance consistency methods for fault-tolerant distributed data fusion," *Information Fusion*, vol. 4, no. 3, pp. 201–215, Sept 2003.

[13] T. Nishimura, "On the a priori information in sequential estimation problems," *IEEE Transactions on Automatic Control*, vol. TAC-11, no. 4, pp. 197–204, April 1966.

[14] S. Julier, D. N. Nicholson and J. K. Uhlmann, "A method for dealing with assignment ambiguity," in *2004 American Control Conference*, Boston, MA, USA, 30 June – 2 July 2004.