Online Filtering For Hybrid Systems

Evgenia Suzdaleva[1], Ivan Nagy[2], and Martin Dungl[1]

[1] Department of Adaptive Systems, Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 18208 Prague, Czech Republic (E-mail: suzdalev@utia.cas.cz, dungl@utia.cas.cz)

[2] Faculty of Transportation Sciences, Czech Technical University, Na Florenci 25, 11000 Prague, Czech Republic (E-mail: nagy@utia.cas.cz)

Abstract. The paper deals with online state estimation for dynamic hybrid systems with mixed continuous and discrete states. The proposed solution is based on a decomposed version of the state-space model and Bayesian filtering. Specialization to Gaussian linear dynamic and multinomial state-space models is described. Experiments with real data illustrating the presented approach are provided.

Keywords: State estimation, online filtering, hybrid systems.

1 Introduction

Dynamic systems whose behavior changes not only continuously in time, but also within some discrete values (i.e., system modes) are met in many fields (for example, target tracking, image processing, speech recognition, traffic control, etc.) Modeling and especially adaptive control of such hybrid systems is a difficult task. Fast online state estimators for hybrid systems are desired in some of these areas.

Many algorithms exist for state estimation of such systems. The well-known approach is the interactive multiple model (IMM) algorithm [1], which performs Kalman filter [2] for each model and then computes a weighted combination of updated state estimates produced by all the filters. An exact filter for a specialized hybrid system is proposed in [3], where discrete state is treated via hidden Markov models (HMM) and a solution is presented as Gaussian sum with explicitly computed specific weights, means and variances. However, a number of statistics grows geometrically in time, which restricts duration of online estimation by number of time steps. Other special cases of dynamic switching models are presented in [4,5]. Iterative techniques nicely presented in [6] can be found in literature as well as the mixture Kalman filter [7] based on sequential Monte Carlo methods. These approaches are closely related to that described at this paper. A part of the proposed work concerned with the estimation of discrete multinomial state is also close to HMM theory [8]. However, the algorithms mentioned run mostly offline and are supported by Monte Carlo computations. The presented paper aims at online state estimation and analytical solution as far as possible.

It means that it applies numerical procedures only in that parts, which can not be computed analytically. The paper exploits a decomposition of state estimate, which enables to consider state as a product of various (here specialized) distributions. An online filter for discrete multinomial state based on fully analytical solution is proposed. The state-space model is taken as the probability (density) function in more general (not reduced) form, including control variables for corresponding distributions.

The layout of the paper is as follows. Section 2 provides a problem formulation and basic facts about models used and Bayesian filtering [9]. Section 3 presents a general solution for online state estimation for hybrid systems and specializes it for normal and multinomial models in Section 3.2. Section 4 demonstrates experiments with real data and comparison with state estimation using HMM. Section 5 contains a conclusion.

2 Problem formulation

A system to be considered exhibits both continuous and discrete behavior and can be treated as a hybrid system. Available observations concerning the system are of a mixed nature:

$$y_t = [y_t^c, y_t^d]', \quad u_t = [u_t^c, u_t^d]',$$

where y_t is an output vector measured at discrete time moments $t = \{1, \ldots, T\} \equiv t^*$, u_t is a known control input vector, and superscript c denotes a continuous type of a variable, while superscript d – a discrete variable. The present paper considers normally distributed variables and discrete multinomial variables. In general, variables with both the superscripts (^{c} and ^{d}) are column vectors.

Having at disposal measured data and assuming that system parameters are known (or estimated offline), a task is to estimate recursively the unobserved state of this system

$$x_t \;=\; [x_t^c, x_t^d]' \;=\; [x_{1,t}^c, \ldots, x_{C,t}^c, x_t^d]'$$

where $x_{i,t}^c$ is a normal entry, $i = 1, \ldots, C$ and x_t^d is a discrete multinomial scalar with number of possible values L.

2.1 Basic facts

State-space model Generally, a state-space model is taken in the form of the following probability (density) functions (p(d)fs),

$$\textit{observation model} \quad f\left(y_t | x_t, u_t\right), \tag{1}$$

$$\textit{state evolution model} \quad f\left(x_{t+1} | x_t, u_t\right). \tag{2}$$

for simplicity denoted as pdfs for random variables throughout the paper.

Bayesian filtering Bayesian filtering, estimating the system state, includes the following recursions. The first is the *data updating*

$$f\left(x_t\middle|\, d^t\right) = \frac{f\left(y_t\middle|\, x_t, u_t\right) f\left(x_t\middle|\, d^{t-1}\right)}{\int_{x^*} f\left(y_t\middle|\, x_t, u_t\right) f\left(x_t\middle|\, d^{t-1}\right)\, dx_t} \propto f\left(y_t\middle|\, x_t, u_t\right) f\left(x_t\middle|\, d^{t-1}\right), \tag{3}$$

which incorporates information contained in data $d^t = (d_1, \ldots, d_t)$, where $d_t \equiv (y_t, u_t)$ into the prior pdf $f\left(x_1\middle|\, d^0\right)$, which starts the recursions. The prior pdf expresses a subjective prior knowledge about the initial state x_1. The relation (3) also comprises the natural conditions of control [9], according to those $f\left(x_t\middle|\, u_t, d^t\right) = f\left(x_t\middle|\, d^t\right)$. $\propto$ means a proportionality. The second is the *time updating*

$$f\left(x_{t+1}\middle|\, d^t\right) = \int_{x^*} f\left(x_{t+1}\middle|\, x_t, u_t\right) f\left(x_t\middle|\, d^t\right) dx_t, \tag{4}$$

that corrects the state estimate updated by measurements.

Kalman filter Bayesian filtering (3)-(4) applied to linear normal state-space model (1)-(2) provides Kalman filter [10]. It means that the models take the following forms

$$y_t = Hx_t + Du_t + v_t, \tag{5}$$

$$x_{t+1} = Ax_t + Bu_t + w_t, \tag{6}$$

where H, D, A and B are parameters supposed to be known, u_t is known control input variable, and v_t is measurement Gaussian noise with zero mean and known covariance R_v and w_t is process Gaussian noise with zero mean and known covariance R_w. For these linear models Kalman filter [10] includes the following equations starting with prior mean $\mu_{1|0}$ and prior covariance matrix $P_{1|0}$.

$$KG = P_{t|t-1}H'(HP_{t|t-1}H' + R_v)^{-1}, \tag{7}$$

$$\mu_{t|t} = \mu_{t|t-1} + KG(y_t - H\mu_{t|t-1} - Du_t), \tag{8}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H'(R_v + HP_{t|t-1}H')HP_{t|t-1},$$

$$\mu_{t+1|t} = A\mu_{t|t} + Bu_t, \tag{9}$$

$$P_{t+1|t} = AP_{t|t-1}A' + R_w, \tag{10}$$

where KG is a Kalman gain, and $\mu_{t+1|t}$ and $P_{t+1|t}$ are the obtained mean and covariance matrix of the needed state estimate.

Chain rule The chain rule [10] is an operation intensively used in the paper, which has a form

$$f\left(a, b\middle|\, c\right) = f\left(a\middle|\, b, c\right) f\left(b\middle|\, c\right) \tag{11}$$

It decomposes the joint pdf $f\left(a, b\middle|\, c\right)$ into a product of conditional pdfs for any random variables a, b and c.

3 Online filtering for hybrid systems

Bayesian filtering is proposed to be done with the simultaneous data and time updating, i.e.,

$$f(x_{t+1}|d^t) \propto \int f(x_{t+1}|x_t,u_t) \left\{ \underbrace{f(y_t|x_t,u_t) f(x_t|d^{t-1})}_{\propto f(x_t|d^t)} \right\} dx_t. \tag{12}$$

3.1 General solution in pdfs

Models in a decomposed form For the considered hybrid system (see the problem formulation in Section 2) it is convenient to decompose models (1)-(2) so that to treat each state individually. Using the chain rule, one decomposes the observation model (1)

$$f(y_t|x_t,u_t) = f(y_t^c, y_t^d | x_t^c, x_t^d, u_t^c, u_t^d) \tag{13}$$

$$= f(y_t^c|y_t^d, x_t^c, x_t^d, u_t^c, u_t^d) f(y_t^d|x_t^c, x_t^d, u_t^c, u_t^d) = f(y_t^c|y_t^d, x_t^c, u_t^c) f(y_t^d|x_t^d, u_t^d)$$

realistically assuming that the past discrete state and a discrete input can be omitted from the condition for the continuous output y_t^c, and continuous entries – from the condition for discrete y_t^d. The obtained decomposition (13) represents a product of distributions. Similarly the state evolution model (2) $f(x_{t+1}|x_t,u_t)$ is decomposed according to the chain rule as

$$f(x_{t+1}^c, x_{t+1}^d | x_t^c, x_t^d, u_t^c, u_t^d) = f(x_{t+1}^c|x_{t+1}^d, x_t^c, x_t^d, u_t^c, u_t^d) f(x_{t+1}^d|x_t^c, x_t^d, u_t^c, u_t^d)$$

$$= f(x_{t+1}^c|x_{t+1}^d, x_t^c, u_t^c) f(x_{t+1}^d|x_t^d, u_t^d). \tag{14}$$

Prior pdfs Prior pdf $f(x_t|d^{t-1})$ is also chosen in the form

$$f(x_t^c, x_t^d|d^{t-1}) = f(x_t^c|x_t^d, d^{t-1}) f(x_t^d|d^{t-1}). \tag{15}$$

Online hybrid filter Substituting models (14), (13) and prior pdf (15) in (12), one obtains

$$f(x_{t+1}|d^t) \propto \int_{x^{c*}} \sum_{x^{d*}} \underbrace{f(x_{t+1}^c|x_{t+1}^d, x_t^c, u_t^c) f(x_{t+1}^d|x_t^d, u_t^d)}_{f(x_{t+1}|x_t,u_t)}$$

$$\times \underbrace{f(y_t^c|y_t^d, x_t^c, u_t^c) f(y_t^d|x_t^d, u_t^d)}_{f(y_t|x_t,u_t)} \underbrace{f(x_t^c|x_t^d, d^{t-1}) f(x_t^d|d^{t-1})}_{prior\ pdf} dx_t^c$$

$$= \underbrace{\sum_{x^{d*}} f(x_{t+1}^d|d^t) f(x_{t+1}^c|x_{t+1}^d, d^t)}_{sum\ of\ distributions} \tag{16}$$

Relation (16) results in the sum of distributions and loses the prescribed original form of the posterior pdf. It is necessary to restore it in order to use in the next step of recursion (12). Kerridge inaccuracy [11], a part of Kullback-Leibler divergence [12], is adopted as a theoretically justified proximity measure. This divergence is known to be an optimal tool within the Bayesian approach [9]. An approximation based on Kerridge inaccuracy is an explicit solution, which restores the original form of the pdf via computation of a specific weighted combination of the pdfs involved in (16). For any random variable a, Kerridge inaccuracy is used to measure the proximity of pdfs $f(a)$ and $\hat{f}(a)$

$$K_a(f(a)||\hat{f}(a)) = \int_{a*} f(a) \ln \frac{1}{\hat{f}(a)} \, da \tag{17}$$

and its minimization allows to find the approximated pdf $\hat{f}(a)$. According to this approximation [9], sum (16) is replaced by the product

$$\hat{f}(x^c_{t+1}|x^d_{t+1}, d^t) f(x^d_{t+1}|d^t), \tag{18}$$

which is used as the prior pdf for the next step of recursive estimation.

The proposed general solution is universal in the sense of exploited distributions. The provided specialization shows usage of the approach with normal and multinomial models.

3.2 Solution for normal and multinomial models

The decomposed observation model (1) is a product of the multivariate K-dimensional (i.e., $y^c_t = [y^c_{1,t}, \dots, y^c_{K,t}]'$) normal output distribution (5) and the multinomial distribution provided by an output table, i.e.,

$$f(y^c_t|y^d_t, x^c_t, u^c_t) f(y^d_t|x^d_t, u^d_t) = \mathcal{N}(Hx^c_t + Du^c_t, R_v)\alpha_{y^d_t|x^d_t u^d_t} \tag{19}$$

where $\alpha_{y^d_t=q|x^d_t=l, u^d_t=n}$ are known probabilities of output $y^d_t = q$ under conditions of $x^d_t = l$ and $u^d_t = n$, and it holds $\sum_q^Q \alpha_{q|ln} = 1$, $\alpha_{q|ln} \geq 0 \;\forall\; q, l, n$ and Q.

The decomposed state evolution model (2) represents a product of the multivariate C-dimensional normal distribution (6) and the multinomial distribution presented by a transition table, i.e.,

$$f(x^c_{t+1}|x^d_{t+1}, x^c_t, u^c_t) f(x^d_{t+1}|x^d_t, u^d_t) = \mathcal{N}(Ax^c_t + Bu^c_t, R_w)\beta_{x^d_{t+1}|x^d_t u^d_t} \tag{20}$$

with known probabilities $\beta_{x^d_{t+1}=l|x^d_t=m, u^d_t=n}$ of transition to state $x^d_{t+1} = l$ under conditions of $x^d_t = m$ and $u^d_t = n$. It holds $\sum_l^L \beta_{l|mn} = 1, \beta_{l|mn} \geq 0 \;\forall\; l, m, n$, and $l = 1, \dots, L$.

Choice of prior distributions The prior pdfs (15) $f(x_t^c|x_t^d, d^{t-1})f(x_t^d|d^{t-1})$ are specialized as

$$\mathcal{N}(\mu_{t|t-1}, P_{t|t-1})p_{x_t^d(t)} \tag{21}$$

of C-dimensional prior normal distribution with initial mean value $\mu_{t|t-1}$ and covariance $P_{t|t-1}$ and of the multinomial distribution in the form of the prior probability $p_{x_t^d=l(t)}$ of $x_t^d = l$, where $\sum_l^L p_{l(t)} = 1$, $p_{l(t)} \geq 0 \ \forall \ l$.

Filter for normal and multinomial models For normal models the solution computationally coincides with the Kalman filter (7)-(10) run for each value of x_t^d.

For discrete multinomial models the filtering takes the following form. For each value l of x_{t+1}^d and with known output $y_t^d = q$ and known input $u_t^d = n$ the probability is computed as

$$p_{x_{t+1}^d(t)} = \beta_{l|1n}\alpha_{q|1n}p_{1(t)} + \beta_{l|2n}\alpha_{q|2n}p_{2(t)} + \ldots + \beta_{l|Ln}\alpha_{q|Ln}p_{L(t)} \tag{22}$$

and then normalized. In this case the resulted relation (16) is mixture distribution. The mixture distribution $\sum_{l=1}^L p_{l(t+1)}\mathcal{N}_l(\mu_{t+1|t}, P_{t+1|t})$ is replaced by the approximated normal distribution based on Kerrigde inaccuracy with

$$\hat{\mu}_{t+1|t} = \sum_{l=1}^L p_{l(t+1)}\mu_{l,t+1|t}, \tag{23}$$

$$\hat{P}_{t+1|t} = \sum_{l=1}^L p_{l(t+1)}P_{l,t+1|t} + \sum_{l=1}^L p_{l(t+1)}(\hat{\mu}_{t+1|t} - \mu_{l,t+1|t-})^2 \tag{24}$$

where $\mu_{l,t+1}$ and $P_{l,t+1}$ denote results of the Kalman filter (7)-(10) obtained for each value l. The approximation (23)-(24) is then used as the prior normal distribution for the next step of the recursion.

4 Illustrative experiments

The presented approach can be found close to the IMM filter [1], comparison with that for simulated data is described in [13]. This paper demonstrates a testing of the algorithm for real traffic-control data from one of the controlled microregions in Prague. The paper compares results of the filtering with those of the HMM algorithms available in standard package of MATLAB.

A normally distributed state x_t^c of the considered hybrid system is a four-dimensional queue length of cars waiting for passing through a traffic microregion. A full dimension of the taken normal state is eight, since occupancy of a measured detector is added to the vector to ensure observability of the model. A discrete state x_t^d (system mode) is a level of service (LoS) of the microregion. It expresses a degree of traffic saturation in that sense how easy cars can pass through the microregion with 4 possible values: from 1 (the best) to 4 (the worst).

The measured data used were: y_t^c – car outgoing intensity along with occupancy of a measured detector; y_t^d – a time mode of a workday (morning peak-hour time, lunch, late afternoon peak-hour time, evening); u_t^c – a relative time of the green light; u_t^d – a discrete variable, reflecting whether the saturated strategy of the adaptive control is used or not. A duration of the online filtering was 1 workday, which corresponds to 288 time periods. The filtering started at midnight that simplifies a choice of prior distributions (i.e., zero queue length and LoS=1).

The car queue length for individual arms of the traffic microregion was estimated using Kalman filter (7)-(10) with traffic linear state-space model available in [14]. Results for one of the arms is shown at Fig. 1. A rest of the arm state estimates is of a similar quality depending on intensity.

The discrete state (i.e., LoS) has been estimated via the proposed hybrid filter and with the help of the HMM algorithm. The same output and transition tables were used firstly for recursions of the hybrid filter and then as initial parameters for the HMM estimator. Results for both the estimators are shown at Fig. 2, at top and bottom figure respectively. The hybrid filter demonstrates better results: a number of correctly point-estimated states for the proposed hybrid filter is 240 for 288 possible ones, while the HMM algorithm gives 114. The experiments have been repeated for other data samples. The hybrid filter again demonstrated a better stability for them. Two peak-

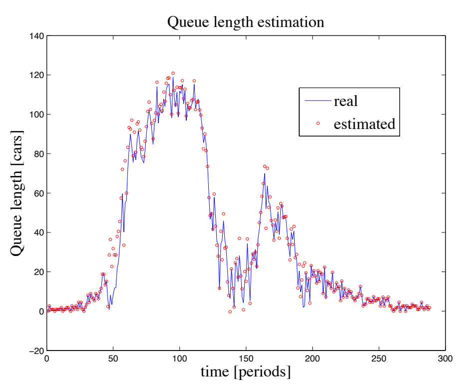

Fig. 1. Online filtering of car queue length

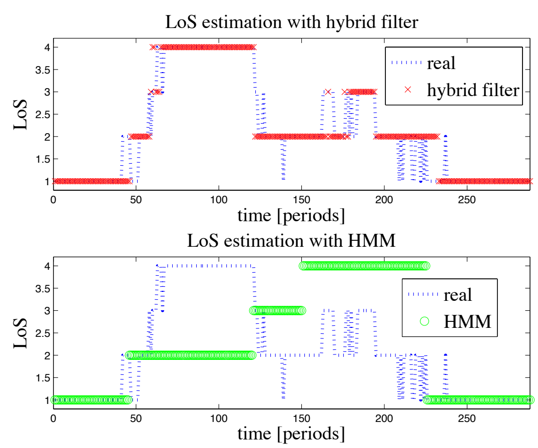

Fig. 2. LoS estimation with the hybrid filter (top) and the HMM (bottom)

hour times of a workday (morning between 50 and 120 time periods and late afternoon between 150 and 200) can be easily seen both at Fig. 1 and Fig. 2 in the course of the queue length and in the switching of the LoS values.

5 Conclusion

The paper describes the recursive state estimation for dynamic hybrid systems. The proposed algorithm runs online and uses explicit solutions. Not only the discrete state is treated as a pointer to the current mode in which the system is located, but the whole state-space model is taken as the discrete one with discrete measurements (output and input). It enables to estimate jointly the normal and discrete multinomial variables. Comparison with one of the counterparts is shown.

Acknowledgements
This research has been supported by grant MŠMT 1M0572.

References

1. Y. Bar-Shalom, T.Kirubarajan and X.-Rong Li, *Estimation with Applications to Tracking and Navigation*, Wiley, New York, NY, USA, 2002.
2. M.S. Grewal and A.P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB. 2nd edition*, Wiley, 2001.
3. F. Dufour R.J. Elliot and D.D. Sworder, "Exact hybrid filters in discrete time," *IEEE Transactions on Automatic Control*, vol. 41, no. 12, pp. 1807–1810, 1996.
4. R.H. Shumway and D.S. Stoffer, "Dynamic linear models with switching," *Journal of the American Statistical Association*, vol. 86, no. 415, pp. 763–769, 1991.
5. Q. Zhang, "Optimal filtering of discrete-time hybrid systems," *Journal of Optimization Theory and Applications*, vol. 100, no. 1, 1999.
6. A. Doucet and C. Andrieu, "Iterative algorithms for state estimation of jump markov linear systems," *IEEE Transactions on Signal Processing*, vol. 49, no. 6, pp. 1216–1227, 2001.
7. R. Chen and J. S. Liu, "Mixture kalman filters," *J. R. Statist. Soc. B*, vol. 62, pp. 493–508, 2000.
8. M. J. Beal, Z. Ghahramani, and C. E. Rasmussen, "The infinite hidden markov model," in *Advances in Neural Information Processing Systems*, 2002, vol. 14.
9. M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař, *Optimized Bayesian Dynamic Advising: Theory and Algorithms*, Springer, London, 2005.
10. V. Peterka, "Bayesian system identification," in *Trends and Progress in System Identification*, P. Eykhoff, Ed., pp. 239–304. Pergamon Press, Oxford, 1981.
11. D.F. Kerridge, "Inaccuracy and inference," *Journal of Royal Statistical Society*, vol. B 23, pp. 284–294, 1961.
12. S. Kullback and R. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, pp. 79–87, 1951.
13. E. Suzdaleva and I. Nagy, "Online soft sensor for hybrid systems with mixed continuous and discrete measurements," *Computers and Chemical Engineering*, 2010, submitted.
14. E. Suzdaleva, "Prior knowledge processing for initial state of Kalman filter", *Int. Journal of Adaptive Control and Signal Processing*, vol. 24, 3, (2010), p. 188–202.