Automatica 57 (2015) 97–104

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Informative windowed forecasting of continuous-time linear systems for mutual information-based sensor planning ☆

Han-Lim Choi [1], Jung-Su Ha

Department of Aerospace Engineering, KAIST, Yuseong, Daejeon, Republic of Korea

ARTICLE INFO

Article history:
Received 31 July 2013
Received in revised form
24 February 2015
Accepted 8 April 2015
Available online 16 May 2015

Keywords:
Mutual information
Windowed forecasting
Informative planning
Continuous-time system

ABSTRACT

This paper presents an expression of mutual information that defines the information gain in planning of sensing resources, when the goal is to reduce the forecast uncertainty of some quantities of interest and the system dynamics is described as a continuous-time linear system. The method extends the smoother approach in Choi and How (2010b) to handle a more general notion of the verification entity—continuous sequence of variables over some finite time window in the future. The expression of mutual information for this windowed forecasting case is derived and quantified, taking advantage of an underlying conditional independence structure and utilizing a two-filter formula for fixed-interval smoothing with correlated noises. Two numerical examples on (a) a two-state linear system with time-varying one-way coupling dynamics, and (b) idealized weather forecasting with moving verification paths demonstrate the validity of the proposed quantification methodology.

1. Introduction

Planning on utilization of sensing resources to gather information out of an environment has been spotlighted in many contexts, the objective of this planning often being uncertainty reduction of some entities of interest—termed *verification* entities herein. Mutual information has been one of the most popular metrics adopted to define/represent this objective in various contexts: tracking of kinematic variables of moving targets through measurement along mobile sensor trajectories (Grocholsky, 2002; Hoffmann & Tomlin, 2010), weather forecast improvement over some region of interest in the future with UAV sensor networks (Choi & How, 2010b, 2011a,b), accurate prediction of a spatially distributed field described by Gaussian processes (Krause, Singh, & Guestrin, 2008), informative management of deployed fixed sensor networks (Choi, How, & Barton, 2013; Choi & Lee, 2015; Lee, Teo, Lim, & B, 2001; Williams, Fisher III, & Willsky, 2007), adaptive landmark selection in simultaneous localization and mapping of mobile robots (Kretzschmar & Stachniss, 2012), and Bayesian belief propagation over the grid-based search space (Julian, Angermann, Schwager, & Rus, 2012). Progresses in computation of mutual information in various formalism (Duncan, 1970; Guo, Shamai, & Verdu, 2005; Kadota, Zakai, & Ziv, 1971; Lee et al., 2001; Mitter & Newton, 2005; Newton, 2006, 2007; Zakai, 2005) have also contributed to this popularity of mutual information in these applications.

While many of these mutual information-based planning studies have dealt with the case where the verification time is same or just one time-step further of the planning horizon, there is a class of problem termed *informative forecasting* that takes particular care for the case where the verification time is significantly greater than the mission horizon. Although less popular in the literature, the informative forecasting problem can handle applications such as (i) adaptive sampling in the context of numerical weather prediction that considers design of sensor networks deployed in the near future (e.g., in 24 h) while the goal is to improve forecast in the far future (e.g., 3–5 days later), and (ii) prediction of indoor contaminant distribution in some future time with wireless indoor sensor networks taken over short period of time. The authors have presented methods to efficiently but correctly quantify the mutual information in this context of informative forecasting for discrete selection case (Choi & How, 2011b), discrete constrained path design (Choi & How, 2011a), and continuous trajectory planning (Choi & How, 2010b), taking advantage of underlying properties of mutual information.

This paper extends the approach in Choi and How (2010b) in that a more general notion of verification quantities is introduced.

☆ The material in this paper was partially presented at the 11th International Conference on Control, Automation, Robotics and Vision (ICARCV), December 7–10, 2010, Singapore. This paper was recommended for publication in revised form by Associate Editor Valery Ugrinovskii under the direction of Editor Ian R. Petersen.

E-mail addresses: hanlimc@kaist.ac.kr (H.-L. Choi), wjdtn1404@kaist.ac.kr (J.-S. Ha).

[1] Tel.: +82 42 350 3727; fax: +82 42 350 3710.

http://dx.doi.org/10.1016/j.automatica.2015.04.011

For some applications, it makes more sense to reduce uncertainty in the entities of interest over some finite window of time instead of a single particular time instance (for example, weather forecast over the weekend). The smoother form in Choi and How (2010b) cannot directly be used for this *windowed forecasting* case, because the mutual information between two continuous random processes (as opposed to one finite-dimensional random vector and one random process) needs to be calculated. This paper presents a formula for the mutual information for this windowed forecasting that is indeed quite similar to the form in Choi and How (2010b), while the only difference is in the process of calculating the conditional initial covariance conditioned on the verification entity. A two-filter form Kalman smoothing formula with correlated noise in Fujita and Fukao (1970) is adapted for this calculation. Two numerical examples are presented to validate the proposed method in terms of necessity and applicability. While a preliminary work (Choi & How, 2010a) proposed a relevant concept of the windowed forecasting, this article presents elaborated and corrected theoretical results for a more general problem setting and also provides much more sophisticated numerical case studies.

2. Problem description

2.1. Continuous-time linear system model

Consider the dynamics of objects/environment of interest with a finite dimensional state vector $X_t \in \mathbb{R}^{n_X}$ that is described by the following linear (time-varying) system:

$$dX_t = A(t)X_t dt + B(t)dW_t \tag{1}$$

where W_t is an n_W-dimensional zero-mean Brownian motion process satisfying $\mathbb{E}[(W_{t_1} - W_0)(W_{t_2} - W_0)'] = \int_0^{\min\{t_1,t_2\}} \Sigma_W(s)ds$ with positive-definite Σ_W. The prime sign (′) throughout the paper denotes the transpose of a matrix. The initial condition of the state, X_0 is normally distributed as $X_0 \sim \mathcal{N}(\mu_0, P_0)$, $P_0 \succ 0$.

An $\mathbb{R}^{n_Z}$-valued continuous-time measurement process is given by

$$dZ_t = C(t)X_t dt + dN_t \tag{2}$$

where N_t is a n_Z-dimensional Brownian motion process with $\mathbb{E}[(N_{t_1} - N_0)(N_{t_2} - N_0)'] = \int_0^{\min\{t_1,t_2\}} \Sigma_N(s)ds$ with positive-definite Σ_N, which is independent of X_t and W_t. The initial value of the measurement process Z_0 is assumed to be normally distributed with covariance $\mathbb{E}[Z_0 Z_0'] = \Gamma_0 \succ 0$; this initial value summarizes all the information collected before the initial time (Mitter & Newton, 2005).[2] Also, a measurement *path* over the time window $[t_1, t_2]$ is defined as

$$\mathcal{Z}_{[t_1,t_2]} = \{Z_t : t \in [t_1, t_2]\}. \tag{3}$$

Definition 1. The verification variables are a (possibly time-varying) linear combination of the state variables whose uncertainty reduction is of interest:

$$V_t = M_V(t)X_t \in \mathbb{R}^{n_V} \tag{4}$$

with $M_V(t) \in \mathbb{R}^{n_V \times n_X}$ termed as *verification matrix*, which is assumed to be *differentiable* herein. A continuous sequence of (time-varying) verification variables, termed *verification path*, is also defined as:

$$\mathcal{V}_{[t_1,t_2]} = \{V_t : t \in [t_1, t_2]\}. \tag{5}$$

[2] Note that this definition of measurement process is an integral of the measurement defined in the authors' prior work (Choi & How, 2010b) in which no information is assumed from the negative time.

2.2. Informative pointwise forecasting (IPF)

One case of interest is when the verification entity is the verification *variables* at some fixed verification time, T. In this case, the informative forecasting problem can be written as the following optimization:

$$\max_{C(t)\in\mathcal{C}:t\in[0,\tau]} \mathcal{I}(V_T; \mathcal{Z}_{[0,\tau]}) \tag{IPF}$$

with some $\tau \in [0, T]$, where $\mathcal{C}$ represents a set of admissible observation functions defined by the characteristics of sensors, and $\mathcal{I}(Y_1; Y_2)$ denotes the mutual information between two random quantities Y_1 and Y_2 (e.g. random variables, random processes, random functions). Note that (2) defines a direct relation between the observation matrix $C(t)$ to the measurement process Z_t; thus, (IPF) finds the (continuous) sequence of $C(t)$ over $[0, \tau]$ that is expected to result in largest reduction of entropy in V_T when noisy measurement is taken accordingly.

Exploiting conditional independence, we proposed an expression for the mutual information, $\mathcal{I}(V_T; \mathcal{Z}_{[0,\tau]})$, as the difference between the unconditioned and the conditioned mutual information for the filtering problem (Choi & How, 2010b):

$$\mathcal{I}(V_T; \mathcal{Z}_{[0,\tau]}) = \mathcal{I}(X_\tau; \mathcal{Z}_{[0,\tau]}) - \mathcal{I}(X_\tau; \mathcal{Z}_{[0,\tau]}|V_T). \tag{6}$$

With (6), the *smoother form* of the mutual information for forecasting is derived as

$$\mathcal{I}(V_T; \mathcal{Z}_{[0,\tau]}) = \mathcal{J}_0(\tau) - \tfrac{1}{2}\,\mathrm{ldet}(I + Q_X(\tau)\Delta_S(\tau)) \tag{7}$$

with $\mathcal{J}_0 \triangleq \frac{1}{2}\mathrm{ldet}S_{X|V} - \frac{1}{2}\mathrm{ldet}S_X$ and $\Delta_S \triangleq S_{X|V} - S_X$, where ldet stands for log det of a positive definite matrix. The matrices $S_X(\tau) \triangleq \mathrm{Cov}^{-1}(X_\tau)$, $S_{X|V}(\tau) \triangleq \mathrm{Cov}^{-1}(X_\tau|V_T)$, and $Q_X(\tau) \triangleq \mathrm{Cov}(X_\tau|\mathcal{Z}_{[0,\tau]})$ are determined by the following matrix differential equations:

$$\dot{S}_X = -S_X A - A'S_X - S_X B\Sigma_W B'S_X \tag{8}$$

$$\dot{S}_{X|V} = S_{X|V}B\Sigma_W B'S_{X|V} - S_{X|V}(A + B\Sigma_W B'S_X) - (A + B\Sigma_W B'S_X)'S_{X|V} \tag{9}$$

$$\dot{Q}_X = AQ_X + Q_X A' + B\Sigma_W B' - Q_X C'\Sigma_N^{-1}CQ_X \tag{10}$$

with initial conditions $S_X(0) = P_0^{-1}$, $S_{X|V}(0) = P_{0|V}^{-1}$, and $Q_X(0) = \left(P_0^{-1} + C(0)'\Gamma_0^{-1}C(0)\right)^{-1}$. The initial value of the Riccati matrix $Q_X(0)$ takes a different value from the authors' prior work (Choi & How, 2010b), as herein the information collected from negative times is also incorporated as suggested in Mitter and Newton (2005). The conditional initial covariance $P_{0|V} \triangleq \mathrm{Cov}(X_0|V_T) \succ 0$ can be calculated in advance by a fixed-point smoothing process, or simply by

$$P_{0|V} = P_0 - P_0\Phi'_{(T,0)}M'_V\left[M_V P_X(T)M'_V\right]^{-1}M_V\Phi_{(T,0)}P_0$$

where $\Phi_{(t_2,t_1)}$ is the state transition matrix from t_1 to t_2, which becomes $e^{A(t_2-t_1)}$ for a time-invariant case.

In Choi and How (2010b), we demonstrated that the smoother form is preferred to the filter form, which explicitly calculates the prior and the posterior entropies of V_T by integrating the Lyapunov and the Riccati equation over $[0, T]$, in terms of the computational efficiency and accessibility to on-the-fly knowledge of information accumulation.

2.3. Informative windowed forecasting (IWF)

This paper newly considers a more general version of the informative forecasting problem where the entity of interest is a verification path over some time window $[T_i, T_f]$. This generalized problem can be written as:

$$\max_{C(t)\in\mathcal{C}:t\in[0,\tau]} \mathcal{I}(\mathcal{V}_{[T_i,T_f]}; \mathcal{Z}_{[0,\tau]}). \tag{IWF}$$