Utility and Privacy in Object Tracking from Video Stream using Kalman Filter

Niladri Das & Dr. Raktim Bhattacharya

Dept. of Aerospace Engineering,
Intelligent Systems Research Laboratory,
Texas A&M University,
College Station, Texas, USA

June 27, 2020

Table of Contents

We consider the problem of maintaining privacy and utility while tracking an object in a video stream using Kalman filtering.

- Privacy: localization accuracy of an object will not improve beyond a certain level.
- Utility: localization accuracy of the same object will always remain under a certain threshold.

Contributions

- We are not aware of any prior works related to privacy and utility in object tracking using filtering from a video stream.
- We proposed two techniques:
  - Method 1: Privacy ensuring
  - Method 2: Utility ensuring

System model

We model the object detection process from a video frame using a linear discrete time stochastic systems $\bar{\mathcal{S}}$ described by the model of the form

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}\boldsymbol{x}_k + \boldsymbol{w}_k, \tag{1a}$$

$$\boldsymbol{y}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{n}_k, \tag{1b}$$

- $k = 0, 1, 2, \ldots$ represents the frame index
- $\boldsymbol{w}_k \in \mathbb{R}^{n_w}$ and $\boldsymbol{n}_k \in \mathbb{R}^{n_y}$ are the process and measurement noise
- $\{\boldsymbol{w}_k\}$ and $\{\boldsymbol{n}_k\}$ are zero-mean, Gaussian, independent white random processes
- $\boldsymbol{w}_k \sim \mathcal{N}(0, \boldsymbol{Q}), \boldsymbol{n}_k \sim \mathcal{N}(0, \boldsymbol{R})$
- $\boldsymbol{R}$ is a diagonal matrix. Inverse of $\boldsymbol{R}$ is the precision matrix

The optimal state estimator for the stochastic system $\bar{\mathcal{S}}$ is the Kalman filter, defined by

$$
\begin{aligned}
\boldsymbol{K}_k &= \boldsymbol{\Sigma}_k^- \boldsymbol{H}^T \left[\boldsymbol{H} \boldsymbol{\Sigma}_k^- \boldsymbol{H}^T + \boldsymbol{R} \right]^{-1}, && \text{(Kalman Gain)} \\
\boldsymbol{\mu}_k^- &= \boldsymbol{F} \boldsymbol{\mu}_{k-1}^+, && \text{(Mean Propagation)} \\
\boldsymbol{\Sigma}_k^- &= \boldsymbol{F} \boldsymbol{\Sigma}_{k-1}^+ \boldsymbol{F}^T + \boldsymbol{Q}, && \text{(Covariance Propagation)} \\
\boldsymbol{\mu}_k^+ &= \boldsymbol{F} \boldsymbol{\mu}_{k-1}^+ + \boldsymbol{K}_k (\boldsymbol{y}_k - \boldsymbol{H} \boldsymbol{\mu}_k^-), && \text{(Mean Update)} \\
\boldsymbol{\Sigma}_k^+ &= (\boldsymbol{I}_{n_x} - \boldsymbol{K}_k \boldsymbol{H}) \boldsymbol{\Sigma}_k^-, && \text{(Covariance Update)} \\
\boldsymbol{\Sigma}_0^+ &= \boldsymbol{\Sigma}_0, && \text{(Initial State Covariance)}
\end{aligned}
$$

- $\boldsymbol{\Sigma}_k^-, \boldsymbol{\Sigma}_k^+ \in \mathbb{R}^{n_x \times n_x}$: the prior and posterior error covariance in frame k.
- $\boldsymbol{\mu}_k^-, \boldsymbol{\mu}_k^+ \in \mathbb{R}^{n_x}$: the prior and posterior mean estimate of the true state $\boldsymbol{x}_k$.

Utility Problem

- Utility of the object detection system can be specified by an upper bound on the steady-state estimation error due to filtering.
- We are interested in the measurement noise $\boldsymbol{R}$ that ensures the steady state prior covariance matrix to be upper-bounded by a prescribed positive definite matrix $\boldsymbol{\Sigma}_{\infty}^{d}$ for the detection system modeled in eqn. 1.
- The parameter $\boldsymbol{R}$ is a measure of maximum inaccuracies allowed in the detection system.

Privacy Problem

- Privacy requirement is centered around a particular frame (say $k+1^{\text{th}}$). It is specified by a lower bound on the estimation error $\boldsymbol{\Sigma}_{k+1}^{+}$ after the Kalman update, for that particular frame.
- Privacy scenario differs from the utility case, where we focus on the steady-state error.
- We are interested in calculating $\boldsymbol{R}$ such that the posterior error covariance matrix $\boldsymbol{\Sigma}_{k+1}^{+}$ is lower-bounded by a prescribed positive definite matrix $\boldsymbol{\Sigma}_{k+1}^{d}$.
- The parameter $\boldsymbol{R}$ is a measure of minimal noise that needs to be artificially added to the $k+1^{\text{th}}$ image frame to ensure privacy with respect to accurate localization.

TEXAS A&M UNIVERSITY

Theorem

Given $\boldsymbol{\Sigma}_\infty^d$, the desired steady-state error variance, the optimal algorithmic precision $\boldsymbol{\Upsilon} := \boldsymbol{R}^{-1}$ that satisfies $\boldsymbol{\Sigma}_\infty \preceq \boldsymbol{\Sigma}_\infty^d$ is given by the following optimization problem,

$$\left.\begin{aligned} &\min_{\boldsymbol{\Upsilon}} \mathbf{tr}\left[\boldsymbol{W}\boldsymbol{\Upsilon}\boldsymbol{W}^T\right] \\ \text{s.t.}\quad &\begin{bmatrix} \boldsymbol{M}_{11} & \boldsymbol{F}\boldsymbol{\Sigma}_\infty^d\boldsymbol{H}^T \\ \boldsymbol{H}\boldsymbol{\Sigma}_\infty^d\boldsymbol{F}^T & \boldsymbol{L}+\boldsymbol{L}\boldsymbol{\Upsilon}\boldsymbol{L} \end{bmatrix} \succeq 0, \end{aligned}\right\} \tag{2}$$

where

$$\boldsymbol{\Upsilon} \succeq 0,\ \boldsymbol{L} := \boldsymbol{H}\boldsymbol{\Sigma}_\infty^d\boldsymbol{H}^T,$$
$$\boldsymbol{M}_{11} := \boldsymbol{\Sigma}_\infty^d - \boldsymbol{F}\boldsymbol{\Sigma}_\infty^d\boldsymbol{F}^T - \boldsymbol{Q} + \boldsymbol{F}\boldsymbol{\Sigma}_\infty^d\boldsymbol{H}^T\boldsymbol{L}^{-1}\boldsymbol{H}\boldsymbol{\Sigma}_\infty^d\boldsymbol{F}^T,$$

with $\boldsymbol{\Upsilon} \in \mathbb{R}^{n_y \times n_y}$ and $\boldsymbol{W} \in \mathbb{R}^{n_y \times n_y}$, is user defined.

- We assume complete detectability of $(\boldsymbol{F}, \boldsymbol{H})$ and stabilizability of $(\boldsymbol{F}, \boldsymbol{Q}^{1/2})$[1] for eqn. 1.
- This ensure existence and uniqueness of the steady state prior covariance matrix $\boldsymbol{\Sigma}_{\infty}$ (for a fixed $\boldsymbol{R}$) for the corresponding DARE.
- The linear matrix inequality (LMI) in eqn. 2 gives the feasible set of $\boldsymbol{R} := \boldsymbol{\Upsilon}^{-1}$.
- We introduced the convex cost function $\mathbf{tr}\left[\boldsymbol{W}\boldsymbol{\Upsilon}\boldsymbol{W}^T\right]$ to calculate the most economical choice of $\boldsymbol{R}$.

[1] Brian DO Anderson and John B Moore. "Optimal filtering". In: *Englewood Cliffs* 21 (1979), pp. 22–95.

The minimal steady-state covariance of the estimate that any object detection setup can achieve modeled as in eqn. 1, is the solution to the following DARE

$$\boldsymbol{\Sigma}_\infty = \boldsymbol{F}\boldsymbol{\Sigma}_\infty\boldsymbol{F}^T + \boldsymbol{Q} - \boldsymbol{F}\boldsymbol{\Sigma}_\infty\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{\Sigma}_\infty\boldsymbol{H}^T\right)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}_\infty\boldsymbol{F}^T \tag{3}$$

This provides a theoretical lower bound on the prescribed $\boldsymbol{\Sigma}^d_\infty$ that we can achieve.

Theorem

Given $\boldsymbol{\Sigma}^d_{k+1}$*, the desired predicted error variance at time* $k+1$*, the optimal measurement noise* $\boldsymbol{R}_p$ *that satisfies* $\boldsymbol{\Sigma}^-_{k+1} \succeq \boldsymbol{\Sigma}^d_{k+1}$ *for a known* $\boldsymbol{\Sigma}^-_k$*, is given by the following optimization problem,*

$$\left.\begin{array}{l} \min\limits_{\boldsymbol{R}_p} \mathbf{tr}\left[\boldsymbol{W}\boldsymbol{R}_p\boldsymbol{W}^T\right] \text{ subject to} \\ \begin{bmatrix} \boldsymbol{M}_{11} & \boldsymbol{L} \\ \boldsymbol{L}^T & \boldsymbol{L}_2 + \boldsymbol{R}_p \end{bmatrix} \succeq 0, \end{array}\right\} \tag{4}$$

$$\boldsymbol{R}_p \succeq 0,\ \boldsymbol{L}_1 := \boldsymbol{F}\boldsymbol{\Sigma}^-_k\boldsymbol{H}^T,\ \boldsymbol{L}_2 := \boldsymbol{H}\boldsymbol{\Sigma}^-_k\boldsymbol{H}^T + \boldsymbol{R}_s \text{ and}$$

$$\boldsymbol{M}_{11} := -\boldsymbol{\Sigma}^d_{k+1} + \boldsymbol{F}\boldsymbol{\Sigma}^-_k\boldsymbol{F}^T + \boldsymbol{Q},$$

with $\boldsymbol{R}_p \in \mathbb{R}^{n_y \times n_y}$. *The variable* $\boldsymbol{W} \in \mathbb{R}^{n_y \times n_y}$, *is user defined .*

- The LMI in eqn. 4 gives the convex feasible set for $\boldsymbol{R}_p$ that ensures lower bound on the posterior covariance in the $k+1^{\text{th}}$ frame.
- We impose a cost convex cost function $\mathbf{tr}\left[\boldsymbol{W}\boldsymbol{R}_p\boldsymbol{W}^T\right]$ to calculate an optimal $\boldsymbol{R}_p$.

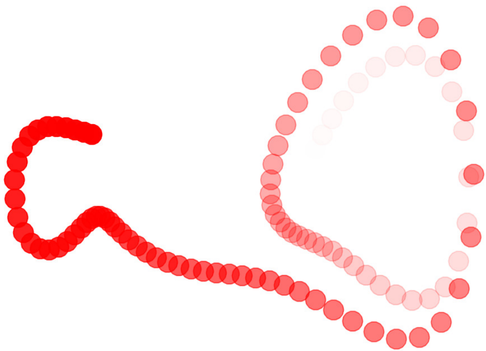

Figure 1: Time evolution of an object with darker shades representing more recent location.

The dynamics in the pixel frame from frame k to $k+1$

$$\underbrace{\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ \delta x_{k+1} \\ \delta y_{k+1} \end{bmatrix}}_{\boldsymbol{x}^p_{k+1}} = \underbrace{\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{F}} \underbrace{\begin{bmatrix} x_k \\ y_k \\ \delta x_k \\ \delta y_k \end{bmatrix}}_{\boldsymbol{x}^p_k} + \boldsymbol{w}_k, \tag{5}$$

$$\boldsymbol{y}_k = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}_{\boldsymbol{H}} \begin{bmatrix} x_k \\ y_k \\ \delta x_k \\ \delta y_k \end{bmatrix} + \boldsymbol{n}_k, \tag{6}$$

where $\boldsymbol{x}^p_k$ is the pixel coordinates of the moving object in the k^{th} frame

- Total of 500 frames in this video with 425 rows and 570 columns in each frame.
- The pair $(\boldsymbol{F}, \boldsymbol{H})$ is completely detectable and $(\boldsymbol{F}, \boldsymbol{Q}^{1/2})$ is completely stabilizable, which ensures existence and uniqueness of positive solution to the induced DARE due to Kalman filtering.

A homography exists between the pixel coordinates $(\boldsymbol{x}^p)$ and the spatial coordinates $(\boldsymbol{x})$. The homography in this numerical problem is represented as an affine map

$$\boldsymbol{x}^{\mathsf{p}} = \underbrace{\begin{bmatrix} 0 & \frac{n_r}{4} \\ -\frac{n_c}{4} & 0 \end{bmatrix}}_{\boldsymbol{U}} \boldsymbol{x} + \begin{bmatrix} \frac{n_r}{2} \\ \frac{n_c}{2} \end{bmatrix}.$$

The affine map induces a covariance relation $\boldsymbol{\Sigma}_{\boldsymbol{x}^{\mathsf{p}}\boldsymbol{x}^{\mathsf{p}}} = \boldsymbol{U}\boldsymbol{\Sigma}_{\boldsymbol{x}\boldsymbol{x}}\boldsymbol{U}^T$ from the pixel to the spatial coordinates.

Utility Results I

- The theoretical lower bound on utility in the pixel coordinates for $\boldsymbol{Q} = \textbf{diag}([0.1\ 0.1\ 50\ 50])$ is $\boldsymbol{\Sigma}^{\mathsf{lb}}_{\boldsymbol{x}^{\mathsf{p}}\boldsymbol{x}^{\mathsf{p}}} = \textbf{diag}([54.891\ 54.891])$
- If we allow for less precise filtering in pixel coordinates which can ensure a error covariance in the estimate of $1.5\boldsymbol{\Sigma}^{\mathsf{lb}}_{\boldsymbol{x}\boldsymbol{x}}$, the convex optimization problem yields an optimal precision requirement of

$$\boldsymbol{\Upsilon}^{*} = \textbf{diag}([0.660\ 0.660])$$

with $\boldsymbol{W}$ chosen to be identity.

Utility Results II

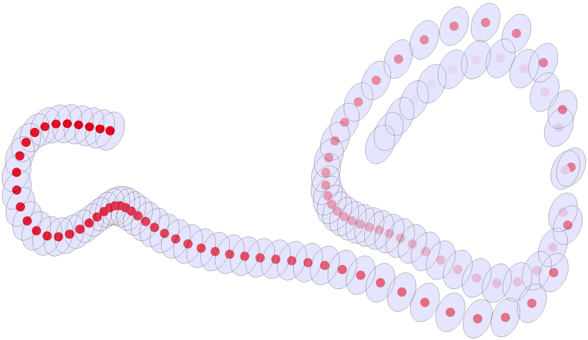

Figure 2: Error covariance averaged over 500 MC runs

Privacy Results I

- We assume that the measurement model has inherent sensor and/or object detection zero mean Gaussian noise ($\boldsymbol{n}_s$).
- We add a synthetic zero mean Gaussian noise ($\boldsymbol{n}_p$) to the image to ensure privacy.
- The noise intensity $\mathbb{E}[\boldsymbol{n}_s \boldsymbol{n}_s^T] = \boldsymbol{R}_s$ is known and $\mathbb{E}[\boldsymbol{n}_p \boldsymbol{n}_p^T] = \boldsymbol{R}_p$ is our design parameter.

Privacy Results II

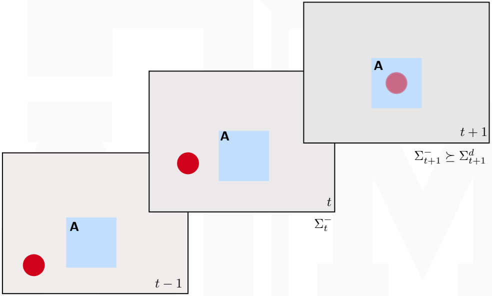

Figure 3: Image frames with privacy in the region **A**

When the tracked red object is in **A** in the $t+1^{\text{th}}$ frame, we want the location estimation error $\mathbf{\Sigma}^-_{t+1} \succeq \mathbf{\Sigma}^d_{t+1}$.

Privacy Results III

- We choose $\boldsymbol{\Sigma}^{d}_{t+1}$ to be $\textbf{diag}([54.891\ 54.891])$ in the pixel frame.
- With $\boldsymbol{\Sigma}^{-}_{t}$, our proposed privacy theorem yields $\boldsymbol{R}_p = \mathbf{I}_2$, with $\boldsymbol{W}$ chosen to be identity. (assuming $\boldsymbol{R}_s = \mathbf{0}$)
- From a data sharing perspective, we would share the image frame at time point $t+1$ with added noise of intensity $\boldsymbol{R}_p$.

Conclusion

- We addressed two questions related to privacy and utility for moving object detection from a video stream using the Kalman filter.
- We modeled them as convex optimization problems based on LMIs.
- The proposed framework was implemented on a numerical problem for two scenarios.
    - First, the purpose was to track an object with an upper bound on estimation error while ensuring utility.
    - Second, we calculated the minimal noise that needs to be injected to a frame to ensure desired privacy prescribed by a lower bound on the localization error of the object.

Acknowledgment

Thank You