GENERALIZED ELL FOR DETECTING AND TRACKING THROUGH ILLUMINATION MODEL CHANGES

*Amit Kale** and Namrata Vaswani***

*Siemens Corporate Technology India, Bangalore , kale.amit@siemens.com
**Dept. of ECE, Iowa State University, Ames, IA 50011, namrata@iastate.edu

ABSTRACT

In previous work, we developed the Illum-PF-MT, which is the PF-MT idea applied to the problem of tracking temporally and spatially varying illumination change. In many practical problems, the rate at which illumination changes varies over time. For e.g. when a car transitions from shadow to sunlight or vice-versa the rate of illumination change is much higher than when it is in shadow or in sunlight. One way to model illumination change in such problems is using a Gaussian random walk model with two values of the change covariance - a large covariance when a "transition" is detected and a much smaller one when "no transition" is detected. But to use such a model, one needs to first detect the transition. The transition is a natural one and so it happens gradually (unlike a sudden manual dimming of the light in the room) and thus existing change detection statistics which are designed only for sudden changes are unable to detect the transition. In this paper, we propose to use the recently proposed generalized ELL (gELL) idea which uses the tracked part of the change to detect it and hence detects such partially trackable changes very quickly. Since gELL detects much before loss of track occurs, one is able to transition to the "transition" model and back without ever losing track. Also, for the first time, we demonstrate the use of gELL in combination with the PF-MT algorithm which is more stable to model change than the original PF.

1. INTRODUCTION

Tracking illumination changes of moving objects is a challenging problem. In absense of illumination changes, motion of a rigid object moving in front of a camera can be tracked using a 3 dimensional vector consisting of x-y translation and uniform scale or more generally using a 6 dimensional affine model as in Condensation [1]. If illumination changes over time, but is constant in space, then one extra dimension gets added. But if different regions of an object experience different lighting conditions (e.g. a face with light falling at different angles on different parts of the face, usually happens when light source is near the object, invalidating assumptions about point light sources at infinity), the maximum dimension of illumination change is equal to the number of image pixels. Of course, the variability is never that large, and as been demonstrated in previous work [2], usually a 3 to 7 dimensional basis suffices for modeling illumination, but even that will increase the total state space dimension to somewhere between 9-16. It is well known that as state space dimension increases, number of particles required to track using a PF increases [3]. This makes PF impractical for dimensions larger than 7 or 8. But, as shown in [4], the conditional posterior of illumination change (conditioned on motion and previous state) is usually unimodal and narrow so that the conditional posterior of illumination can be replaced by a Dirac delta function at its posterior mode, with little error. Furthermore, this mode computation is very efficient, since it turns out to be the solution of a regularized least squares problem. This one step, reduces the importance sampling dimension to 3 instead of 10, drastically reducing the number of particles required. The idea, called Illum PF-MT, was demonstrated in our recent paper [4].Now, in certain problems, the rate at which illumination changes varies over time. For e.g. when a car transitions from shadow to sunlight or vice-versa the rate of illumination change is much higher than when it is in shadow (see the first row of Fig. 1). One good way to model illumination change in such problems is using a Gaussian random walk model with two values of the change covariance - a large covariance (or in effect a weak prior) when a "transition" is detected and a much smaller covariance (learnt from training data) when "no transition" is detected. Note that, since the transition itself from small to large covariance happens gradually (since it is a natural one), even though to keep our modeling simple, we use a single change point to model it. Since the transition is not a sudden one (e.g. as would happen if the light in a room was manually suddenly dimmed to a third of its original value), even without any correction step, the PF-MT algorithm is able to partially track it. Such changes which get partially tracked (are not sudden enough) are usually missed by loss-of-track based statistics such as tracking error [5] or averaged likelihood [6] or score function (see [7] for a survey of sudden change detection methods using particle filters). Note that the tracking error plots in the last columns of Figure 1 miss the change.

But, the ELL statistic [8, 9] was designed for detecting exactly such gradual changes. It uses the partially tracked part of the change to detect it and hence is able to detect gradual transitions much better than existing statistics [5, 6, 7]. In fact, it detects much before loss of track. Now, in problems, such as ours, where the nominal model is nonstationary and has continuously increasing prior state variance, the sensitivity of ELL reduces with time. In this paper, we demonstrate the use of a recently proposed generalization of ELL, called gELL [9] (which was developed for detecting changes in nonstationary nominal models), to detect the changes in the rate of illumination change. Also, unlike ELL, the gELL is able to detect a sequence of changes, for e.g., in our case, the increase the decrease of the change covariance as shown in Fig. 1, last column.

Note that this is the first application where gELL (and not ELL) is used for change detection (so far only one proof-of-concept simulation was shown in [9]). In addition, we successfully demonstrate the use of gELL not only to detect illumination model change, but also to increase illumination change covariance to a large value when the transition is detected, and then reduce it to its original value when gELL again goes below a threshold. Since gELL detects much be-

This research was partially supported by funds from NSF under grant ECCS-0725849

fore loss of track occurs, one is able to transition to the next model and back without ever losing track (see Figure 1). Also, for the first time, we demonstrate the use of gELL in combination with the PF-MT algorithm - in past work [9, 10], ELL was used in combination with only the original PF [11]. This is important because PF-MT (and also some other PFs such as [12]) importance sample using a density that depends on the current observation. For this reason, PF-MT is much more stable to model changes than original PF, i.e. is able to partially track them better than original PF. This fact reduces the delay in the detection using gELL compared to using it with original PF.

2. STATE SPACE MODEL AND THE PROBLEM

We briefly describe below the state space model for illumination and motion change over time. This is taken from our previous work [4] where we introduced the PF-MT algorithm for illumination and motion tracking.

System Model: The state, X_t, consists of a 3-dimensional motion vector u_t which contains x-y translation and scale, and a 7 dimensional illumination coefficients vector (illumination is parameterized using a Legendre basis) as in [2], i.e. $X_t = [u_t' \ \ \Lambda_t']'$ The system model is a random walk model on object motion, u_t and on illumination coefficients, Λ_t i.e.

$$u_{t+1} = u_t + \nu_{ut}, \ \ \nu_{ut} \sim h(.) \quad (1)$$
$$\Lambda_{t+1} = \Lambda_t + \nu_{\Lambda t}, \ \ \nu_{\Lambda t} \sim \mathcal{N}(0, \Pi) \quad (2)$$

where $\Pi_{N_\Lambda \times N_\Lambda}$ is a diagonal covariance matrix (variance of individual components of Λ) and $h(.)$ denotes the pdf of ν_{ut} described in [4].

Observation Model: Let T_0 denote the original template and let M denote the number of pixels in it. The observation at time t, Y_t, is the image at t. It assumes the following image formation process: the image intensities of the region that contains the object, are illumination scaled versions of the intensities of the original template, T_0, plus Gaussian noise. The region containing the object is the original template region scaled and translated using the current elements of u_t. The rest of the image (which does not contain the object) is independent of the object intensity or "shape", and hence is not used in defining the observation likelihood. Thus we have the following observation model:

$$Y_t \left(\mathbf{J} u_t + \begin{bmatrix} \mathbf{X}_0 \\ \mathbf{Y}_0 \end{bmatrix} \right) = [f_{T_0}(\mathbf{P}\Lambda_t)]_{vec} + \psi_t, \ \ \psi_t \sim \mathcal{N}(0, V) \quad (3)$$

where the notation $[.]_{vec}$ denotes arranging a two dimensional matrix as a column vector; $(V)_{M \times M}$ is a diagonal covariance matrix (variance of individual pixel noise); $\mathbf{P}$ contains the Legendre basis directions as its column vectors (defined in (5) of [4]) and

$$f_{T_0}(P\Lambda_t) \triangleq T_0 + T_0 .* \mathbf{P}\Lambda_t,$$
$$\mathbf{J} \triangleq \begin{bmatrix} \mathbf{X}_0 - \bar{x}_0 & \mathbf{1} & \mathbf{0} \\ \mathbf{Y}_0 - \bar{y}_0 & \mathbf{0} & \mathbf{1} \end{bmatrix} \quad (4)$$

where $.*$ is the MATLAB notation, $\mathbf{X}_0$ and $\mathbf{Y}_0$ denote the x and y coordinates of each point on the template and $\bar{x}_0$ and $\bar{y}_0$ denote the corresponding means. $\mathbf{1}$ and $\mathbf{0}$ denote a vector of ones and zeros of size M respectively. Thus the observation likelihood (OL) is:

$$p(Y_t|X_t) = p(Y_t|u_t, \Lambda_t) = e^{-||[G_t^{u_t} - f_{T_0}(\mathbf{P}\Lambda_t)]_{vec}||_V^2} \quad (5)$$

where $||a||_V \triangleq a^T V^{-1} a$ for a vector a and

$$G_t^{u_t} \triangleq Y_t \left(\mathbf{J} u_t + \begin{bmatrix} \mathbf{X}_0 \\ \mathbf{Y}_0 \end{bmatrix} \right) \quad (6)$$

The PF-MT algorithm for tracking using the above state space model, when the illumination change covariance, Π, is a constant, was proposed in [4]. It importance sampled on motion (since it had a large variance and multimodal state transition prior and since it often resulting in a multimodal observation likelihood), while mode tracking on illumination (whose change covariance was much smaller and the observation likelihood was mostly unimodal conditioned on motion). It is summarized in the first few steps of Algorithm 1.

In the current work, we consider the problem where the model of (2) can change with time. In particular, Π can take two possible values (small and large) and the time when the transition between them happens is unknown. The goal is to detect when to change Π from small to large (for the shadow-light transition frames) and when to change it back. The value Π_{small} is known (learnt from training data), but the value Π_{large} for the transition frames is not known.

3. DETECTING AND CHANGING THE SYSTEM MODEL

In many tracking applications, the system model parameters are not time-invariant. For our problem, consider the random walk model on illumination coefficients given in (2). As explained in the introduction, the rate of change of illumination over time (quantified by the illumination change covariance) is much larger when the car transitions from a shadowy region to a bright/sunlit region or vice versa than when it is in a shadowy or bright region. One good model for this situation is (2) with a small noise covariance value when the car is in the shadowy or the sunlit region, but a large noise covariance value when it transitions from shadow to sunlight or vice versa.

3.1. Computing Generalized-ELL (gELL) and gELL-max

To use the above model for tracking, one first needs to be able to detect the change time (the time when the illumination change covariance needs to be increased or reduced), as quickly as possible. To avoid having to re-initialize the tracker, one would like to detect this change before significant loss-of-track occurs. In [9], the Expected (negative) Log-Likelihood of state (ELL) statistic was introduced to detect changes before they resulted in significant loss-of-track. The key idea was to use the "tracked part of the change" to detect it. Thus ELL requires the change to be partially tracked in order to detect it and it often does not detect very sudden changes that result in immediate loss of track. Such sudden statistics do not occur in our problem, but if they do, tracking error [5] or averaged likelihood [6] can be used to also detect them in combination with ELL.

ELL can be interpreted as the Kerridge Inaccuracy (proportional to Kullback-Leibler divergence) between the posterior at the current time, $\pi_{t|t}(X_t) = p(X_t|Y_{1:t})$, and the prior state distribution at t, which is equal to the t step ahead prediction distribution, $\pi_{t|0}(X_t) = p(X_t)$. As explained in [9], ELL cannot detect multiple changes in a sequence and its sensitivity reduces with time in many problems such as ours where the nominal model is nonstationary (because the variance of $p(X_t)$ increases with t). To handle this, a generalization of ELL was defined in [9]. Generalized ELL (gELL) is the Kerridge Inaccuracy between $\pi_{t|t}$ and the $\Delta < t$ step ahead prediction distribution, $\pi_{t|t-\Delta}(X_t) = p(X_t|Y_{1:t-\Delta})$, i.e.

$$gELL(t, \Delta) \triangleq -\mathbb{E}_{\pi_{t|t}}[-\log \pi_{t|t-\Delta}(X_t)] \quad (7)$$

Note that gELL and ELL may be computed for the entire state X_t or for a part of it. In our problem, we need to detect changes in illumination and hence we compute gELL, defined in (7), only for the posterior of Λ_t. Note, the same idea can also be used to detect changes in x or y direction velocity of the object (in that case we would define gELL for only for posterior of u_t). To compute the gELL, we need a closed form expression for $\pi_{t|t-\Delta}$. To get that, we propose to approximate the PF estimate of the posterior at $t-\Delta$, $\pi_{t-\Delta|t-\Delta}(X_t)$, by a Gaussian density, i.e. $\pi^N_{t-\Delta|t-\Delta}(X_t) \approx \mathcal{N}(\mu^N_{t-\Delta|t-\Delta}, \Sigma^N_{t-\Delta|t-\Delta})$ where the parameters are estimated as the empirical mean and covariance of the weighted particle set comprising of $\pi^N_{t-\Delta|t-\Delta}(X_t)$. With this approximation, the prediction distribution, $\pi_{t|t-\Delta}(X_t)$, which is obtained by applying the system model of Λ_t given in (2) Δ times to $\pi_{t-\Delta|t-\Delta}(X_t)$, is also Gaussian, i.e. $\pi_{t|t-\Delta}(X_t) \approx \mathcal{N}(\mu^N_{t|t-\Delta}, \Sigma^N_{t|t-\Delta})$ with parameters defined below. Thus, in summary, gELL is computed as:

$$
\begin{aligned}
gELL(t,\Delta) &\triangleq \sum_{i=1}^{N} w_t^i (\Lambda_t^i - \mu^N_{t|t-\Delta})^T {\Sigma^N_{t|t-\Delta}}^{-1} (\Lambda_t^i - \mu^N_{t|t-\Delta}), \\
\mu^N_{t|t-\Delta} &= \mu^N_{t-\Delta|t-\Delta} \triangleq \sum_{i=1}^{N} w^i_{t-\Delta} \Lambda^i_{t-\Delta} \\
\Sigma^N_{t|t-\Delta} &\triangleq \Sigma^N_{t-\Delta|t-\Delta} + \Delta\Pi, \\
\Sigma^N_{t-\Delta|t-\Delta} &\triangleq \sum_{i=1}^{N} w^i_{t-\Delta} (\Lambda^i_{t-\Delta} - \mu^N_{t|t-\Delta})(\Lambda^i_{t-\Delta} - \mu^N_{t|t-\Delta})^T
\end{aligned}
\quad (8)
$$

The choice of Δ in the above expression is not clear. If it is too small, the change between time t and $t-\Delta$ may not be large enough, i.e. the numerator, $(\Lambda_t^i - \mu^N_{t|t-\Delta})$, may be too small. If it is too large, the prediction covariance, $\Sigma^N_{t|t-\Delta} = \Sigma^N_{t-\Delta|t-\Delta} + \Delta\Pi$ may be larger than needed, thus reducing its sensitivity to smaller changes (a problem similar to that of ELL which uses $\Delta = t$). Thus, a statistic that is always more sensitive than $gELL(t,\Delta)$ (i.e. its detection delay is smaller than or equal to that of $gELL(t,\Delta)$) is

$$
gELL\text{-}max(t) \triangleq \max_{\Delta=1,2,\ldots t} gELL(t,\Delta) \quad (9)
$$

Of course it may also generate a few extra false alarms. In Sec. 4, we show experiments with both $gELL(t,\Delta)$, for the car sequence which is faster moving and thus has faster rate of change of illumination covariance, and with the more sensitive, $gELL\text{-}max(t)$, for the face sequence in which the changes are slower.

3.2. Using gELL or gELL-max to Change System Model

We begin by tracking using PF-MT that uses (2) with a small covariance $\Pi = \Pi_{small}$ (learnt only from the shadow sequence) and we keep computing the $\Delta = 5$ step ahead gELL (or gELL-max) at every t. gELL (or gELL-max) exceeding its detection threshold is used as a cue to increase the value of Π to Π_{large}. Π_{large} can heuristically set to a large value (or even to ∞ to allow PF-MT to only use the observations) or if enough training data for the transition frames is available, it can be learnt from it. A large value of Π models a weak prior, i.e. the tracker mostly follows the changing observations. It uses these observations to latch on to the sunlight illumination. When the car has fully transitioned to the sunlight region, the value of gELL falls below its threshold, and this is used as a cue to again reduce Π to Π_{small}. The complete algorithm in summarized in Algorithm 1.

Algorithm 1 Change Compensated Aux PF-MT.

At each t, do

1. *Auxiliary Resampling:* $\forall i$, compute g_t^i using $g_t^i = w^i_{t-1} p(Y_t|X_t = X^i_{t-1})$ and resample X^i_{t-1} according to it. Reset the weights of the resampled particle to $(w^i_{t-1})^{new} = \frac{w^i_{t-1}}{N g_t^i} = \frac{p(Y_t|X_t = X^i_{t-1})}{N}$.
2. *Importance Sample (IS) on effective basis:* $\forall i$, sample ${\nu_u}_t^i \sim h(u)$ and compute $u_t^i = u^i_{t-1} + {\nu_u}_t^i$.
3. *Mode Tracking (MT) in residual space:* $\forall i$, compute m_t^i using $m_t^i = \Lambda^i_{t-1} + (\Pi^{-1} + A^T_{T_0} V^{-1} A_{T_0})^{-1} A^T_{T_0} V^{-1} D$ $D \triangleq [G_t^{u_t^i}]_{vec} - f_{T_0}(\mathbf{P}\Lambda^i_{t-1})$ and set $\Lambda_t^i = m_t^i$.
4. *Weighting:* Compute w_t^i using $w_t^i = \frac{\tilde{w}_t^i}{\sum_{j=1}^N \tilde{w}_t^j}$, $\tilde{w}_t^i = p(Y_t|u_t^i, \Lambda_t^i) p(\Lambda_t^i|\Lambda^i_{t-1})$.
5. *gELL Computation:* Compute gELL using (8) or $gELL\text{-}max(t)$ using (9).
6. *Change* Π*:* If gELL exceeds threshold, set $\Pi = \Pi_{large}$, when it goes below threshold set $\Pi = \Pi_{learnt}$.

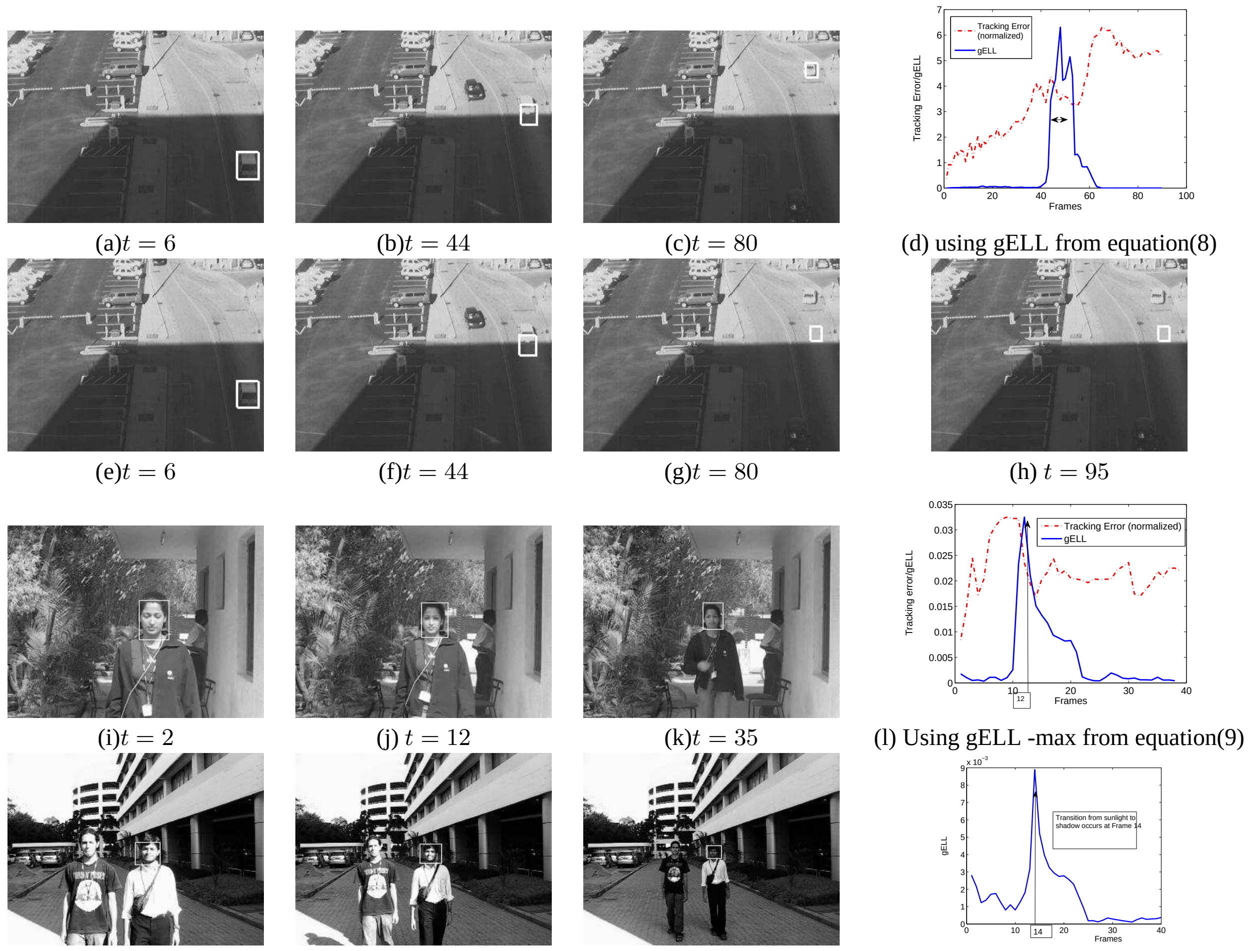
4. EXPERIMENTAL RESULTS

We now demonstrate the utility of the proposed approach for two different datasets. The car dataset was generated from a camera observing a road from above as cars approach an intersection and move in and out of shadow. The second dataset contained several subjects moving through different illuminations in an outdoor environment.

In Figure 1 we show the results of using gELL and PFMT using 100 particles. Around frame 40, when the car starts to move from shadow to sunlight (as indicated by the double arrow in Figure **??**(d)) the gELL value starts to increase from its shadow value. When it does we set $\Pi = \Pi_{large}$ in Algorithm1. We use $\Pi_{large} = \infty$. When gELL decreases again, we reset Π to Π_{learnt}. The tracking is shown in first row of Figure 1. If we do not use gELL to detect the transitions and increase Π, the tracker fails(Figure 1e-h). We also show the use of normalized tracking error (normalized by ratio of peaks of actual tracking error to gELL) for change detection in Figure 1(d). As can be seen, tracking error does not show any sharp change around frame 40 unlike gELL whose change is clearly detectable. This makes the task of switching Π to Π_{large} difficult leading to loss of track. For the face tracking case Figure 1 i-p, we tried the use of both gELL(8) and gELL-max (9). gELL max works better for this case since the rate of change of illumination are slower than in the face case (See Section 3.1). The changeover from changeover from sunlight to shadow (arrows indicate frame of changeover) are again detected accurately. However in this case using PF-MT succeeds even across illumination changes.

5. CONCLUSION

In this paper, we proposed to use the recently proposed generalized ELL (gELL) idea which uses the tracked part of the change to detect it and hence detects such partially trackable changes very quickly. Since gELL detects the illumination change before loss of track occurs, one is able to transition to the “transition” model and back without ever losing track. Furthermore we demonstrate the use of gELL in combination with the Illumination PF-MT algorithm [4] which is more stable to model change than the original PF. We demonstrated



(a) $t = 6$ (b) $t = 44$ (c) $t = 80$ (d) using gELL from equation(8)

(e) $t = 6$ (f) $t = 44$ (g) $t = 80$ (h) $t = 95$

(i) $t = 2$ (j) $t = 12$ (k) $t = 35$ (l) Using gELL -max from equation(9)

(m) $t = 2$ (n) $t = 14$ (o) $t = 40$ (p) Using gELL -max from equation(9)

Fig. 1. Tracking using PF-MT and gELL as objects move through different lighting conditions. The white box corresponds to the MAP estimate of the "shape vector". (a) (b) (c) show tracking of a car (d) shows the comparison of the tracking error with gELL. The second row shows the tracking of the car as it moves from shadow to sunlight when gELL is not used to detect change and PF-MT is left unaltered. Bottom two rows show the face tracking cases where gELL max detects illum changes correctly

the algorithm for tracking faces and cars across drastic illumination changes.

6. REFERENCES

[1] M. Isard and A. Blake, "Condensation: Conditional Density Propagation for Visual Tracking," *Intl. Journal of Comp. Vision*, pp. 5–28, 1998.

[2] A. Kale and C. Jaynes, "A joint illumination and shape model for visual tracking," *Proceedings of IEEE CVPR*, pp. 602–609, 2006.

[3] A. Doucet, N. deFreitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer, 2001.

[4] A. Kale, N. Vaswani, and C. Jaynes, "Particle filter with mode tracker (pf-mt) for visual tracking across illumination change," in *ICASSP*, 2007.

[5] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, Academic Press, 1988.

[6] Q. Zhang, F. Campillo, F. Cerou, and F. LeGland, "Nonlinear system fault detection and isolation based on bootstrap particle filters," December.

[7] C. Andrieu, A. Doucet, S.S. Singh, and V.B. Tadic, "Particle methods for change detection, system identification, and control," *Proceedings of the IEEE*, vol. 93, pp. 423– 438, March 2004.

[8] N. Vaswani, "Change detection in partially observed nonlinear dynamic systems with unknown change parameters," in *American Control Conference (ACC)*, 2004.

[9] N. Vaswani, "Additive change detection in nonlinear systems with unknown change parameters," pp. 859–872, March 2007.

[10] N. Vaswani, A. RoyChowdhury, and R. Chellappa, ""Shape Activity": A Continuous State HMM for Moving/Deforming Shapes with Application to Abnormal Activity Detection," *IEEE Trans. Image Proc.*, pp. 1603–1616, October 2005.

[11] N.J. Gordon, D.J. Salmond, and A.F.M. Smith, "Novel approach to nonlinear/nongaussian bayesian state estimation," *IEE Proceedings-F (Radar and Signal Processing)*, pp. 140(2):107–113, 1993.

[12] A. Doucet, "On sequential monte carlo sampling methods for bayesian filtering," in *Technical Report CUED/F-INFENG/TR. 310, Cambridge University Department of Engineering*, 1998.