Tracking (Optimal filtering) on Large Dimensional State Spaces (LDSS)

Namrata Vaswani

Dept of Electrical & Computer Engineering

Iowa State University

http://www.ece.iastate.edu/~namrata

IOWA STATE UNIVERSITY
Becoming the best.

HMM Model & Tracking

- Hidden State sequence X_t, Observations Y_t
  - $\{X_t\}$ is a Markov chain
  - $X_t \rightarrow Y_t$ is a Markov chain at each t
  - $p(X_t|X_{t-1})$: state transition prior (STP) : known
  - $p(Y_t|X_t)$: observation likelihood (OL) : known

- Tracking (Optimal filtering): Get the "optimal" estimate of X_t based on observations $Y_{1:t}$ (causally)
  - Compute/approx the posterior, $\pi_t(X_t) = p(X_t|Y_{1:t})$
  - Use π_t to compute any "optimal" state estimate

LDSS Tracking Problems

- Image Sequences
  - – Boundary contour of a moving/deforming object
  - – Rigid motion & Illumination variation (over space & time)
  - – Optical flow (motion of each pixel)
- Sensor Networks
  - – Spatially varying physical quantities, e.g. temperature
  - – Boundary of a chemical spill or target emissions
- Time-varying system transfer functions
  - – Time-varying STRF: repr. for neuronal transfer functions
  - – Time varying AR model for speech (e.g STV-PARCOR)

Problem Setup

- Observation Likelihood (OL) is often multimodal
  - e.g. clutter, occlusions, low contrast images
  - e.g. some sensors fail or nonlinear sensors
  - If STP narrow, posterior unimodal: can adapt KF, EKF
  - If STP broad (fast changing sequence): require a Particle Filter (PF)

- Large dimensional state space (LDSS)
  - e.g. deformable contour tracking
  - e.g. tracking temperature in a large area
  - PF expensive: requires impractically large N

Temperature tracking: bimodal OL

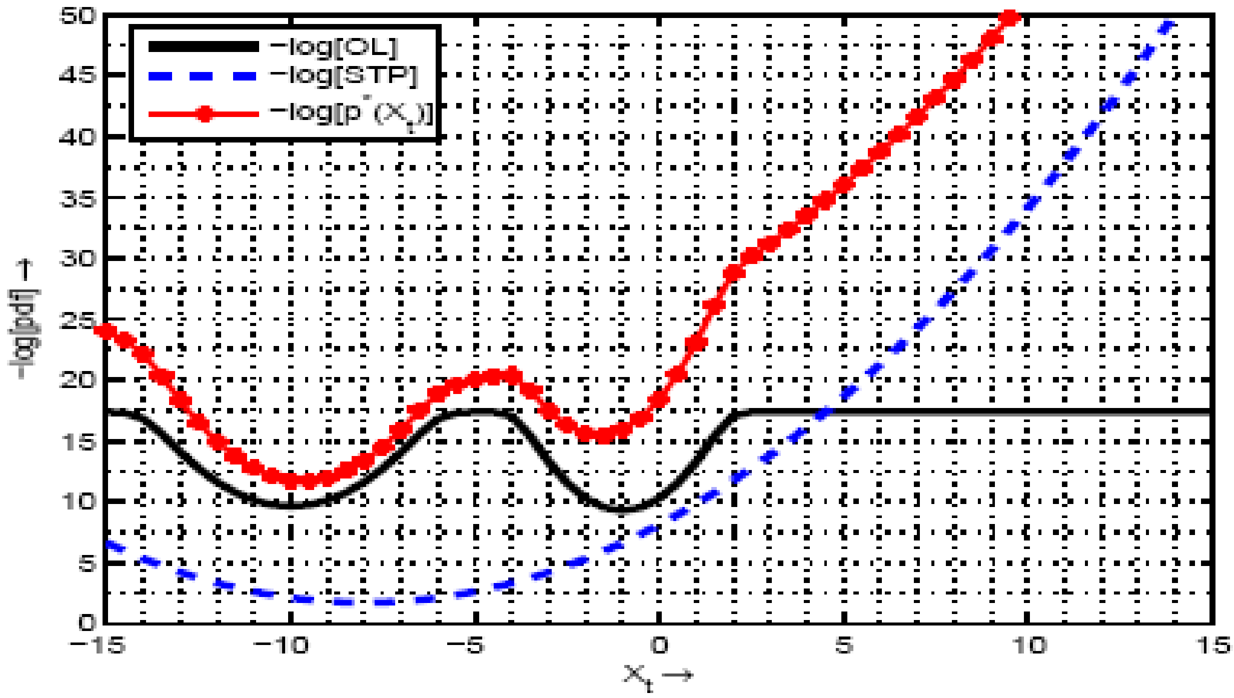
- Nonlinear sensor (measures square of temp.)

$Y_t = X_t^2 + w_t, \qquad w_t \sim N(0,\sigma^2)$

- Whenever $Y_t>0$, $p(Y_t|X_t)$ is bimodal as a function of X_t with modes at $X_t = Y_t^{1/2}$ & $X_t = -Y_t^{1/2}$

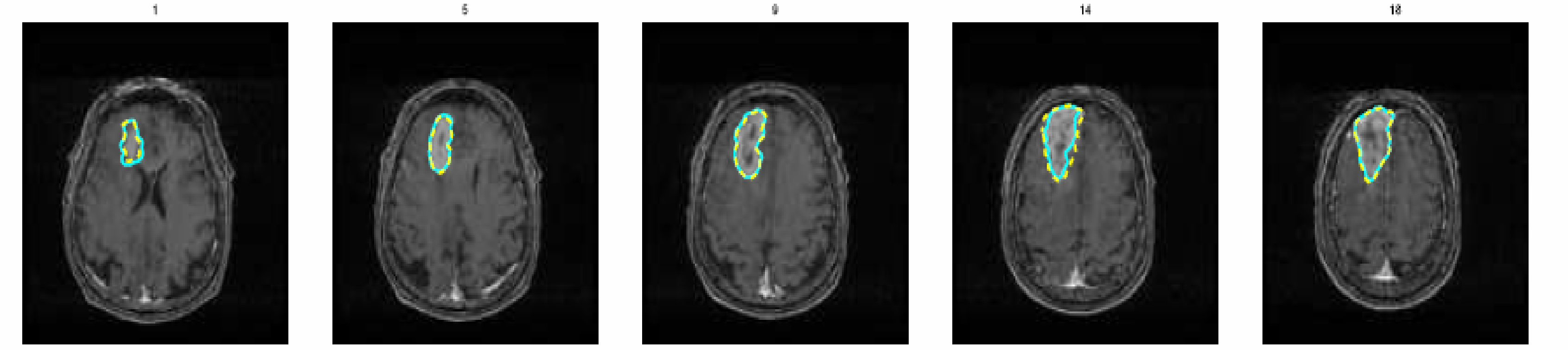
Temperature tracking: bimodal OL

Temperature measured with 2 sensors, each with some probability of failure. Bimodal OL if one of them fails. Bimodal posterior when STP broad

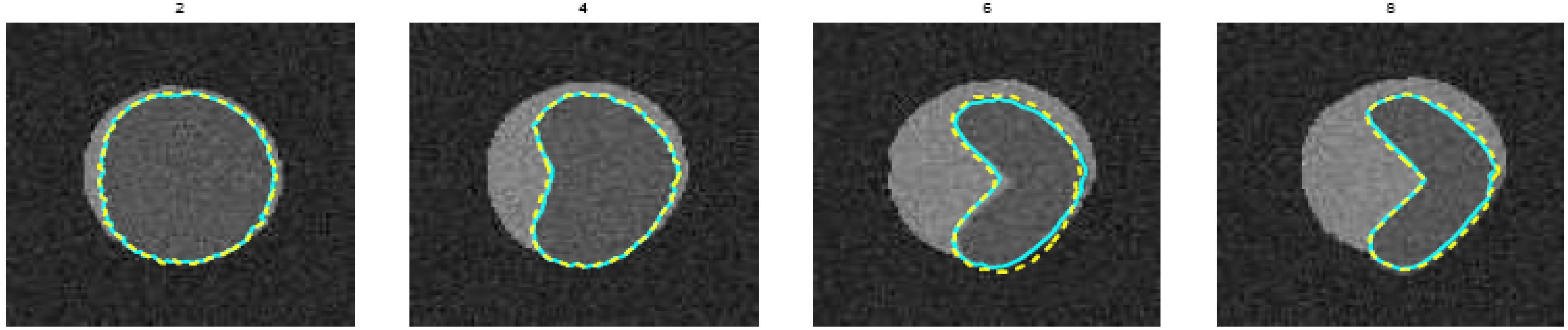


Contour tracking: multimodal OL

Low contrast images (tumor region in brain MRI)



Overlapping background clutter



Particle Filter [GSS'93]

- Sequential Monte Carlo technique to approx Bayes' recursion for computing the posterior $\pi_t(X_{1:t}) = p(X_{1:t}|Y_{1:t})$
  - Approx approaches true posterior as the # of M.C. samples ("particles")→ ∞, for a large class of nonlinear/non-Gaussian problems

- Does this sequentially at each t using Sequential Importance Sampling along with a Resampling step (to throw away particles with very small importance weights)

Monte Carlo, Importance Sampling

- Goal: compute $E_p [\gamma(X)] = \int_x \gamma(x)\, p(x)\, dx$

(compute expected value of any function of X)

- Monte Carlo:

$$E_p [\gamma(X)] = \int_x \gamma(x)\, p(x)\, dx \approx (1/N) \sum_i \gamma(X^i), \quad X^i \sim p$$

- Imp Sampling: If cannot sample from p,

$$E_p [\gamma(X)] = E_q [\gamma(x)\, p(x)/q(x)] \approx (1/N) \sum_i \gamma(X^i)\, p(X^i) / q(X^i), \quad X^i \sim q$$

Bayesian Importance Sampling

- Goal: compute $E[\gamma(X)|Y] = E_{p(X|Y)}[\gamma(X)]$
(compute posterior expectation of any function of X)

- Choose Imp Samp. density q_Y & rewrite above as

$$E_{p(X|Y)}[\gamma(X)] = N/D$$

$$N = E_{q_Y}[\gamma(X)\, p(Y|X)\, p(X) / q_Y(X)]$$

$$D = E_{q_Y}[p(Y|X)\, p(X) / q_Y(X)]$$

  - Imp Sample: $X^i \sim q_Y$
  - Weight: $w^i \propto p(Y|X^i)\, p(X^i) / q_Y(X^i)$
  - Posterior, $p(X|Y) \approx \sum_i w^i \delta(X - X^i)$

$$E_{p(X|Y)}[\gamma(X)] \approx \sum_i \gamma(X^i)\, w^i$$

Particle Filter: Seq. Imp Sampling

- Sequential Imp Sampling for HMM model
  - Replace Y by $Y_{1:t}$, Replace X by $X_{1:t}$
  - Choose Imp Sampling density s.t. it factorizes as

$q_{t,Y_{1:t}}(X_{1:t}) = q_{t-1,Y_{1:t-1}}(X_{1:t-1})\ q_{X_{t-1},Y_t}(X_t)$

  - Allows for recursive computation of weights

- Seq Imp Sampling: At each t, for each i,
  - Importance Sample: $X_t^i \sim q_{X_{t-1}^i,Y_t}(X_t)$
  - Weight: $w_t^i \propto w_{t-1}^i\ p(Y_t|X_t^i)\ p(X_t^i|X_{t-1}^i) / q_{X_{t-1}^i,Y_t}(X_t^i)$
  - Posterior, $\pi_t(X_{1:t}) \approx \pi_t^N(X_{1:t}) = \sum_i w_t^i\ \delta(X_{1:t} - X_{1:t}^i)$

Particle Filter: Resampling step

- Seq IS: gives a weighted delta function estimate of posterior: $\pi_t^N(X_{1:t}) = \sum_i w_t^i \, \delta(X_{1:t} - X_{1:t}^i)$
  - With just Seq IS, as t increases, most weights become very small (particles "wasted")
- "Resample": Sample N times from π_t^N to get an equally weighted delta function estimate of posterior: $\pi_t^{N,new}(X_t) = \sum_i (1/N) \, \delta(X_t - X_t^{i,new})$
  - Effect: High weight particles repeated multiple times, very low weight ones thrown away

Outline

- Goal, Existing Work & Key Ideas
- Proposed algorithms: PF-EIS, PF-EIS-MT
- Open Issues
- Applications
- Ongoing Work (System Id)

Goal, Existing Work & Key Ideas

Our Goal

- Design efficient importance sampling techniques for PF, when
  - OL is multimodal & STP is broad

  and/or

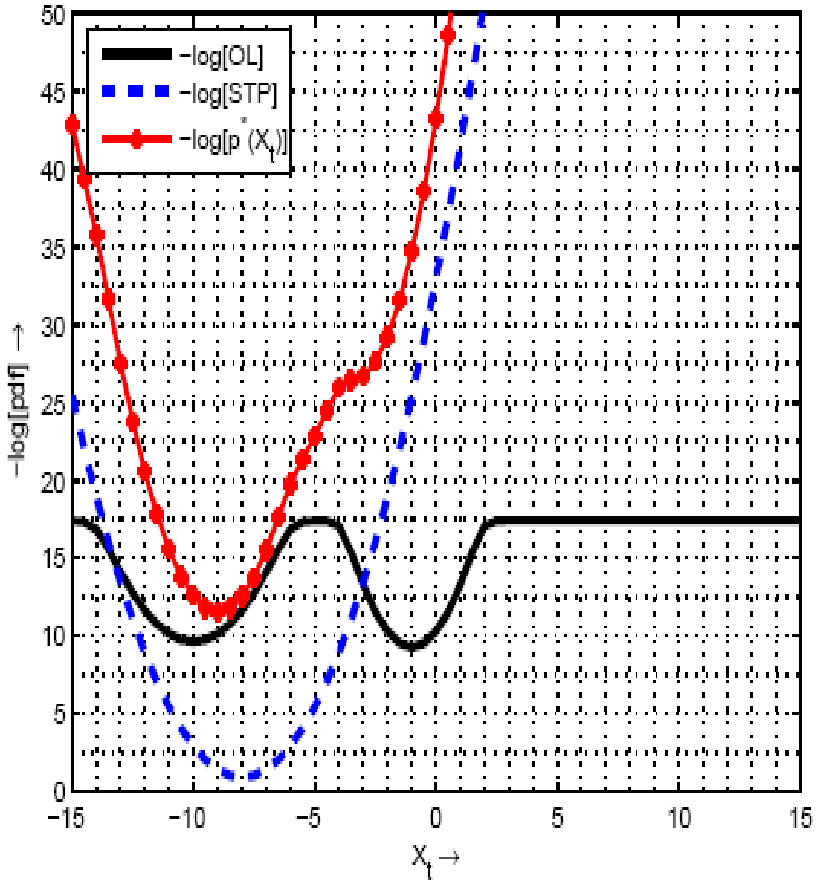
  - Large dimensional state space (LDSS)

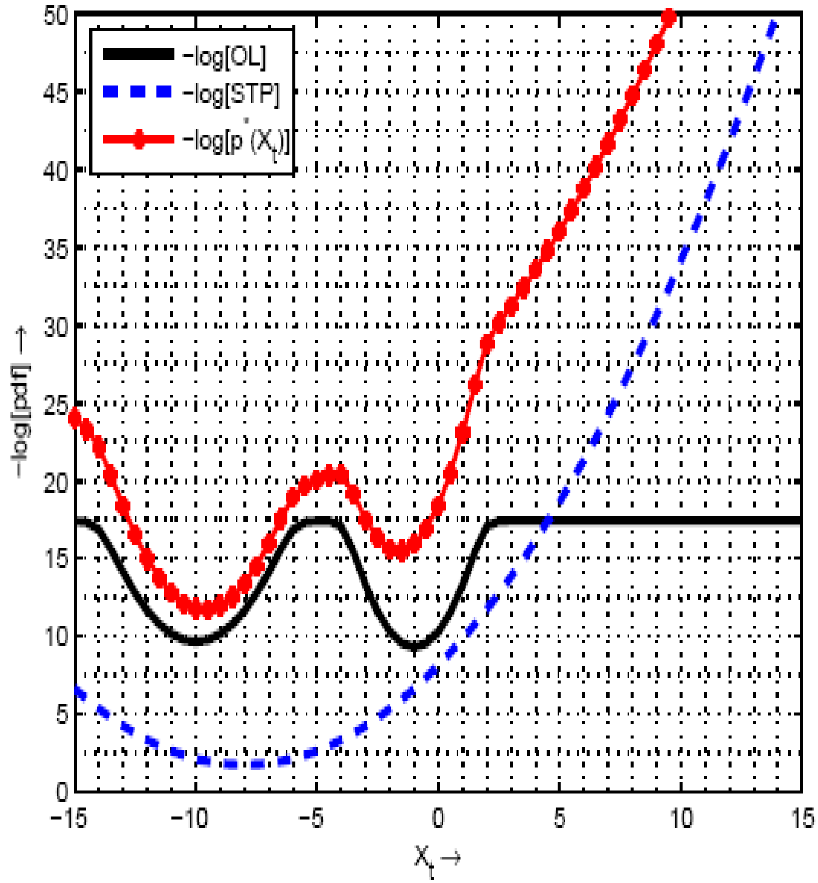
OL multimodal & STP broad

- "OL multimodal": $p(Y_t|X_t)$ has multiple local maxima as a function of X_t
- If OL multimodal but STP narrow, posterior given previous state (p*) is unimodal
  - If the posterior is also unimodal: can adapt KF or Posterior Mode Trackers
  - Efficient importance sampling methods for PF exist
- If OL multimodal & STP broad: p* multimodal
  - Original PF (sample from STP): but inefficient
  - ?

Narrow STP: Unimodal p*

Broad STP: Multimodal p*





Temperature measured with 2 types of sensors, each with some failure prob

Existing Work

- No assumptions reqd, but inefficient
  - PF-Original: Imp Sample from STP [GSS'93]
- Optimal Imp Sampling density: $p^* = p(X_t | X_{t-1}, Y_t)$
  - Cannot be computed in closed form most cases [D'98]
- When p* is unimodal
  - KF/EKF/UKF, PMT: Use OL mode nearest to predicted state as new observation [BIR,CDC'94][TZ'92][JYS, CDC'04]
  - **PF-D: Imp Sample from Gaussian approx to p*** [D'98]
  - PF-EKF/UPF: UKF/EKF to approx to p* [VDDW,NIPS'01]
- Number of OL modes small & known
  - MHT, IMM, Gaussian Sum PF, …

Large Dim State Space (LDSS)

- As dimension increases, N required for accurate tracking also increases (effective particle size reduces)

- Regular PF impractical for > 10-12 dims
  - Resample multiple times within a time interval [MB,ICCV'98][OM, MCIP, Ch 13]: increases particle degeneracy
  - If large part of state space conditionally Linear Gaussian or can be vector quantized into a few discrete centroids: RB-PF [CL,JASA'00], [SGN,TSP'05]
  - If not, ?

Key Idea 1: "LDSS property"

- State space dim may be large, but in most cases,
  - At any given time, most of the state change occurs in a small # of dims (effective basis) while the state change in the rest of the dims (residual space) is small
    - Different from dimension reduction, which is useful only if state sequence is stationary
  - Effective basis dimension can change with time

Key Idea 2: "Unimodality"

- If residual state change small (residual STP narrow) enough compared to distance b/w OL modes: "residual posterior" (p**) is unimodal
  - p** = posterior of residual state conditioned on previous state & effective basis states
  - p** = p* conditioned on effective basis states

- If p** is unimodal, modify PF-D:
  - Imp Sample effective basis states from STP
  - Imp Sample residual states from Gaussian approx to p** about its unique mode

Key Idea 3: "IS-MT"

- If residual state change still smaller (residual STP still narrower), the residual posterior is unimodal & also narrow
  - Above usually true for a subset of residual space

- If an imp sampling (IS) density is unimodal & narrow, any sample from it is close to its mode with high probability
  - A valid approx: just use its mode as the sample: Mode Tracking (MT) approx of IS or IS-MT

PF with Efficient Imp Sampling (PF-EIS)

PF-Efficient IS (PF-EIS)

[Vaswani, ICASSP'07, 06]

- LDSS problems with multimodal OL
  - # of OL modes large: MHT, GSPF impractical
  - STP broad in at least some dims: p* multimodal
  - Can Imp Sample from STP: inefficient
- But, LDSS property (residual STP narrow)
  - Choose $X_{t,s}$ (effective basis states) & $X_{t,r}$ (residual states) s.t. p** (p* conditioned on $X_{t,s}$) is unimodal
    - Imp Sample $X_{t,s}$ from its STP
    - Imp Sample $X_{t,r}$ from Gaussian approx of p**

Sufficient Conditions for PF-EIS (Unimodality of p**) [Vaswani, ICASSP'07]

- STP($X_{t,r}$) log-concave

  and

- Predicted $X_{t,r}$ close enough to a mode of OL($X_{t,s}^i$,$X_{t,r}$) s.t. its –log is locally convex around it. Denote this mode: X_r^*

  and

- STP($X_{t,r}$) narrow "enough", i.e. its maximum variance smaller than an upper bound, Δ^*
  - Δ^* increases as distance of X_r^* to next nearest OL mode increases or as strength of that mode decreases
  - $\Delta^* = \infty$ if OL($X_{t,s}^i$,$X_{t,r}$) is log-concave

Expression for Δ^* [Vaswani, ICASSP'07]

$$\max_p \Delta_{r,p} < \inf_{\cap_{p=1}^{M_r}(\mathcal{A}_p \cup \mathcal{Z}_p)} \max_p \gamma_p(X_{t,r}) \triangleq \Delta^*$$

$$\gamma_p(X_{t,r}) \triangleq \begin{cases} \frac{|[\nabla D_{num}]_p|}{\epsilon_0 + |[\nabla E]_p|}, & if \quad X_{t,r} \in \mathcal{A}_p \\ \frac{|[\nabla D_{num}]_p|}{\epsilon_0 - |[\nabla E]_p|}, & if \quad X_{t,r} \in \mathcal{Z}_p \end{cases}$$

$E(X_{t,r}) = -\log OL(X_{t,s}^i, X_{t,r})$

f_r^i = predicted $X_{t,r}$

$$\nabla D_{num} \triangleq X_{t,r} - f_r^i$$

R_{LC}: largest convex region in neighborhood of f_r^i where E is convex

$$\mathcal{Z}_p \triangleq \{X_{t,r} \in \mathcal{R}_{LC}^c : [\nabla E]_p.[\nabla D]_p \geq 0 \ \& \ |[\nabla E]_p| < \epsilon_0\}$$

$$\mathcal{A}_p \triangleq \{X_{t,r} \in \mathcal{R}_{LC}^c : [\nabla D_{num}]_p.[\nabla E]_p < 0\}$$

Temperature tracking

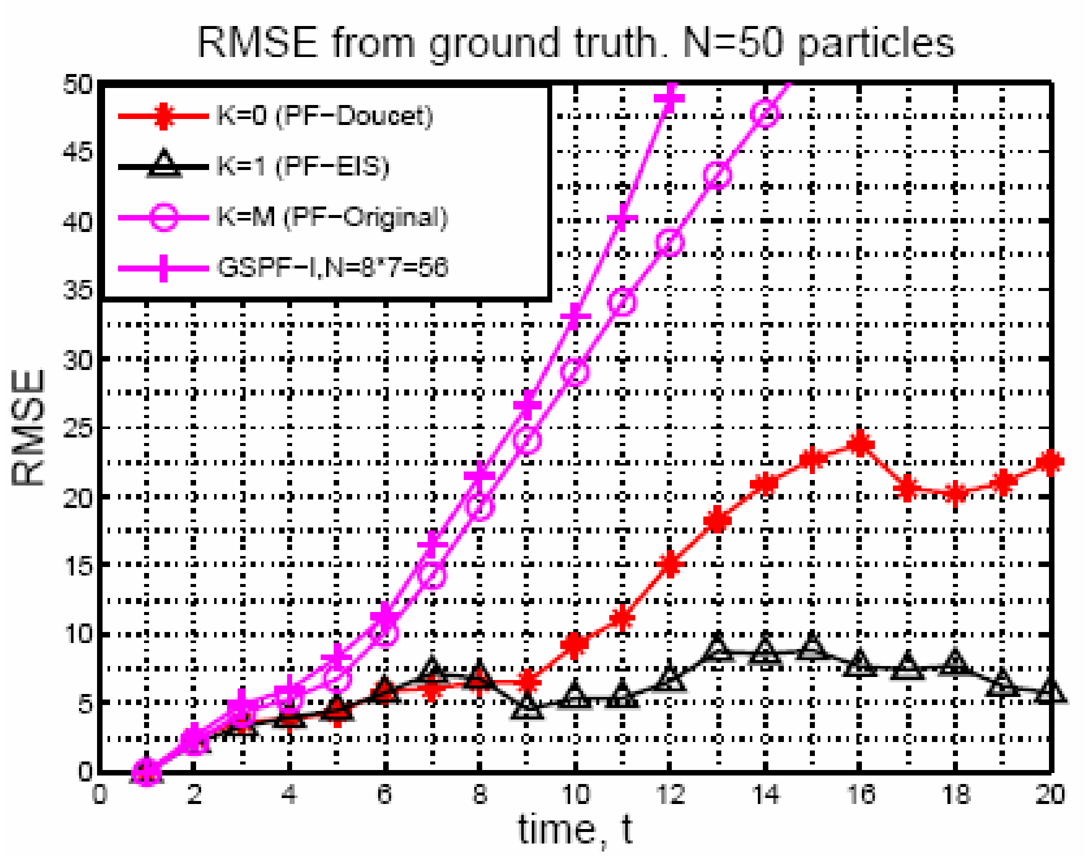
- Track temperature at a set of nodes in a large area using a network of sensors
- STP: state, $X_t = [C_t, v_t]$
  - Temp change, v_t, spatially correlated & follows a Gauss-Markov model. Temp, $C_t = C_{t-1} + v_t$
- OL: observation, Y_t = sensor measurements
  - Diff. sensor meas. independent given actual temp (C_t)
  - Working sensor: meas. C_t corrupted by Gaussian noise
  - With some small probability, any sensor can fail
  - OL multimodal w.r.t. temp at a node, if its sensor fails

Applying PF-EIS: Choosing $X_{t,s}$

Choose $X_{t,s}$ s.t. p** most likely to be unimodal

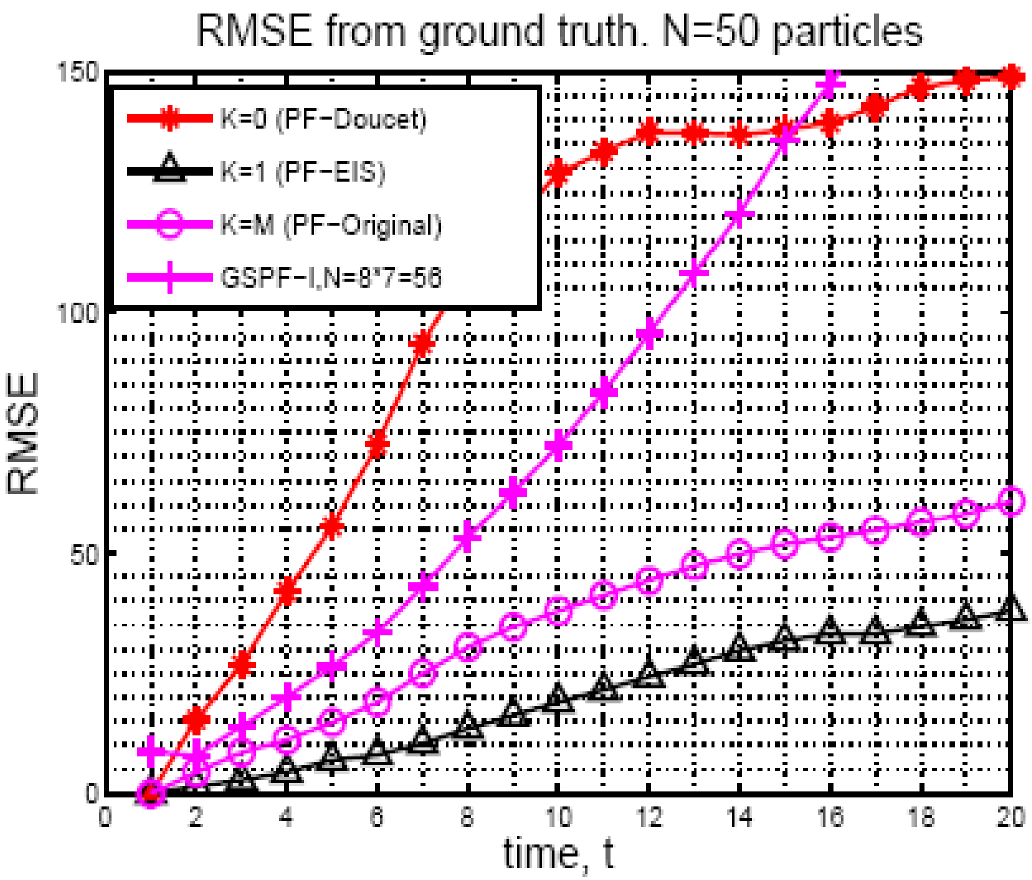
- Get eigen-decomp^n of covariance of temp change
- If a node has older sensors (higher failure prob) than other nodes: choose temp change along eigen-directions most strongly correlated to temp at this node as $X_{t,s}$
- If all sensors have equal failure prob: choose coeff. along the K eigen-directions with highest eigenvalues as $X_{t,s}$

Simulation Results: Sensor failure



- Tracking temp at M=3 sensor nodes, each with 2 sensors. Node 1 has much higher failure prob than rest.
- PF-EIS uses K=1 dim effective basis
- PF-EIS (K=1) outperforms PF-D (K=0), PF-Original (K=3) & GSPF

Simulation Results: Nonlinear sensor



- Tracking temp at M=3 nodes, each with 1 sensor per node
- Node 1 has a squared sensor (measures square of temp plus Gaussian noise)
  - OL multimodal when $Y_t > 0$ (almost always for t > 3)
- PF-EIS (K=1) outperforms all others

PF-EIS with Mode Tracker (PF-EIS-MT)

PF-EIS Mode Tracking (PF-EIS-MT)

- If for part of the residual state, the residual posterior is unimodal & narrow enough,
  - It can be approx by a Dirac delta function at its mode
  - Happens if residual STP narrow enough (LDSS property)
- Above: MT approx of Imp Sampling (IS) or IS-MT
  - MT is an approx to IS: introduces some error
  - But MT reduces IS dim by a large amount (improves effective particle size): much lower error for a given N
  - Net effect if chosen carefully: lower error when N is small

PF-EIS-MT algorithm

Choose $X_{t,s}$, $X_{t,r} = [\, X_{t,r,s}, X_{t,r,r} \,]$. For each t, for each i, do

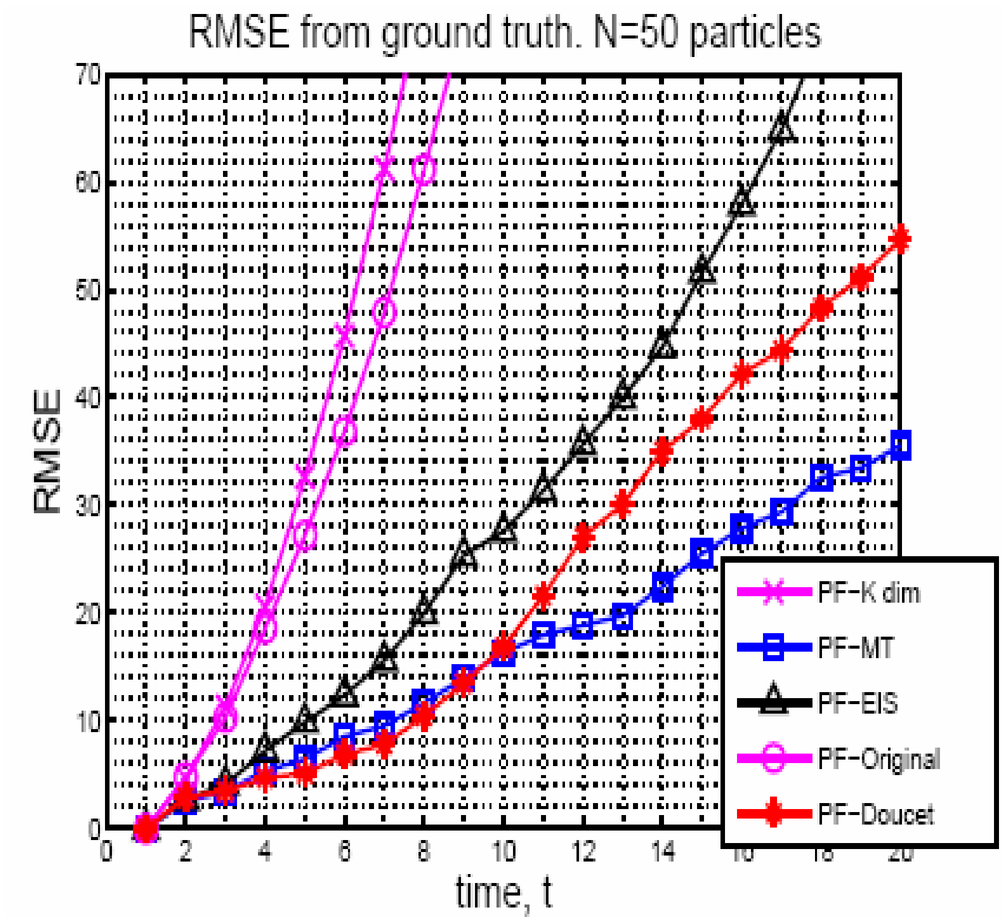
- Imp Sample $X_{t,s}^i$ from STP
- Compute $(p^{**})_G(X_{t,r})$: Gaussian approx to $p^{**}(X_{t,r})$ which is the posterior of $X_{t,r}$ given X_{t-1}^i, $X_{t,s}^i$
- Efficient Imp Sample $X_{t,r,s}^i \sim (p^{**})_G$
- Compute mode of $p^{**}(X_{t,r,r})$
- Set $X_{t,r,r}^i$ equal to this mode
- Weight & Resample

$$w_t^i \propto w_{t-1}^i \, OL(X_t^i) \, STP(X_{t,r}^i) \, / \, (p^{**})_G(X_{t,r}^i)$$

PF-MT

- PF-MT: computationally simpler version of PF-EIS-MT
  - Combine $X_{t,s}$ & $X_{t,r,s}$ & Imp Sample from STP for both. Mode Track $X_{t,r,r}$

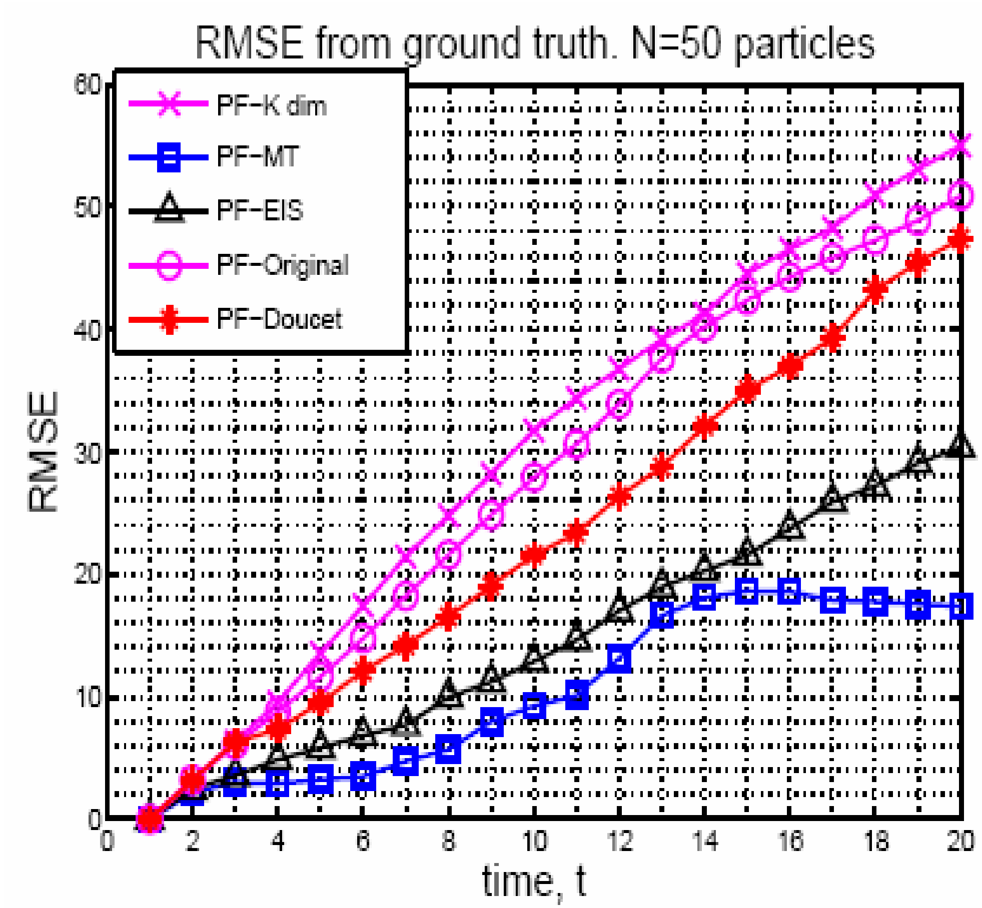
Simulation Results: Sensor failure



- Tracking on M=10 sensor nodes, each with two sensors per node. Node 1 has much higher failure prob than rest
- PF-MT (blue) has least RMSE
  - Using K=1 dim effective basis

IOWA STATE UNIVERSITY
Becoming the best.

Simulation Results: Nonlinear sensor



- Tracking on M=10 sensor nodes, each with one sensor per node. Node 1 has a squared sensor.
- PF-MT (blue) has least RMSE
  - Using K=1 dim effective basis

IOWA STATE UNIVERSITY
Becoming the best.

Summary

Efficient IS techniques for LDSS w/ multimodal OL

- Generalized existing work that assumed unimodality of posterior given previous state (p*)
- Derived sufficient conditions to test for unimodality of residual posterior, p** & used these to guide the choice of $X_{t,s}$, $X_{t,r}$ for PF-EIS
- If STP of $X_{t,r}$ narrow enough, p** will be unimodal & also very narrow: approx by Dirac delta function at its mode: IS-MT
  - Some extra error, but improves effective particle size

Open Issues

- PF-EIS much more expensive than original PF
  - Make mode computation faster
  - Choose effective basis dimension to min computation complexity (not N) for given error
- Compute Δ^* (bound on residual STP) efficiently & use it to choose $X_{t,s}$ on-the-fly
  - Or derive sufficient conditions to choose $X_{t,s}$ to max probability that p** will be unimodal offline
- Analyze IS-MT: systematic way to choose “narrowness” bound, when is net error lower?
- Extensions to PF-Smoothing (for offline problems)

Applications

Applications

- Deformable contour tracking
  - Affine PF-MT [Rathi et al, CVPR’05, PAMI (to appear)]
  - Deform PF-MT [Vaswani et al, CDC’06]
- Tracking spatially varying illumination change of moving objects
  - Moving into lighted room, face tracking [Kale et al, ICASSP’07]
- Tracking change in spatially varying physical quantities using sensor networks
  - Tracking temperature change [Vaswani,ICASSP’07]

Illumination Tracking: PF-MT

[Kale et al, ICASSP'07]

- State = Motion (3 dim) + Illumination (7dim)

- IS on motion (3 dim) & MT on illumination
  - Illumination changes very slowly
  - OL usually unimodal as a function of illumination
  - If OL multimodal as a fn of illumination (e.g. occlusions), modes usually far apart compared to illumination change variance

Face tracking results [Kale et al, ICASSP'07]

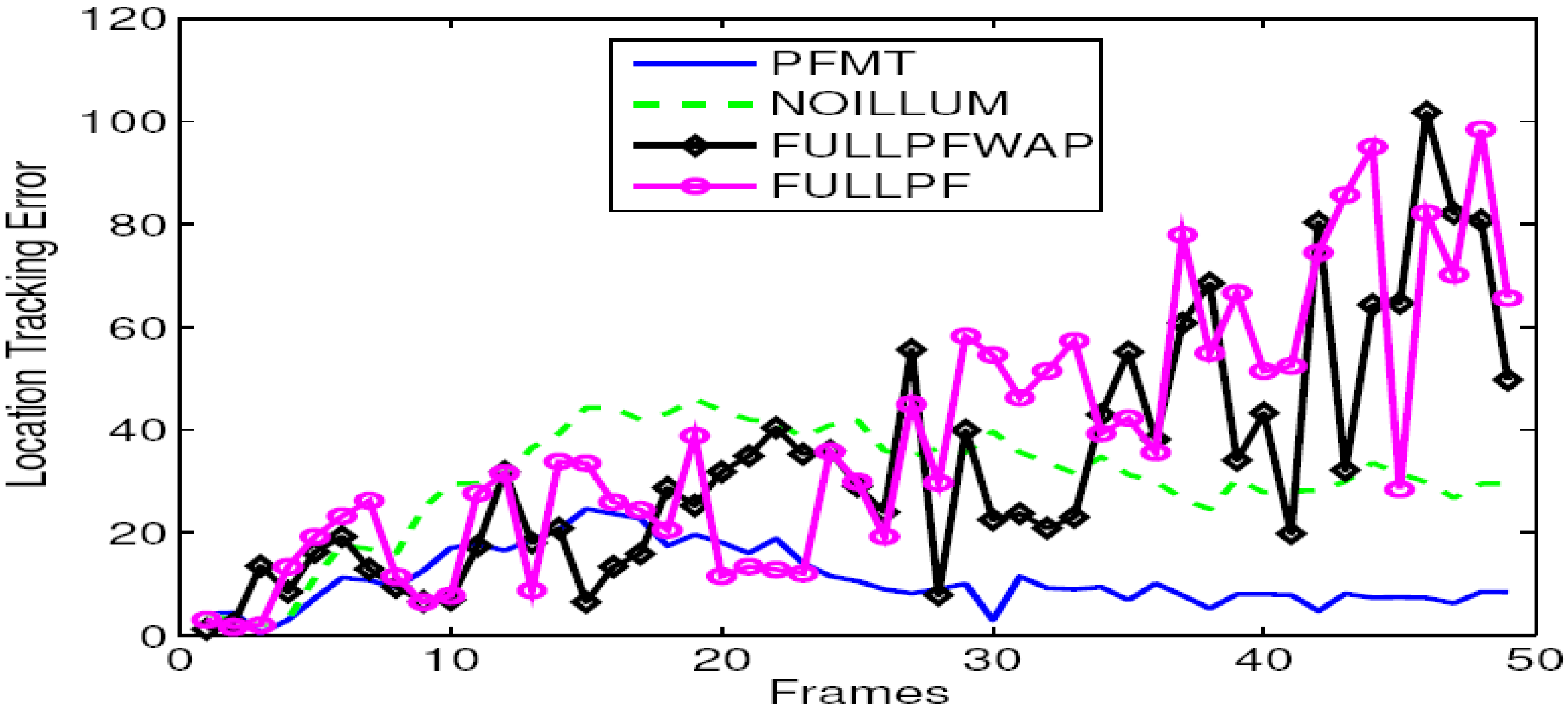
PF-MT

3 dim PF (no illum)

10-dim Auxiliary PF

Face tracking: RMSE from ground truth

[Kale et al, ICASSP'07]



Comparing PF-MT with 10 dim regular PFs (original, auxiliary) & with PF- K dim (not track illumination at all). N = 100

Deformable Contour Tracking

- State: contour, contour point velocities
- Observation: image intensity and/or edge map

- OL: segmentation energies (region or edge based)
  - Region based: observation is the image intensity. OL is the probability of the image being generated by the contour. Assumes a certain object/bgnd intensity model

  - Edge based: observation is the edge locations (edge map). OL is the probability of a subset of these edges being generated by the contour, and of others being generated by clutter or low contrast

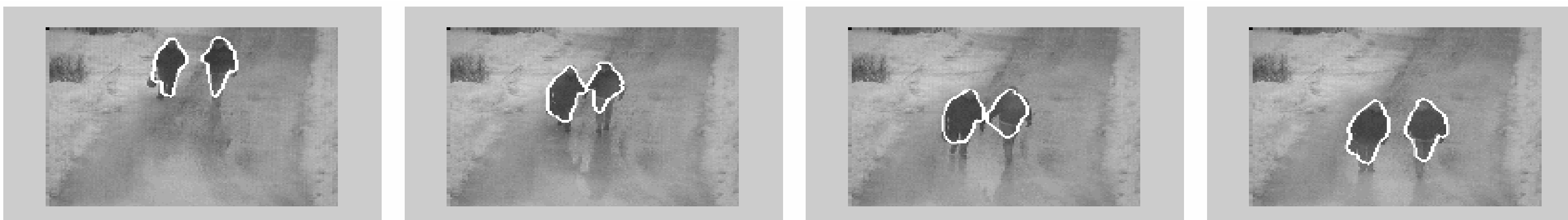
Two proposed PF-MT algorithms

- Affine PF-MT [Rathi et al, CVPR'05, PAMI (to appear)]
  - Effective basis sp: 6-dim space of affine deformations
  - Assumes OL modes separated only by affine deformation **or** small non-affine deformation per frame

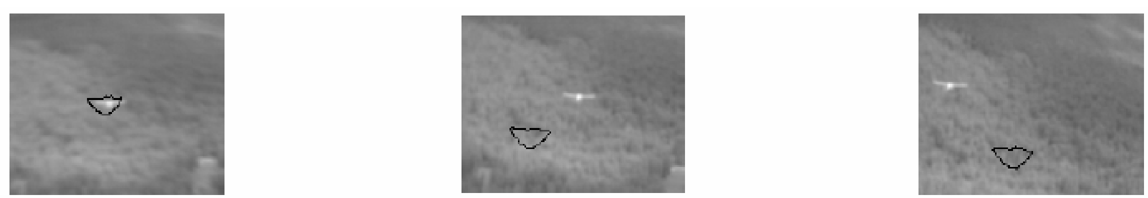
- Deform PF-MT [Vaswani et al, CDC'06]
  - Effective basis sp: translation & deformation at K sub-sampled locations around the contour. K can change
  - Useful when OL modes separated by non-affine def (e.g. due to overlapping clutter or low contrast) & large non-affine deformation per frame (fast deforming seq)

Low contrast images, small deformation per frame: use Affine PF-MT [Rathi etal,CVPR'05]

- Tracking humans from a distance (small def per frame)
- Deformation due to perspective camera effects (changing viewpoints), e.g. UAV tracking a plane

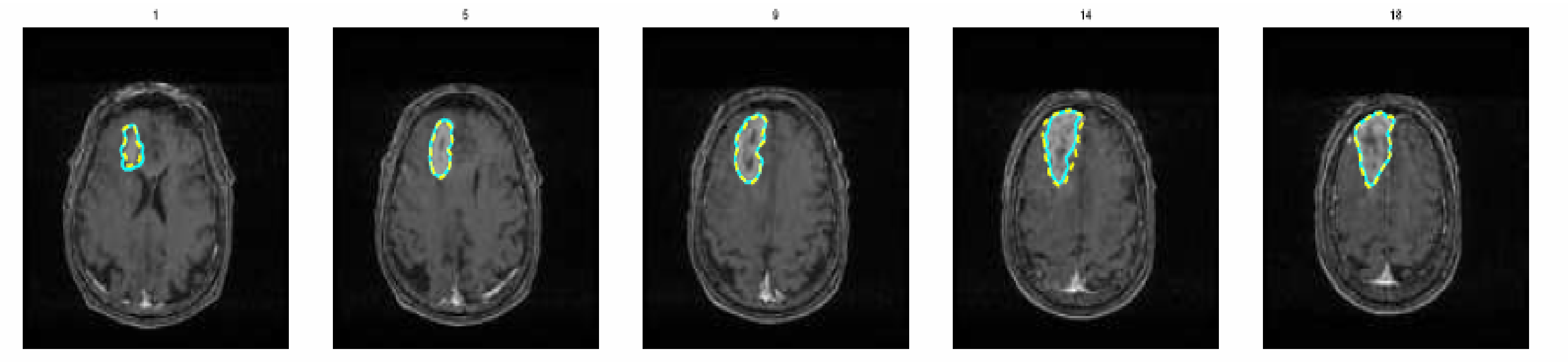


Condensation (PF K-dim) fails



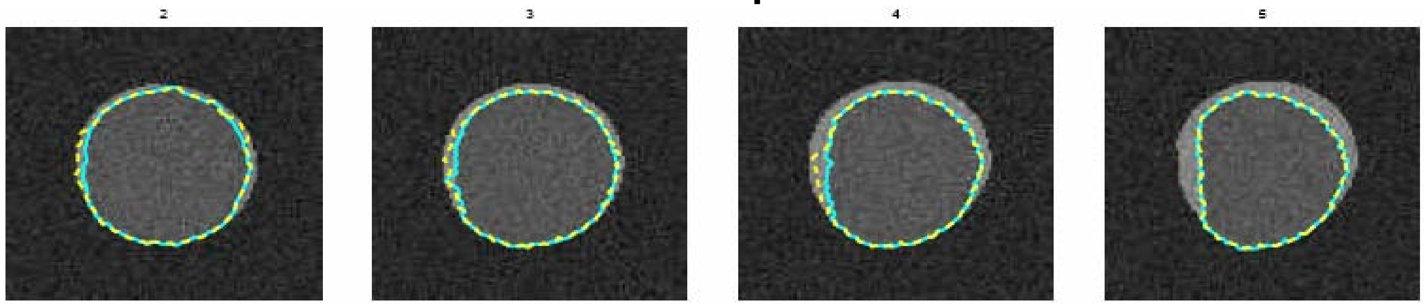
Low contrast images, large def per frame: Brain slices (Tumor Sequence)

- Multiple nearby OL modes of non-affine deformation: due to low contrast
- Tracking with Deform PF-MT [Vaswani et al,CDC'06]

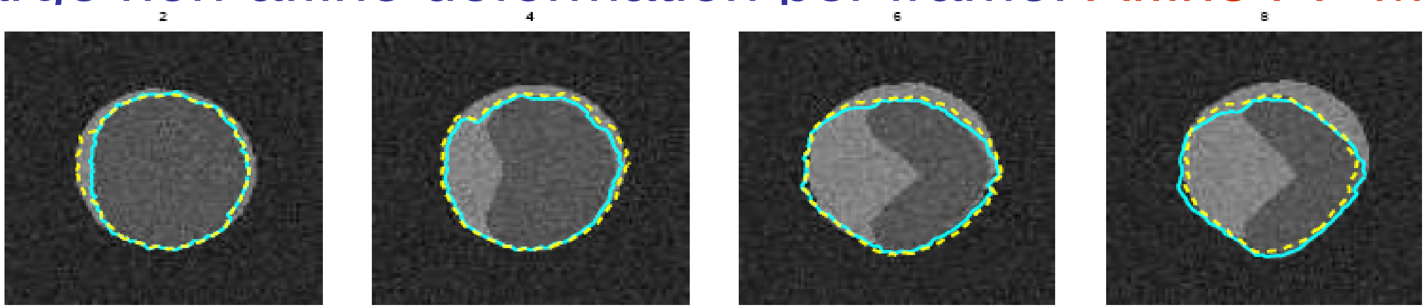


Overlapping Background clutter

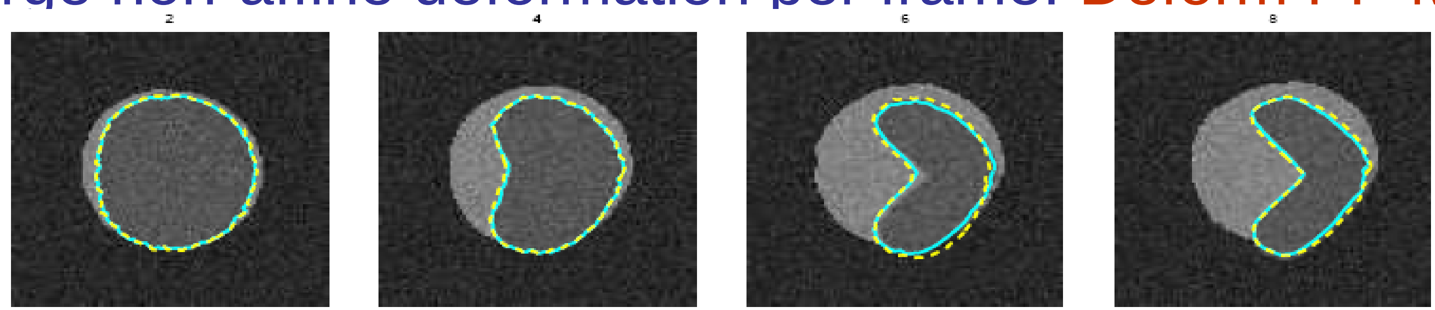
Small non-affine deformation per frame: Affine PF-MT works



Large non-affine deformation per frame: Affine PF-MT fails



Large non-affine deformation per frame: Deform PF-MT works



Partial Occlusions (car)

- 3 dominant modes (many weak modes) of edge based OL due to partial occlusion
- Tracking using Deform PF-MT [Vaswani et al, CDC'06]

(a) Tracking the contour for the car to left of the pole using Algorithm 2

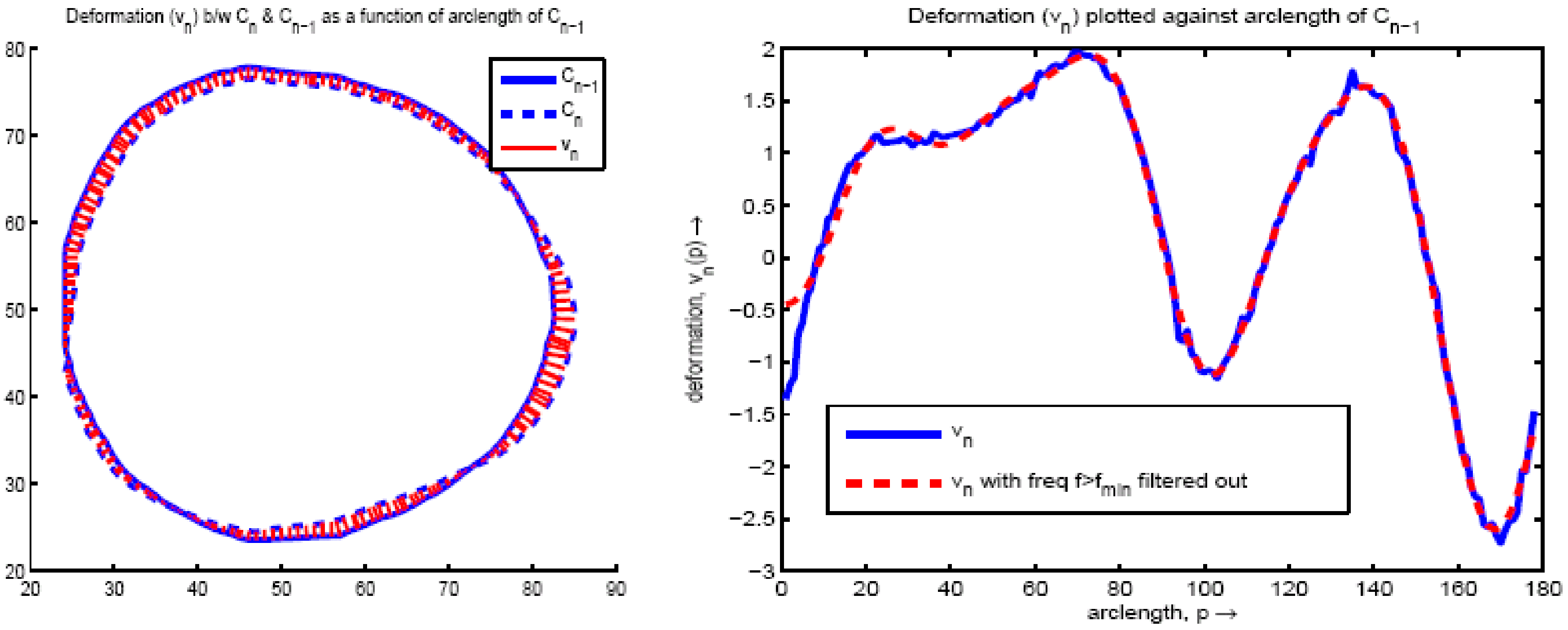
Ongoing Work: System Id

System Id

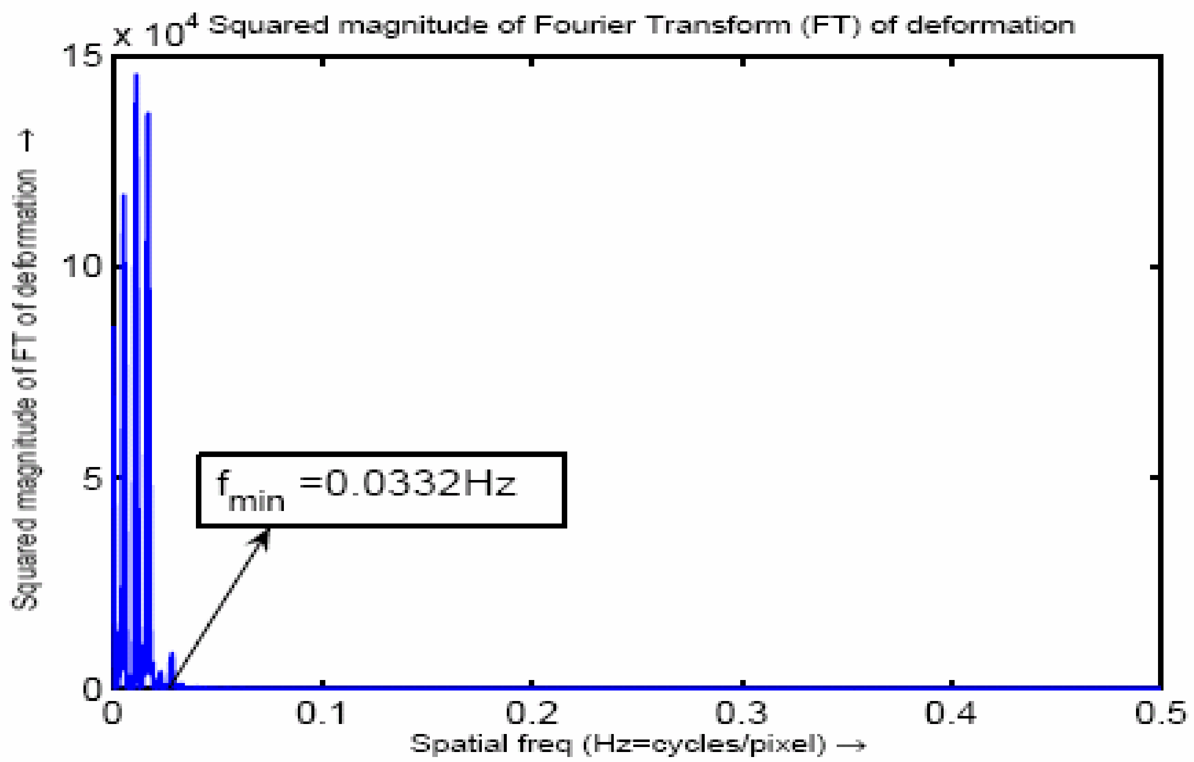
- LDSS: Time sequences of discretized spatial "signals" (usually heavily oversampled)
  - "signal": spatially stationary or p.w. stationary
- System Id problem has 2 parts
  - Estimate effective basis dim: use PSD of spatial signal & its r% cut-off frequency
  - Learn Temporal dynamics: AR model on LPF'ed Fourier coefficients or on subsampled spatial signal

An Example: Estimating K

M = 178 dimensional contour deformation “signal”



Using PSD to estimate K



f_{min} = 0.03Hz for residual deformation = 0.05% of total deformation. M = 178, K = $\lceil M.\ 2\ f_{min} \rceil$ = 12

System Id: Open Issues

- Temporally piecewise stationary sequences
  - Detect changes in PSD (spatial)
  - Detect change in effective basis dimension, K
  - Detect change in temporal dynamics (if K same)
  - Do all the above while tracking
- Spatial nonstationarity
- Contour deformation sequences: spatial axis (arclength) warps over time: total L changes, distance b/w points changes
  - Effects not studied in regular signal processing

Collaborators

- Deformable contour tracking
  - Anthony Yezzi, Georgia Tech
  - Yogesh Rathi, Georgia Tech
  - Allen Tannenbaum, Georgia Tech

- Illumination tracking
  - Amit Kale, Siemens Corporate Tech, Bangalore

- System Id for time sequences of spatial signals
  - Ongoing work with my student, Wei Lu