Statistical Models for Deformable Contour Tracking

Namrata Vaswani
Dept. of Electrical & Computer Engineering
Iowa State University
http://www.ece.iastate.edu/~namrata

Based on joint work with Rathi,Tannenbaum & Yezzi

Goal

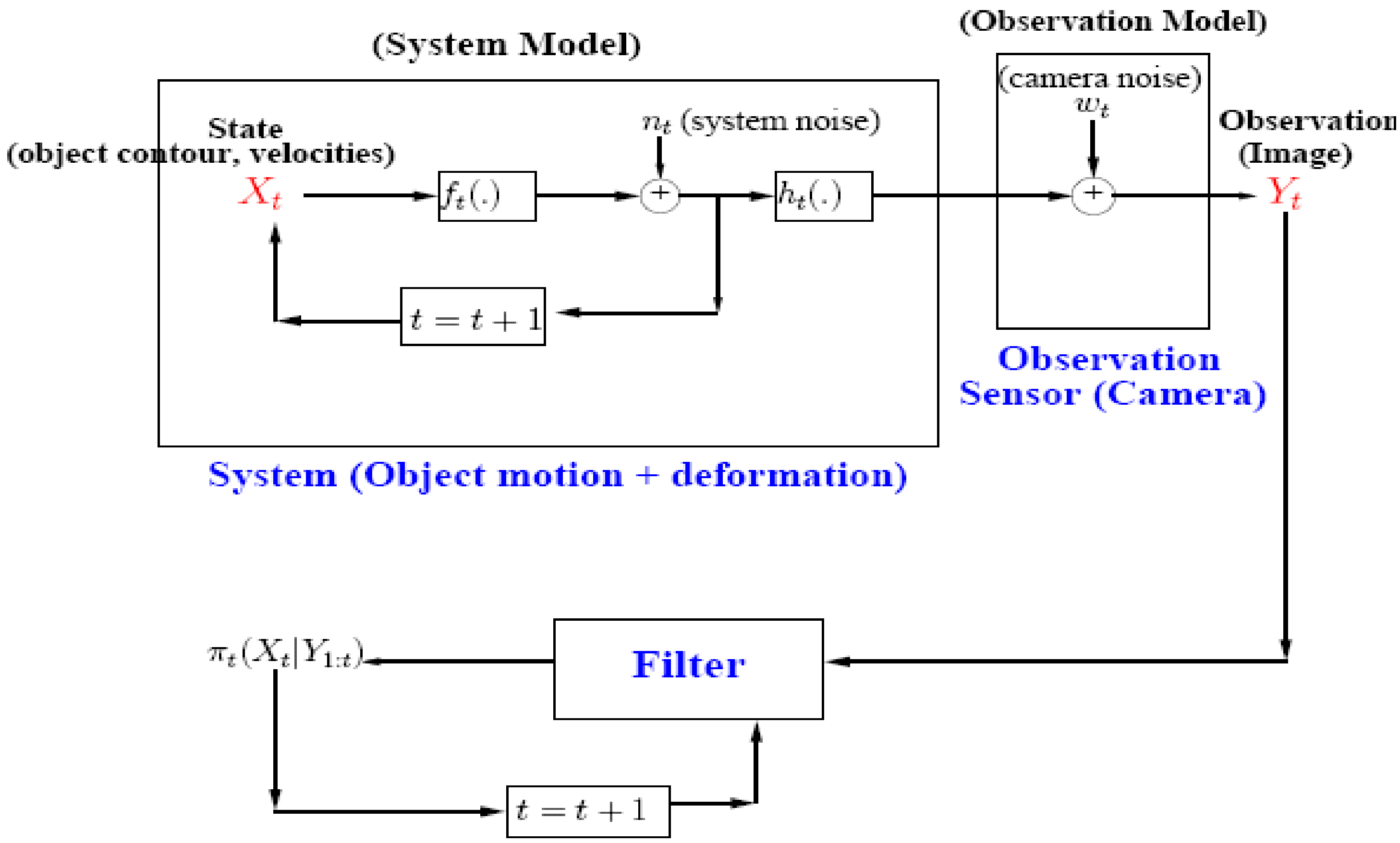
- Sequentially segment moving & deforming objects or Regions of Interest (ROIs) from video or spatial image sequences

- Deforming contours occur due to
  - Deforming objects, e.g. ROIs in heart, brain slices
  - Partial occlusions
  - Perspective effects: changing camera viewpoints

Problem Formulation

- Formulate sequential segmentation as a causal Bayesian estimation of the “optimal” state: “Tracking” or “Optimal Filtering”
  – State = contour, contour velocity at t
    - Contour: represented using level set method
    - Contour velocity = global motion + local deformation
    - Local deformation = large (infinite) dim.
  – Observation = image at t

State Space Model & Tracker



Observation Likelihood (OL)

- Image = noisy & nonlinear function of contour
  - $p(Y_t|C_t) \propto \exp[-E(Y_t,C_t)/\sigma^2]$
  - E = any segmentation energy functional
    - e.g. E=Chan-Vese energy, edge energy, or sum of both
- OL may be highly non-Gaussian
    - e.g. multimodal or heavy tailed or flat

Examples: Non-Gaussian OL

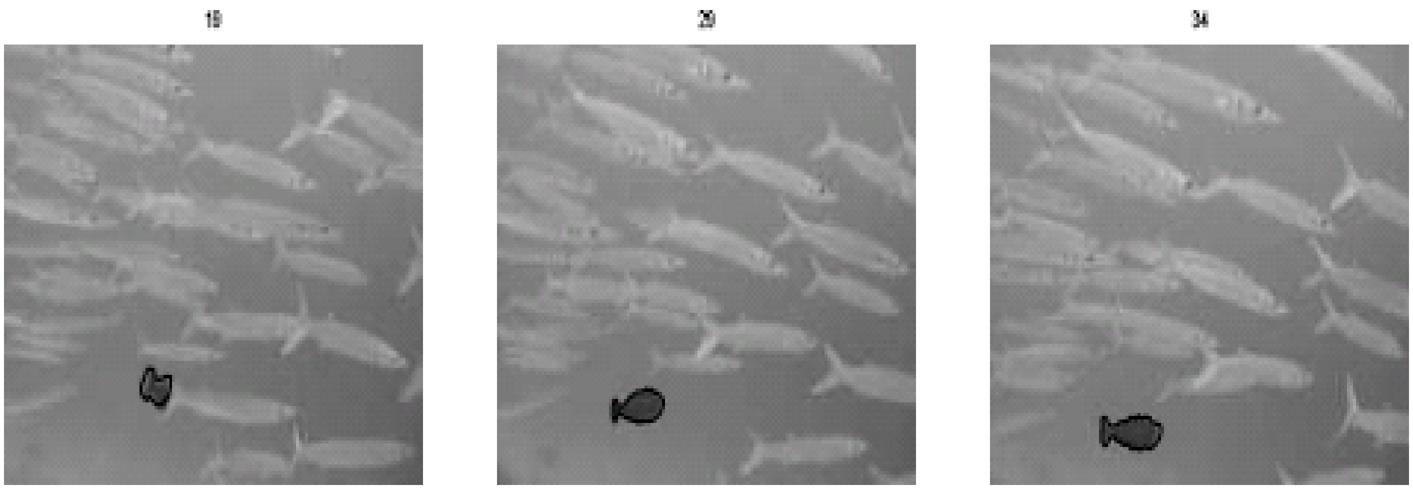
- Multimodal OL
  - Multiple separate objects
  - Background clutter
  - Partial occlusions

- Heavy tailed or flat OL
  - Low contrast b/w object & background
  - Outliers

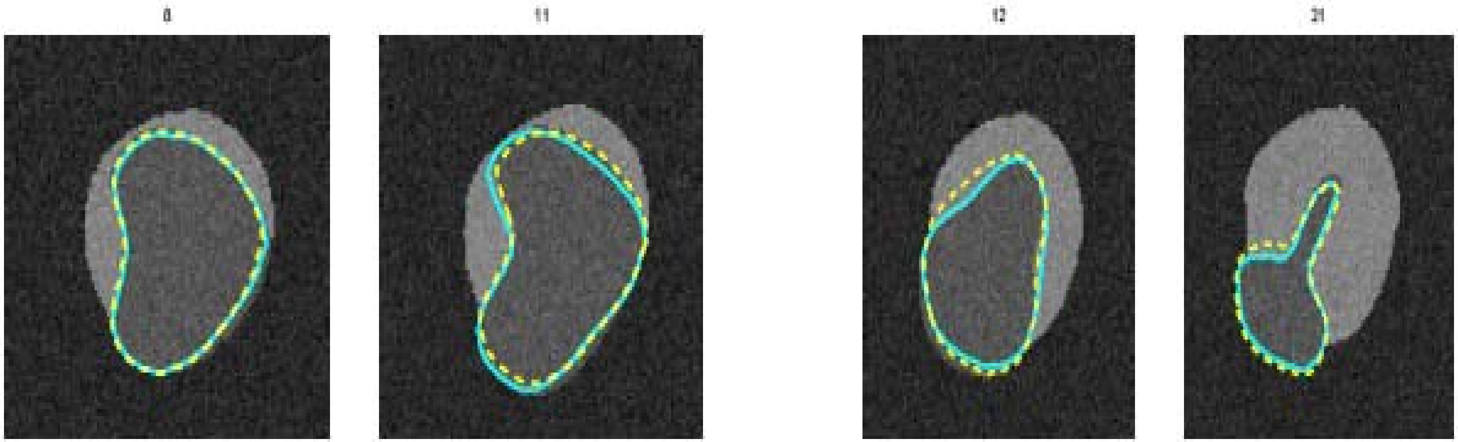
Examples: Deformation

- Actual deformations
  - Human tracking: surveillance, sports videos,...
  - Animals such as a fish
  - Medical sequences: ROIs in brain or heart
- Changing region of partial occlusions
  - Automatic vehicle navigation
  - Robot navigation
- Frequently changing camera viewpoint
  - Tracking using a UAV

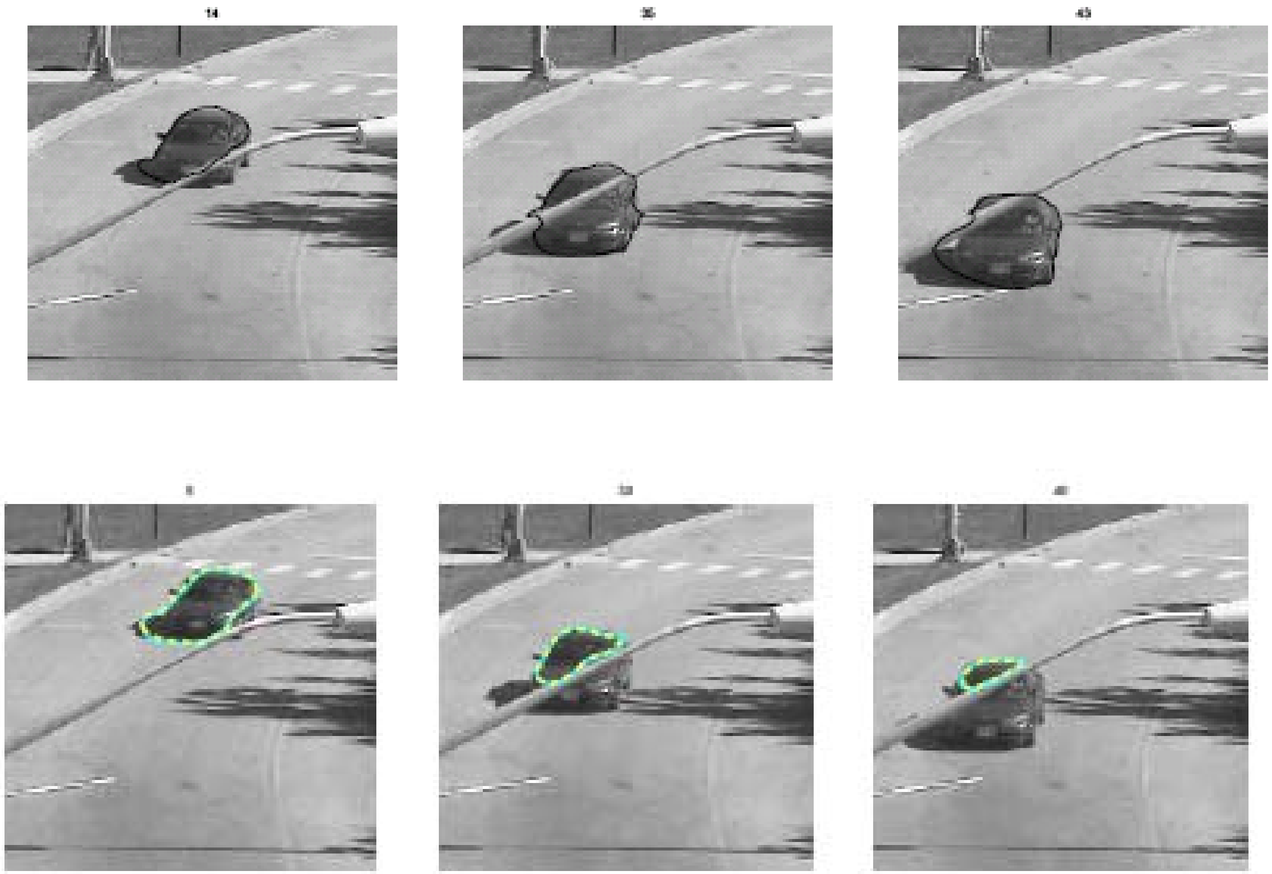
Multiple fishes, partial occlusions + deformation



Background clutter (light grey object) + deformation

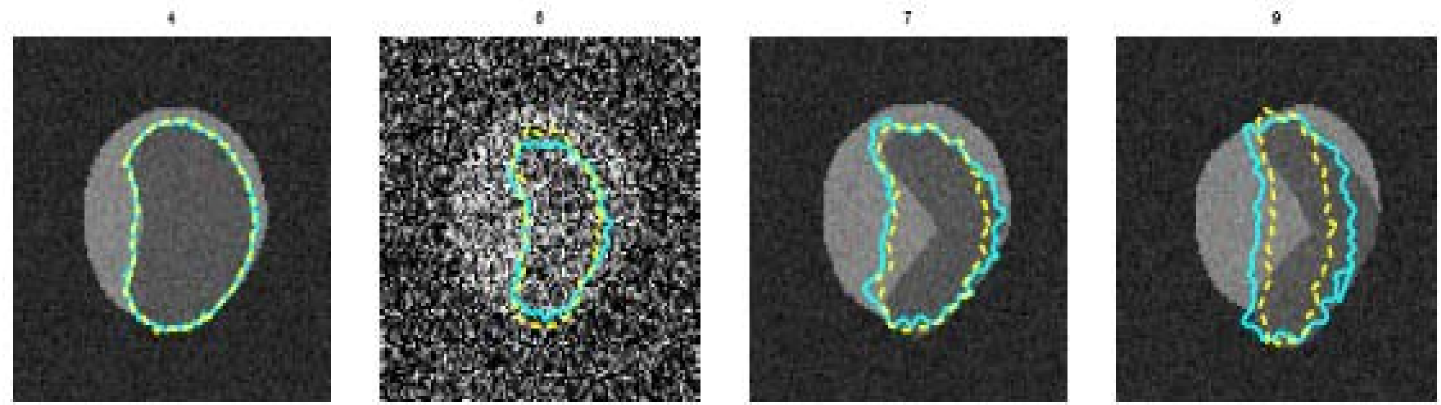


Partial occlusion of car by street light: Multiple, deforming contours

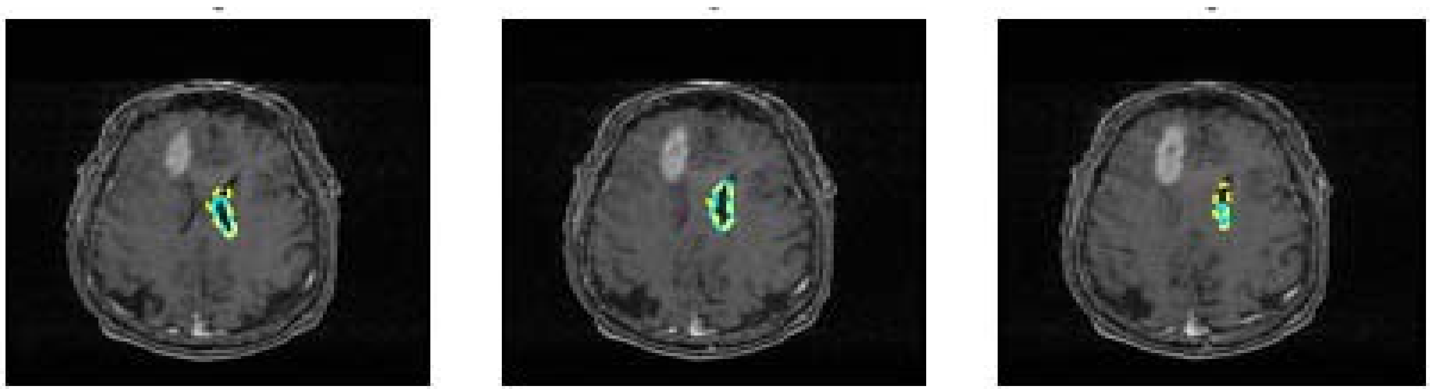


Low contrast + Frequent viewpoint changes

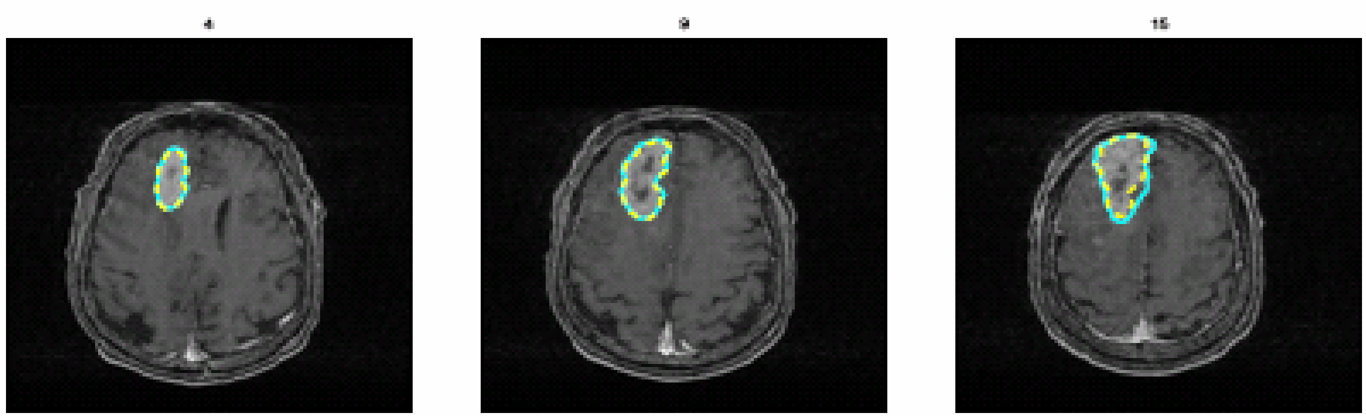
Outlier image+ background clutter + deformation



Deforming ROI: Right ventricle tracking



Low Contrast + Deforming ROI: Tumor tracking



Outline

- Main Issues to Address
- Particle Filtering (PF)
- Solution 1: PF on space of affine deform^n, MT for non-affine
- Solution 2: PF on space of translations & sub-sampled local deformation, MT for rest

Three Issues

- Nonlinearities
  - Space of contours: not a vector space
  - Affine/similarity group applied to contour
- Non-Gaussian Observation Likelihood
  - Multimodal or heavy tailed
- Local deformation: Large dim state space

Contour Tracking Literature

- Tracking as Bayesian state estimation
  - Fixed finite representation of contour
    - Affine tracking: Condensation & precursors
    - Deformable templates, Landmark based,…
  - Exemplars to handle non-affine deformations
  - Separate linear observers (posterior mode trackers) for global & local deformation separately
- Other formulations: OF, region tracking,...

Past Work: Problems

- Finite dim. representation: do not handle large changes in contour length/ topology
- Exemplars: very restrictive
- Posterior mode trackers (approx. linear observers
  - Assume unimodality of posterior
  - Uncoupled observers for global & local deformation (coupled observer: nonlinear)

A Possible Solution

- Replace the approx. linear observer by a particle filter
  - Can track nonlinear systems: able to use a coupled observer
  - Can handle multimodal observation likelihoods (multimodal posteriors)

Particle Filtering (PF)

- Sequential Importance Sampling (SIS)
  - IS: MC method to approx $p(x|y) \propto p(y|x)\, p(x)$
    - For $i=1,..N$, sample: $x^i \sim q(x|y)$: importance density
    - Weight x^i by $w^i \propto p(y|x^i)\, p(x^i) \,/\, q(x^i|y)$
    - $\mathrm{E}[f(x)|y] = \sum_i f(x^i)\, w^i$
  - Sequential IS:
    - At t, use $x_t \equiv X_{0:t}|Y_{1:t-1}$, $y_t \equiv Y_t$,
    - Choose q_t: $q_t = q_{t-1} q\,(X_t|X_{t-1}, Y_t)$ → recursive update

- \+ Resampling to reduce degeneracy

Particle Filter Algorithm

$$\text{Given } \pi_{t-1}^{N}(x_{0:t-1} \mid Y_{1:t-1}) = \sum_i w_{t-1}^i \delta(x_{0:t-1} - x_{0:t-1}^i). \text{ At } t, \text{ do}$$

$$\text{- Imp. Sample: For } i = 1,2,..N \text{ do, } \; x_t^i \sim q(x_t \mid x_{t-1}^i, Y_t)$$

$$\text{- Weight: For } i = 1,2,..N \text{ do, } \breve{w}_t^i = w_{t-1}^i \frac{p(Y_t \mid x_t^i)\, p(x_t^i \mid x_{t-1}^i)}{q(x_t \mid x_{t-1}^i, Y_t)}, \;\; w_t^i = \frac{\breve{w}_t^i}{\sum_i \breve{w}_t^i}$$

$$\pi_t^N(x_{0:t} \mid Y_{1:t}) = \sum_i w_t^i \delta(x_{0:t} - x_{0:t}^i)$$

- Resample

Solution 1: PF + MT for affine basis

PF-MT: Main Idea [CVPR'05]

- Use a PF to track affine deformation and use an approx. Mode Tracker (MT) that computes mode of posterior of local deformation given the affine deformation
  - PF on large dim state spaces is expensive & inaccurate (for manageable *N*)

State Space Model

- State: $X_t = [C_t, A_t]$,
  - C_t = contour at t, A_t = affine deform^n b/w C_{t-1} & C_t
  - A_t : random walk motion model
  - $\hat{C}_t = A_t(C_{t-1})$
  - $p(C_t | \hat{C}_t) \propto exp[- d^2(C_t, \hat{C}_t)/\Delta]$
- Observation: Y_t = image at t
  - OL: $p(Y_t|C_t) \propto exp -[E(C_t, Y_t)]$
  - E = Chan-Vese energy + edge energy (Condensation)
  - Multimodal: clutter/occlusions/multiple objects

PF - MT (Exact)

- At t, for each particle $i=1,2,\dots N$
  - IS for A_t: $A_t^i \sim$ state transition pdf
  - Compute $\hat{C}_t^i = A_t^i (C_{t-1}^i)$
  - Compute the single mode of $p(C_t | \hat{C}_t^i, Y_t)$

    $m_t^i = \arg\min_C [E(C) + d^2(C, \hat{C}_t^i) / \Delta]$
  - IS for C_t: $C_t^i \sim N(C; m_t^i, \Delta I)$
  - Compute IS weights & resample

$$w_t^i \propto \frac{exp[-E(C_t^i)]exp[-d^2(C_t^i, \hat{C}_t^i)]}{\mathcal{N}(C_t^i; m_t^i, \Delta I)}$$

PF – MT: Approx. 1

- At t, for each particle $i=1,2,\ldots N$
  - IS for A_t: $A_t^i \sim$ state transition pdf
  - Compute $\hat{C}_t^i = A_t^i(C_{t-1}^i)$
  - Compute the single mode of $p(C_t | \hat{C}_t^i, Y_t)$

    $m_t^i = \arg\min_C \, [\, E(C) + d^2(C, \hat{C}_t^i) / \Delta \,]$
  - IS for C_t: Deterministically set $C_t^i = m_t^i$
  - Compute IS weights & resample

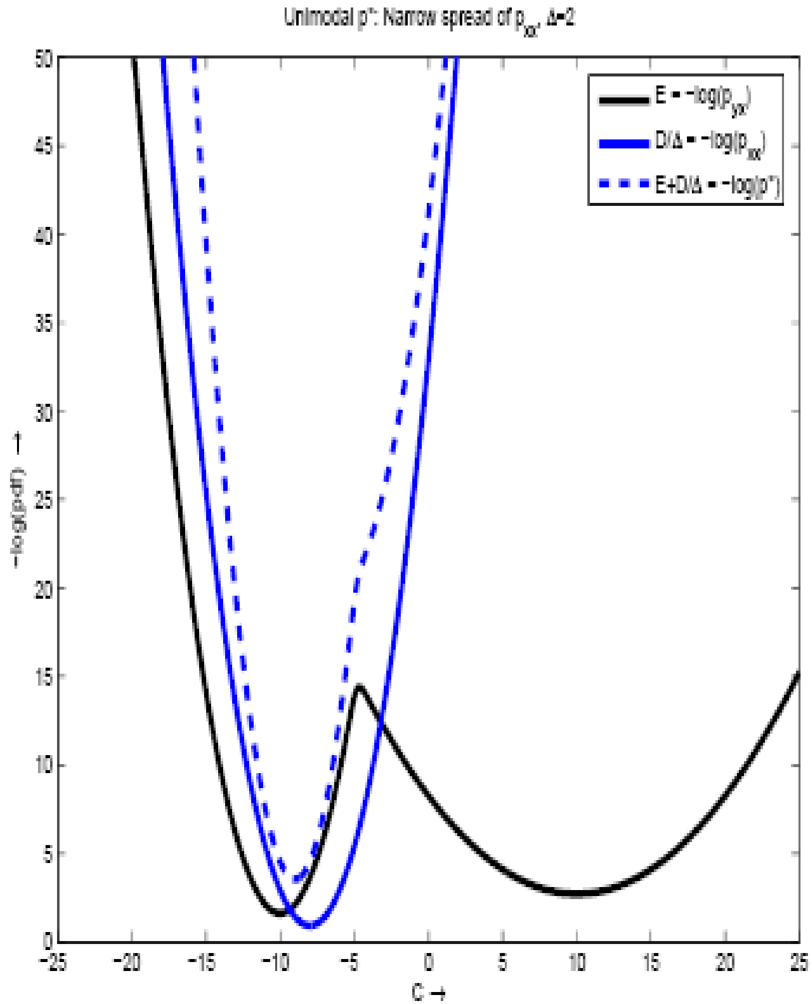
$$w_t^i \propto \frac{exp[-E(C_t^i)]exp[-d^2(C_t^i,\widehat{C}_t^i)]}{\mathcal{N}(C_t^i;m_t^i,\Delta I)}$$

Implicit Assumption [ICASSP'06]

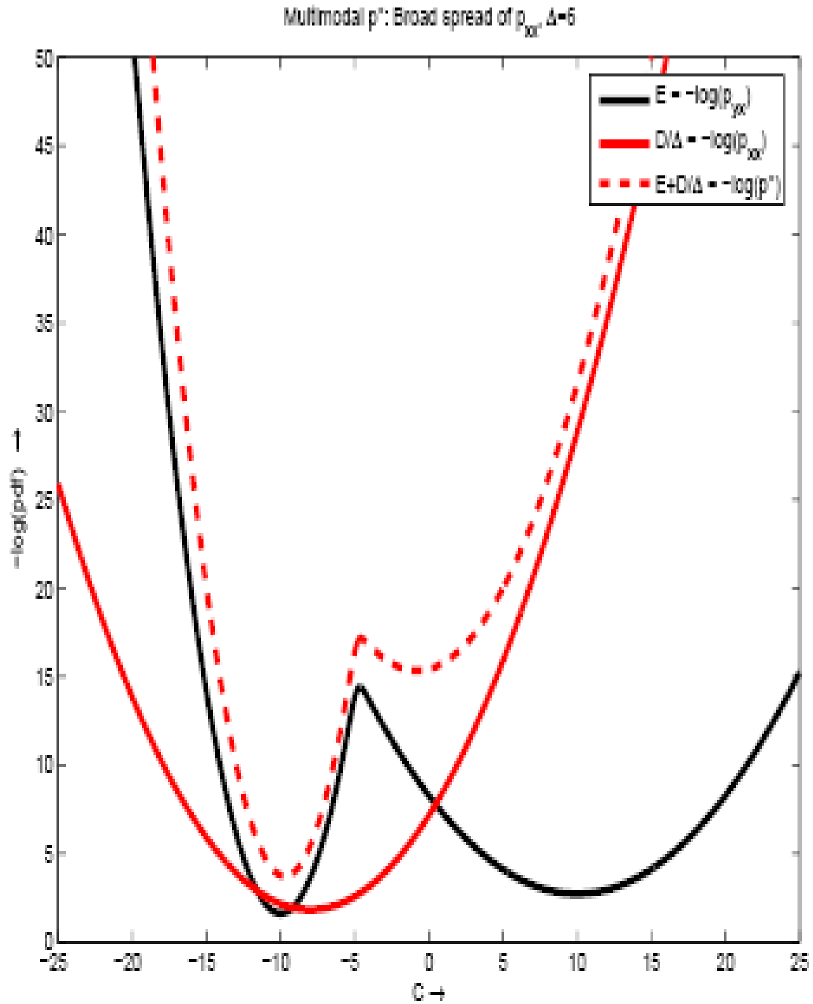
- Non-affine deformation per frame "small" enough compared to distance b/w modes of OL at same affine location, to ensure unimodality of $p(C_t | \hat{C}_t, Y_t)$

$$p(C_t | \hat{C}_t, Y_t) \propto p(Y_t | C_t)\, p(C_t | \hat{C}_t)$$
$$\propto \exp -[\, E(C_t) + d^2(C_t, \hat{C}_t)/\Delta \,]$$

  – Δ small enough, so that $E + d^2/\Delta$ has a single minimum on space of non-affine deform^n



Unimodal p*: Narrow spread of $p_{xx'}$, Δ=2
E = −log(p_{yx})
D/Δ = −log($p_{xx'}$)
E+D/Δ = −log(p*)
−log(pdf) →
C →



Multimodal p*: Broad spread of $p_{xx'}$, Δ=6
E = −log(p_{yx})
D/Δ = −log($p_{xx'}$)
E+D/Δ = −log(p*)
−log(pdf) →
C →

Ensuring Unimodality of $p(C_t | \hat{C}_t, Y_t)$

- $p(C_t | \hat{C}_t Y_t) \propto p(Y_t | C_t)\, p(C_t | \hat{C}_t)$
  $\propto \exp -[E(C_t) + d^2(C_t, \hat{C}_t)/\Delta]$
- $E + d^2/\Delta$ has a single minimum if
  - $\hat{C}_t$ lies in a locally convex region of E
  - $$\Delta < \Delta^* = \min_{C \in A} \max_{p} \frac{|(\nabla_C D(C, \hat{C}_t))(p)|}{|(\nabla_C E(C))(p)|}$$
    - $D = d^2$
    - $A = \{C \in R^c : \nabla D\,(p) \,.\, \nabla E\,(p) < 0,\ \forall p\}$
    - R = {largest region of Q containing $\hat{C}_t$ & where E is locally convex}, $Q = \{\hat{C}_t$ + non-affine deformations}

Approx. MT [ICASSP'06]

- Approximate: $m_t^i = arg\ min_C [E(C) + d^2(C, \hat{C}_t^i) / \Delta]$

- m_t^i satisfies $E(C_{min}) < E(m_t^i) < E(\hat{C}_t^i)$
  - C_{min} = minimizer of *E in R* (locally convex region of *E* containing $\hat{C}_t^i$)
  - Start from $\hat{C}_t^i$ & perform Gradient Descent (GD) to minimize *E*: GD will go towards C_{min}
  - If GD iterations slow enough, will cross m_t^i at "some" iteration
- Approx. solution: starting with $\hat{C}_t^i$, run "some" iterations of GD to minimize *E*

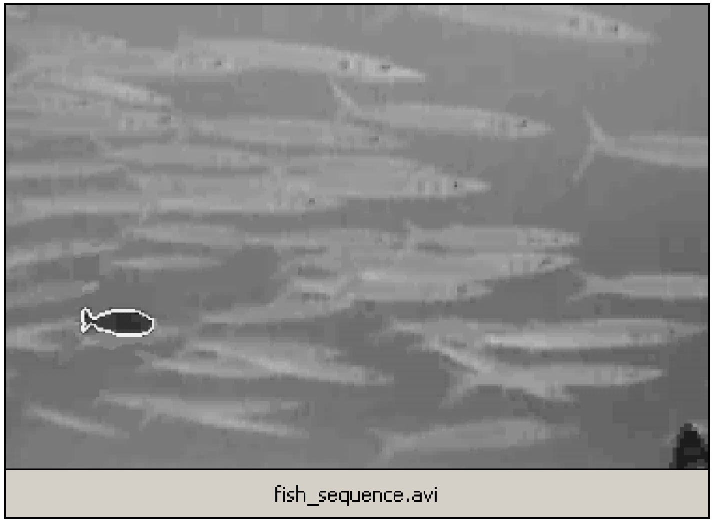
A New PF Technique?

- Extends optimal IS [Doucet'98] for multimodal $p(X_t|X_{t-1}, Y_t)$ satisfying "$p(X_t|X_{t-1}, X_{t,s}, Y_t)$ is unimodal" assumption & for which Δ is small
- Fast PF technique for large dim states: sample only from a small dim subspace, MT for rest.

Validity of “Assumption”

- “Non-affine deformation per frame small compared to distance b/w modes of OL at same affine location” valid for
  - Distinct objects (separated by translation)
  - Concentric contours (separated by scale)
  - Low contrast & camera viewpoint changes (small non-affine deformation per frame)

Multiple fishes, partial occlusions + deformation (Modes separated by translation)



Plane Sequence taken from a UAV: Low contrast & Frequent viewpoint changes (Small non-affine deformation per frame)

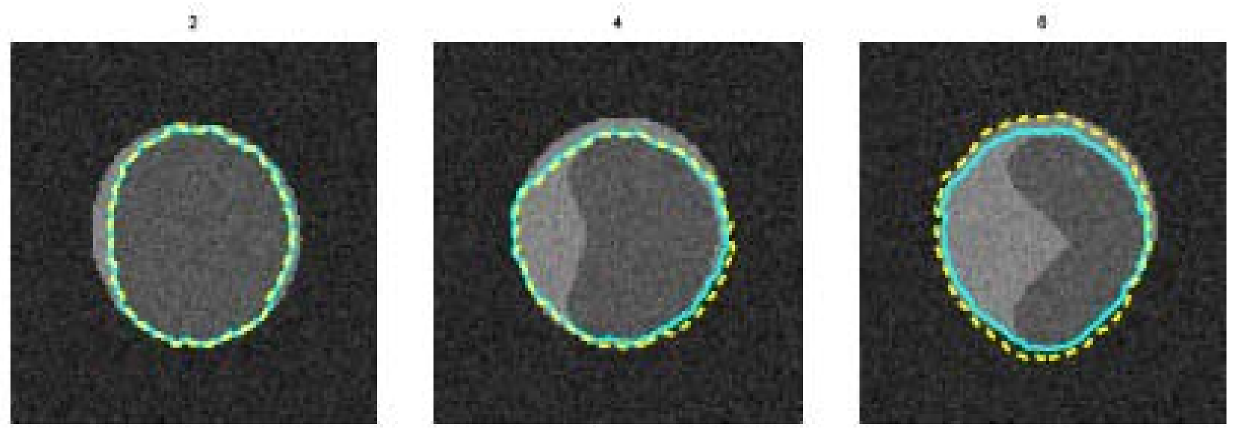
Assumption Fails when

- 2 or more OL modes at same affine location & large deformation per frame
  - e.g. car sequence, medical image sequences

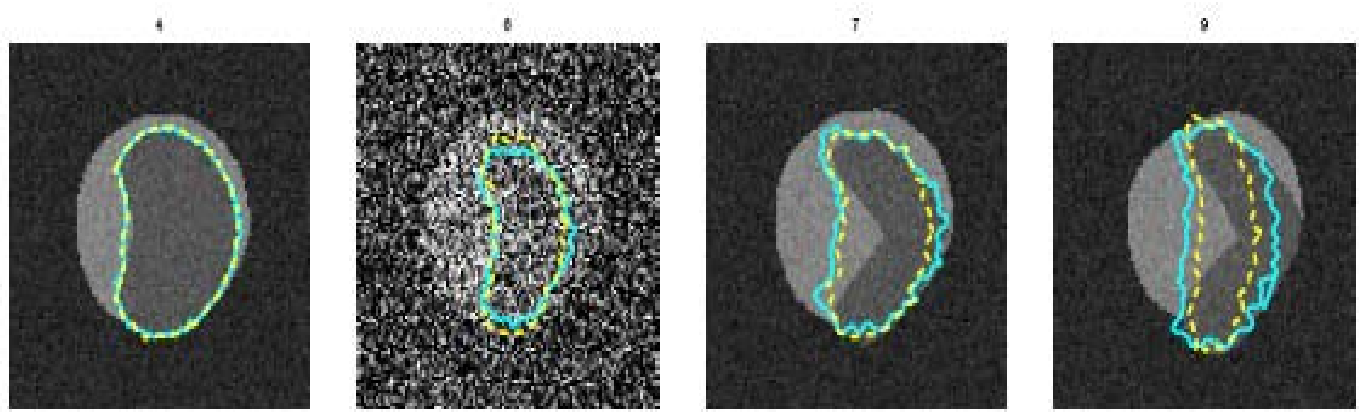
- Outlier observations (multiple modes of OL very close) & large deformation per frame

- Both cases: Contour attracted to wrong mode in MT step

Background clutter – light grey object (2 OL modes at same affine location) & Large non-affine deformation per frame

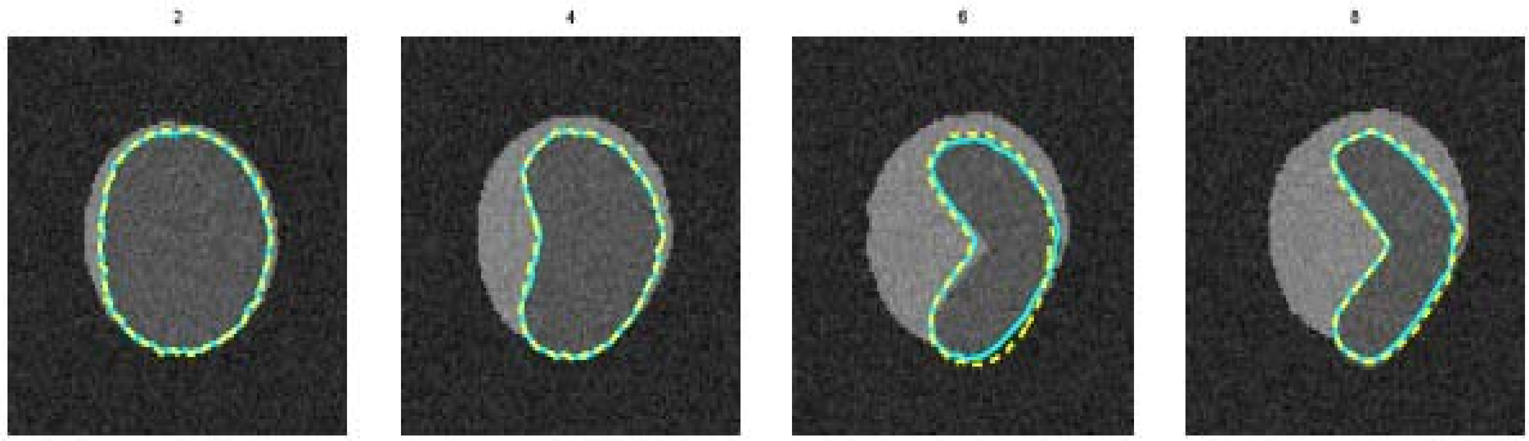


Outlier observations at every even frame (Multiple OL modes at same affine location) & Large non-affine deformation per frame

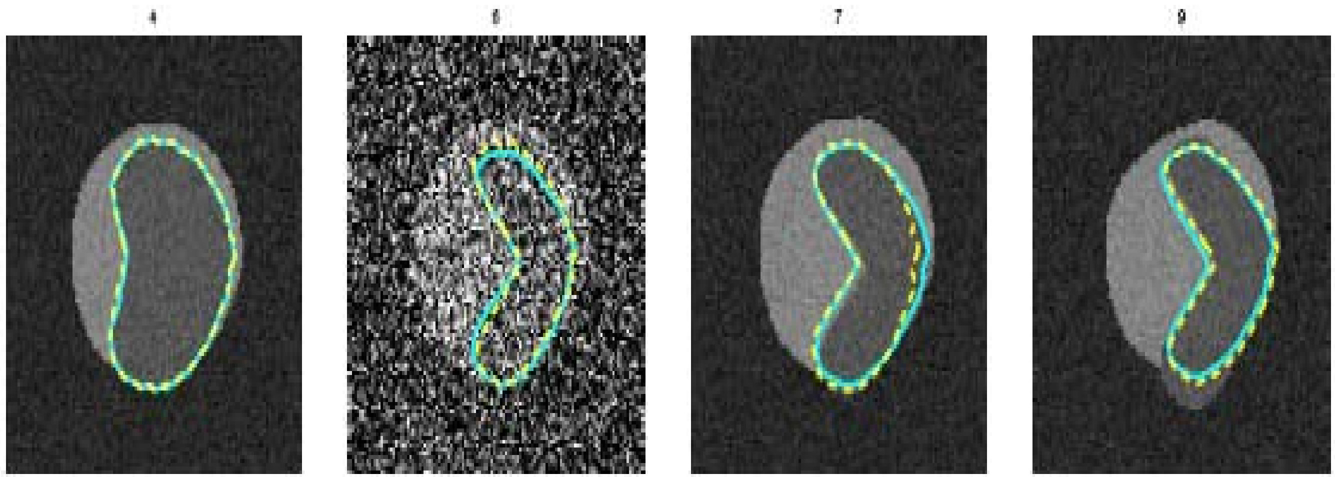


Solution 2: PF for sub-sampled local deformation + MT for rest

Background clutter
(2 OL modes at same affine location) &
Large non-affine deformation per frame



Outlier observations at every even frame (Multiple OL modes at same affine location) & Large non-affine deformation per frame



Car left of pole

Full car

Weakening the Assumption

- Previous assumption fails when
  - Non-affine deformation separates modes of OL
  - Outlier observations

- But, if space of deformations is a separable Hilbert space, it has a countable orthogonal basis
  - There exists a *K*-dim subspace (*K*-dim basis) in which "most" of the deformation occurs
    - Can achieve any approx error, Δ, for "residual deformation" by choosing *K* large enough

Spatial Frequency Interpretation

- Contour length is finite and spatial frequency of contour deformation is approx band-limited, so that $K = L/(2f_{max})$ or a bit larger (in practice): Nyquist criterion

- Using a smaller K ⇔ low pass filtering (estimating a smoothed contour)

- Using a much larger K ⇔ estimating noise

PF-MT-NonAffine: Main Idea

- Replace affine basis by a *K*-dim B-spline basis to parameterize contour deformation velocity, *K* can change with *t*
  - B-spline basis only for contour velocity: need a much smaller *K* than for contour

- Run a PF to track velocity at *K* control points (random walk model on velocity), run approx MT for rest, detect need to change *K*

State Dynamics

$$\tau = \text{frame interval}, \quad \vec{N} = \text{normal}, \quad B_s = \text{basis}$$

$$v_{t,s} \in R^K, \rho_{t,s} \in R^2, C_t \text{ in a manifold of } R^{M_t}$$

$$v_{t,s} = v_{t-1,s} + w_n, \quad w_n \sim N(0, \Sigma\tau), \qquad \rho_{t,s} = \rho_{t-1,s} + w_{n,\rho}$$

$$k = ceil(\tau / \tau_{CFL}), \tau_{CFL} = 1/\max_j (v_{t,s,j})$$

$$\text{For } m = 1,2,...k : \hat{C}_{t-1} = C_{t-1}$$

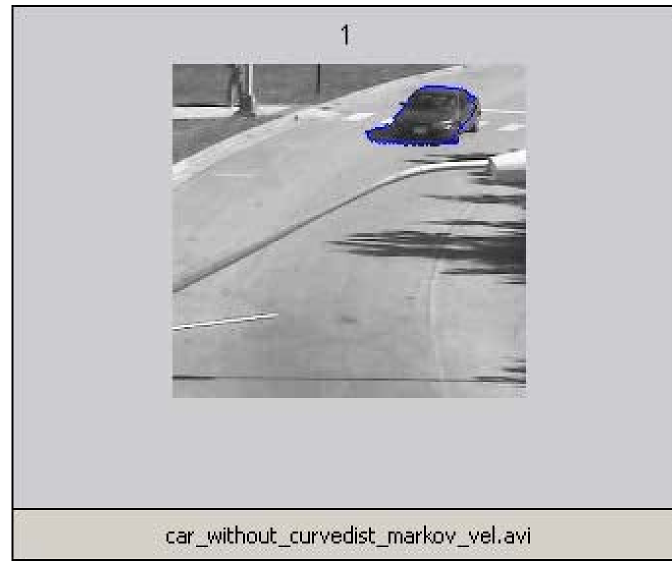
$$\hat{C}_{t-1+\frac{m}{k}}(p) = \hat{C}_{t-1+\frac{m-1}{k}}(p) + \tfrac{\tau}{k} \vec{N}(p)[B_s(p) v_{t,s} + \vec{N}^T \rho_{t,s}]$$

$$p(C_t \mid \hat{C}_t) = (const) \exp[-\frac{d^2(C_t, \hat{C}_t)}{\Delta}]$$

PF-MT-NonAffine Algorithm

- Importance Sample velocity at K control points
- Interpolate to get contour velocity, $v_{t,s}^i$
- Move C_{t-1}^i by $v_{t,s}^i$ to get $\hat{C}_t^i$
  - Move slowly enough to satisfy CFL condition: multiple level set evolution iterations for one $v_{t,s}^i$
- Approx. MT for "rest of deformation": $C_t^i = m_t^i$
- Weight & Resample
- Detect if K-dim basis suffices, else increase K

Deformation due to partial occlusions



Full car

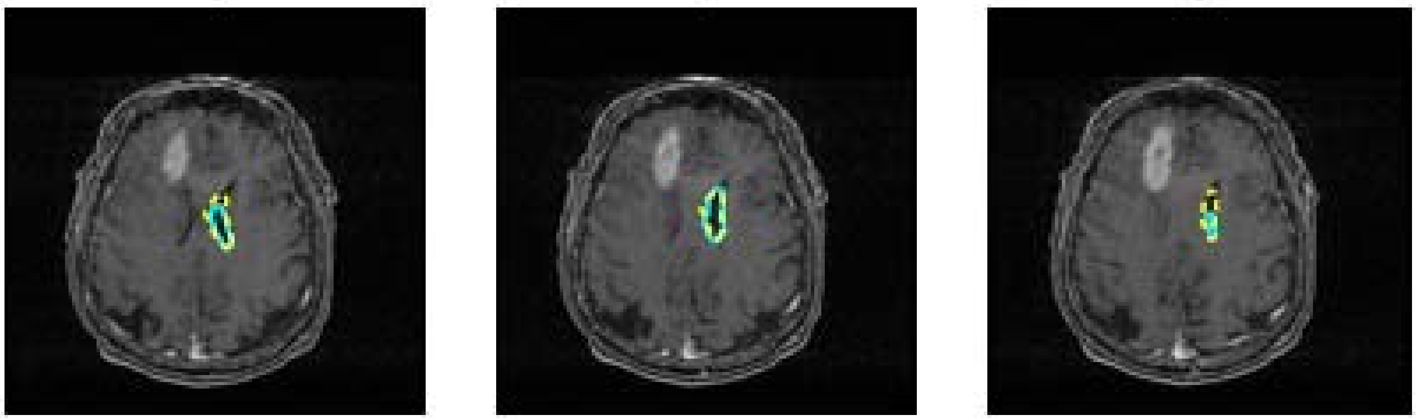
Car left of pole

Advantages over PF-MT-Affine

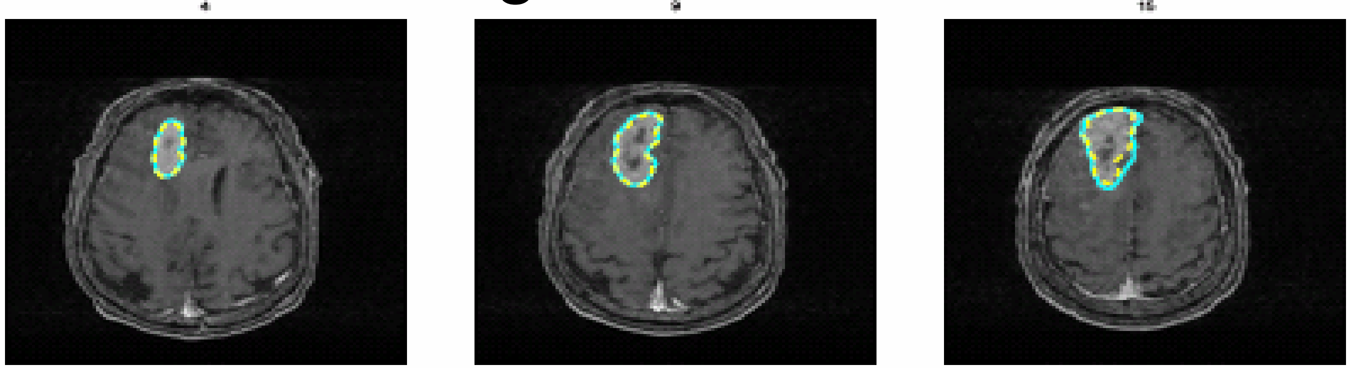
- “Rest of deformation”: much smaller, lesser GD iterations needed even for large non-affine deformations
- Handles two OL modes at “similar” affine “location”
- Back in track easily after outliers: able to return to correct mode

Ongoing Work

Brain MRI: Tracking the right ventricle



Brain MRI: Tracking the tumor



Choosing *K*: Ongoing Work

- Choose *K* large enough s.t. expected residual deformation, Δ_K, small enough to ensure unimodality of $p(C_t|\hat{C}_t, Y_t)$, i.e. $\Delta_K < \Delta^*$

- Problem: Δ^* also depends on *K*, so run an iterative procedure to find *K*

Choosing K: Ongoing Work

- Given a training sequence of contours
  - Learn a K_0-dim approx to deformation, $v_{t,s}$, $\forall t$
  - Learn Δ_{K_0} = expected residual deformation
  - Using Y_t, compute Δ_{t,K_0}^* at each t
  - If $\Delta_{K_0} > min_t\, \Delta_{t,K_0}^*$, set $\varepsilon = min_t\, \Delta_{t,K_0}^*$
    - Find $K>K_0$ large enough so that $\Delta_K<\varepsilon$: K exists because of countable orthogonal basis assumption
  - $\Delta_{t,K}^*$ non-decreasing with K: $\Delta_{t,K}^* \geq \Delta_{t,K_0}^* > \Delta_K$

Basis Change: Ongoing Work

- Assuming *K* is piecewise constant with time
- Change K when *L* changes or f_{max} changes
  - Require methods to detect this & estimate new K
- Non-uniformly allocate control points based on prior information: space varying f_{max}
- Issues related to stability of PF algorithm when delay or errors in estimating new basis

Future Work

- Basis to parameterize velocity w.r.t. contour arclength or w.r.t. x-y location?
  - Parametrizing velocity w.r.t x-y locations handles topology change, but velocities may not satisfy comparison principle

- Basis change detection & estimation
- Proving that Δ_K^* is non-decreasing in K
- Choosing an orthogonal basis?