Deformable Contour Tracking

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

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

  – Do this offline (“smoothing”) or online (“tracking”)

  – Output of a tracker is input to a smoothing algorithm
Brain MRI slices: Tumor sequence (actual deformations)

Partial occlusion of car by street light

Perspective effect: Plane tracked by UAV (frequent camera viewpoint changes)
Outline

• The Tracking Problem
• Main Issues & Limitations of Existing Work
• Our Key Ideas
• Affine PF-MT & Deform PF-MT
• Summary & Open Issues
Tracking Framework
Definitions

- **State transition prior (STP):** $p( X_t | X_{t-1} )$
  - PDF of $X_t$ conditioned on a value of $X_{t-1}$

- **Observation Likelihood (OL):** $p( Y_t | X_t )$
  - Probability of $Y_t$ taking a certain value conditioned on a value of $X_t$
    - “OL multimodal”: OL has multiple local maxima (modes) as a function of $X_t$
Main Issues
Main Issues

- **Observation Likelihood (OL) is often multimodal**
  - e.g. clutter, occlusions, low contrast images
  - If STP narrow enough, posterior is unimodal: adapt KF
  - If STP broad (fast deforming sequence): require a Particle Filter (PF)

- **Deforming contours: Large dim state space (LDSS)**
  - If constrained motion, e.g. rigid/affine: easy to use PF
  - LDSS: PF expensive (requires impractically large N)
Narrow prior: Unimodal posterior

Broad prior: Multimodal posterior
Low contrast + deforming contours (large deformation per frame)

Low contrast + Frequent viewpoint changes (small non-affine deformation per frame)
Overlapping clutter (light grey object) + deformation

Separate clutter (multiple fishes) + deformation
Multimodal OL

• As a function of affine deformation
  – Background clutter due to separate objects
  – Background clutter due to concentric contours

• As a function of non-affine deformation
  – Low contrast (weak edges: multiple edge responses)
  – Overlapping background clutter
  – Partial occlusions
  – Outlier noise
Deforming contours

- Actual deformations: biological images
  - Humans: surveillance, sports videos, ...
  - Animals
  - 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
Existing Work: Limitations

- Different trackers & diff contour representations

- **Condensation** [IB, ECCV’96]
  - Used a particle filter to track multimodal posteriors, but
  - Tracked only on a small (6) dimensional space of affine deformations

- **Approx. linear observers + level set repr.** [JYS,’04]
  - Handled large dimensional deformation, but
  - Did not handle multimodal posteriors
  - Uncoupled observers for global & local deformation
Proposed Solution [Rathi et al, CVPR’05]

• Use level set representation of contour

• Replace the linear observer by a particle filter
  – Can handle multimodal posteriors & defines an asymptotically optimal coupled observer
  – Problem: Deforming contours large dimensional: PF expensive
  – Solution: Imp sample only on a smaller “effective basis” space. Track posterior mode on the “residual space”
Particle Filter (PF) [Gordon et al,’93]

• A sequential Monte Carlo technique to approximate Bayes’ recursion for computing the posterior
  \[ \pi_t(X_{1:t}) = p(X_{1:t}|Y_{1:t}) \]

• Does this sequentially at each time, 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 \( X \), when \( X \sim p \))

• **Monte Carlo:** Sample \( X^i \sim p \)
  \( E_p [\gamma(X)] = \int_x \gamma(x) \ p(x) \ dx \approx (1/N)\sum_i \gamma(X^i) \)

• **Imp Sample:** If cannot compute or sample from \( p \):
  Choose a \( q \), sample \( X^i \sim q \)
  \( E_p [\gamma(X)] = E_q [\gamma(x) \ p(x)/q(x) ] \approx (1/N)\sum_i w^i \gamma(X^i), \)

• **Bayesian Imp Sample:** uses same idea to approx
  numerator & denominator of a posterior expectation
Particle Filter

Goal: To approx \( p( X_{1:t} \mid Y_{1:t} ) \)

- Choose Imp Sampling density s.t. it factorizes as
  \[
  q_{t,Y_{1:t}}(X_{1:t}) = q_{t-1,Y_{1:t-1}}(X_{1:t-1}) \cdot q_{X_{t-1},Y_{t}}(X_t)
  \]
  Allows for recursive computation of weight

- At each \( t \), for each particle \( i \),
  
  - Importance Sample: \( X_{t,i} \sim q_{X_{t-1},Y_{t}}(X_t) \)
  
  - Weight:
    \[
    w_{t,i} \propto w_{t-1,i} \cdot p(Y_t \mid X_{t,i}) \cdot p(X_{t,i} \mid X_{t-1,i}) / q_{X_{t-1,i},Y_{t}}(X_{t,i})
    \]
    
  - Resample to eliminate low weight particles
The Problem is…

- Brute force particle filtering on space of deforming contours (very large dimensional space): very expensive
  - “effective particle size” reduces with dimension: need more particles for the same tracking accuracy
Key Ideas of Our Work
Key Idea 1: “LDSS” [Vaswani et al, ICASSP’06]

• Even though space of contour deformation is very large dimensional,
  – At any given time, most of the deformation occurs in a small # of dims (effective basis) while the deformation in the rest of the dims (residual space) is small
    – Different from dimension reduction (PCA) assumption
  
  – Effective basis dim can change with time
Key Idea 2: “Unimodality” [Vaswani, ICASSP’07]

• If residual deformation small enough (its STP narrow enough) compared to distance b/w OL modes, can show that the “residual posterior” is unimodal
  – “residual posterior”: posterior of residual deformation conditioned on effective basis states

• Ensure this by choosing enough dims as part of effective basis
Key Idea 3: “IS-MT” [Vaswani, ICASSP’07]

• If residual deformation still smaller, the residual posterior is unimodal & also narrow

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

• Resulting algorithm is called PF-MT
Affine PF-MT
Affine PF-MT [Rathi et al, CVPR’05, PAMI, Aug’07]

• Contour represented using level set method

• Use Importance Sampling to track on the 6-dim space of affine deformations (effective basis)

• For each affine deformed contour particle, track the unique mode of the posterior of non-affine deformation (residual space): Mode Tracking (MT)

• Very efficient: Imp sampling dim was only K=6, small N sufficed for given accuracy
Affine PF-MT assumes

- Non-affine deformation per frame small enough, s.t. its residual posterior is unimodal & “narrow enough” to replace IS by MT
  - Much weaker than assuming the same thing for the full posterior

- This is satisfied whenever either
  - Small non-affine deformation per frame
  - or
  - OL modes (contour modes in image) separated only by translation or scale or other affine deformation
Low contrast, viewpoint changes

- Deformation due to perspective camera effects (changing viewpoints), e.g. UAV tracking a plane

Condensation fails

Affine PF-MT works
Condensation (30 & 1200 particles)

Affine PF-MT (30 particles)

Condensation only tracks affine deformation
Assumption fails when...

- Large non-affine deformation per frame (fast deforming sequence) & OL multimodal as a function of non-affine deformation (low contrast images, overlapping clutter or partial occlusions)

- Results in multimodal residual posterior of non-affine deformation
Small non-affine deformation per frame: Affine PF-MT works

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

Deform PF-MT works
Deform PF-MT
Deform PF-MT [Vaswani et al, CDC’06]

• For multimodal posterior of non-affine deformation, need an importance sampling step in PF that also samples on space of local deformations
  – This is a very large dim space: regular PF inefficient

• Again use PF-MT but with deformation at a subsampled set of K contour points (basis points) & translation as the effective basis

• $K = \# \text{ of basis points}: \text{ fixed or time varying}$
Multiple Issues

• Imp. Sample + Level Set Rep of contour: CFL?
  – Interpolate & compute extension velocity efficiently
  – Using multiple iterations to deform a contour: slow

• How to parameterize contour deformation?
  – As a function of arclength: Expensive to implement using level sets, Cannot handle topology change
  – As function of radial/tangent angle: Fails if 2 contour points far along arclength are close along angle

• Estimate/Change effective basis dimension?
  – Use spatial freq response of deformation “signal”
Low contrast images, large def per frame: Brain MRI (Tumor, Ventricle)

- Multiple nearby modes due to low contrast

(b) Attempt to track the right ventricle (black region in the center) using Algorithm 2. Notice the low contrast imagery.
Outlier: multiple nearby modes

- Every even frame: outlier frame
- Multiple nearby modes separated by non-affine deformation

Affine PF-MT fails

Deform PF-MT works
Relation to Other Work

- **PF-MT**
  - Extension of PF-Doucet [Doucet’98]
  - Approx to Rao-Blackwellized PF [Chen-Liu’00]

- **PF for tracking Heart LV** [Sun et al, MICCAI’04]
  - PF-MT with PCA effective basis + retaining MAP particle

- **Approx linear observer + level sets** [JYS, CDC’04]
  - PF-MT with zero dimensional effective basis

- **Condensation** [IB, ECCV’96]
  - PF-MT with zero dimensional residual space

- **Stochastic Active Contours** [JPK, VLSM’04]
  - Annealing for segmentation
Summary & Open Issues

- **Affine PF-MT**
  - Use when small non-affine deformation per frame

- **Deform PF-MT & Smoother: Fixed or changing K**
  - Human body contour tracking
  - Heart LV, Brain ROIs, Lung ROIs

- Observation models, tracking intensity variations
- PF smoother: better estimate for offline problems

- **Level Set Rep. + Imp Sampling: speedup, CFL**
- **Extensions to surface tracking**
Collaborators

• Affine & Deform PF-MT
  – Yogesh Rathi, Anthony Yezzi & Allen Tannenbaum at Georgia Tech

• System Identification
  – Ongoing work with grad student, Wei Lu
Other applications of PF-MT

• Spatially varying illumination change of moving objects
  – Moving into lighted room, face tracking [Kale et al, ICASSP’07]
  – Vehicle tracking through changing illuminations

• Change in spatially varying physical quantities using sensor networks
  – Tracking temperature change [Vaswani, ICASSP’07]

• Deformations of shapes of landmark points using the nonstationary shape activity model
Illumination Tracking: PF-MT [Kale et al’07]

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

- PF on motion (3 dim) & MT on illumination
  - Illumination change very slow
  - OL usually unimodal as a function of illumination
  - If OL multimodal (e.g. occlusions), modes usually far apart compared to illumination change variance
Face tracking results

- PF-MT
- 3 dim PF (no illum)
- 10-dim Auxiliary PF
Error from ground truth

Comparing with 10 dim regular PFs (original, Auxiliary) & with PF- K dim (not track illumination at all)
Sensor nets: Temperature tracking

- \( \text{Dim}(X_t) = 10 \)
- \( K = 1, \) i.e. \( \Delta_s = 10, \Delta_r = 1, \) & OL multimodal
- \( N = 50 \) particles
- Plotting RMSE from ground truth
- PF-MT better than all full PFs (PF-EIS, PF-D) & PF-K dim (dim reduced PF)
Landmark Shape Tracking

- Tracking deformations of shapes of landmark points using nonstationary SA (NSSA) model

- NSSA better models larger & nonstationary shape changes than existing methods (ASMs)
  - Existing ASM work uses piecewise ASMs to track long sequences, e.g., separate ASM for systolic & diastolic heart motion, or hierarchical ASMs
  - Cannot model transitions b/w pieces very well
  - Cannot detect change while tracking
Landmark Shape Tracking
Landmark Shape Tracking

• Compared modeling error of our method (NSSA) with Active Shape Models for CMU MOCAP dataset (human action sequences)

• For all sequences, modeling error of our method much smaller than ASM
Modeling Error Comparison

• Defined 10 dimensional PCA space for ASM and for shape velocity (our method)

• Defined AR model for ASM & for shape velocity. Total modeling error
  
  Crawl:  ASM: 0.00870, Shape Velocity: 0.00030
  Sit:  ASM: 0.00760, Shape Velocity: 0.00005
  Interview: ASM: 0.00450, Shape Velocity: 0.00020