Recognition of Moving Patterns from Moving Observers

Rama Chellappa

Center for Automation Research

And

Electrical and Computer Engineering

University of Maryland

Issues

- Simultaneous object and observer motions.
- Illumination and pose variations.
- Behavior-based recognition is possible.
- Recognition of a rare event.
- Time is of the essence!
- Unambiguous ID of objects and activities possible
- Motion – Shape – Recognition- How they influence each other
- Basic survival skill

Opportunities

- Persistent observations – continuous sensing, tracking and identification.
- Peek-a-boo, I see you – Surprising an object by an observer (Active sensing).
- Behavior modeling (Dynamic Bayesian networks).
- Advances in structure from motion.
- Old tools (from the 50's) in modern computers (MCMC techniques).
- Role of invariants/quasi-invariants understood better.

Simultaneous Modeling of Object and Observer Motions

- Is there one or more combinations of observer and object motions for best recognition ?
- Keeping the object in observer's cross wire.
- The problem of critical motion sequences
  - Motions that result in many reconstructions

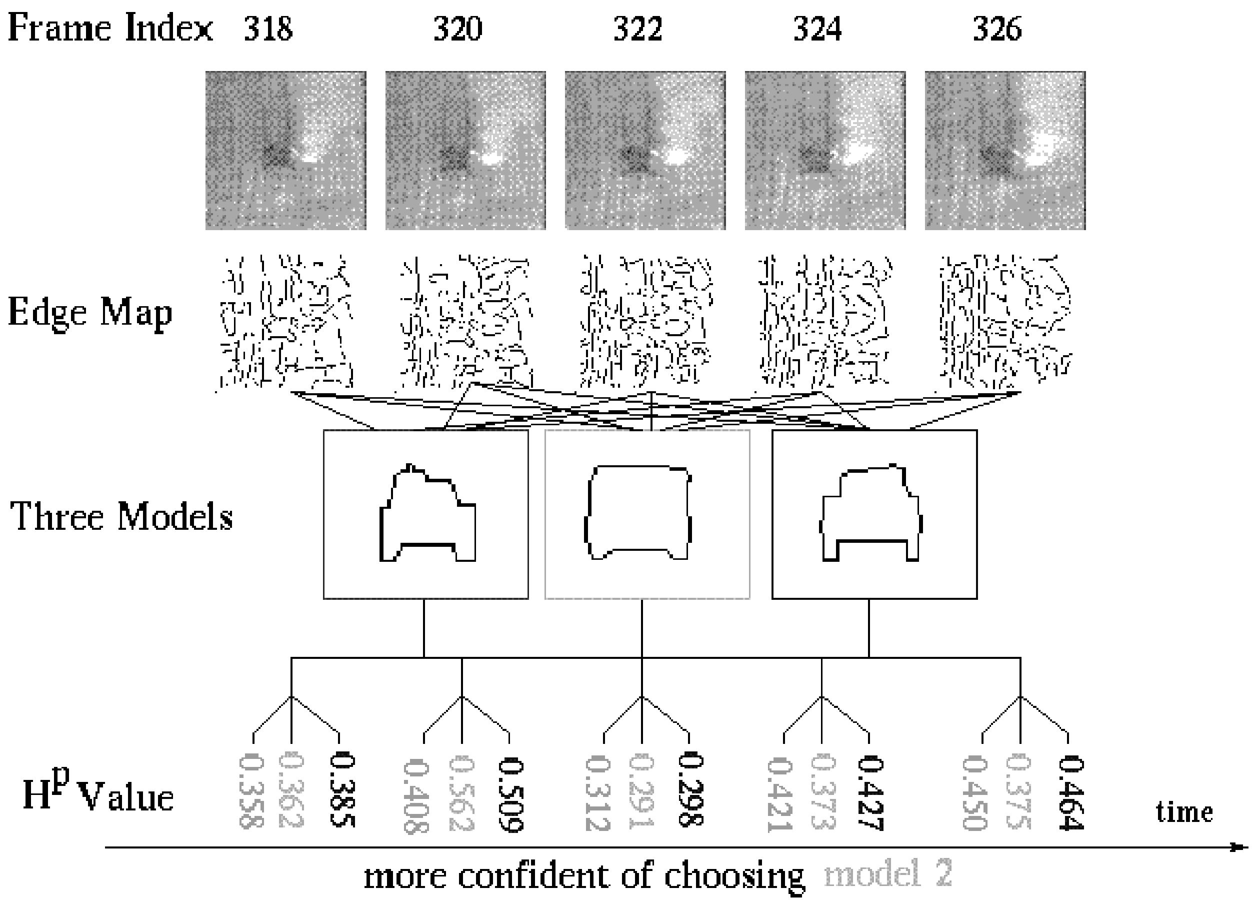
Continuous Tracking and Recognition -1

- “Fingerprinting” the object. What and how many ?
- How to accumulate evidence as the object is persistently observed ?
  - Evidence accumulation may have to be non-monotonic
- Not all objects are of equal interest – costs of incorrect decision unknown and non-uniform across the object space.
- How many frames are needed for achieving the required performance ?
  - Need to generalize Wald’s Sequential Probability Ratio Test

Continuous Tracking and Recognition -2

- Temporal Identification/Verification
  - Is this what I saw the last time?
  - How confident am I?
  - Recovery after temporary loss of tracking
- Hausdorff Metric
  - Measures the similarity between point sets
  - Using edge maps as input
  - Account for residual motion

Continuous Tracking and Recognition - 3



Continuous Tracking and Recognition -4

- MCMC (static) : a general method
  - Monte Carlo integration + Markov Chain
- Sequential importance sampling (SIS): dynamic MCMC
  - assuming the system is varying slowly
  - random samples at time *t* will be re-used at time *t+1*
  - have some good properties of both importance sampling and MCMC
  - a recent tool for analyzing non-Gaussian/nonlinear systems (e.g. Carlin *et al*.'92, Kitagawa '96, Gordon *et al*. '93, Isard *et al*. '96, Liu and Chen, 1998, Liu 2001)

Marginalization Over Motion

- Gallery: $H = \{I_1, I_2, ..., I_N\}$
- Identity variable:

  depicting the identity at time instant t n_t
- Motion state:

  characterizing the dynamics at time instant t

$$\theta_t \quad \Theta$$

Time Series Model for Recognition (II)

- Prior distribution:

$$p(\theta_o \mid y_o) \quad p(n_o \mid y_o)$$

- Noise distribution:

  $p(u_t)$ or state transition prob. $p(\theta_t \mid \theta_{t-1})$

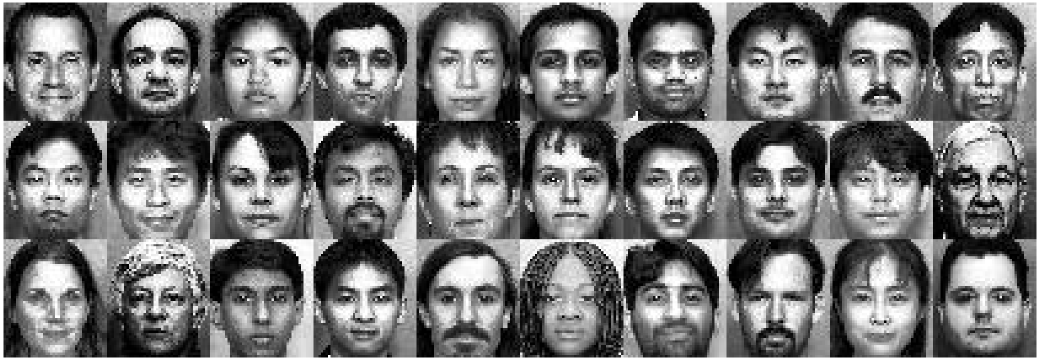
  $p(v_t)$ or likelihood $p(y_t \mid n_t, \theta_t)$

- Statistical independence:

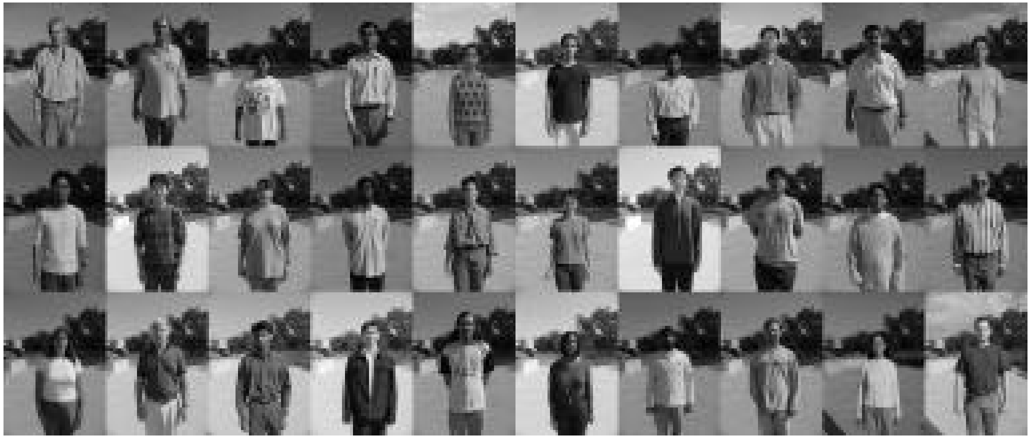
$$n_0 \quad \theta_0, u_t \quad u_s, v_t \quad v_s, t \neq s$$

Galleries

Face Gallery



Upper Body Gallery



Tracking results




Continuous Tracking and Recognition - 5

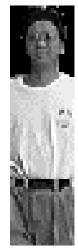
- Original sequence

- Tracking results for true hypothesis

- Tracking results for false hypothesis

Continuous Tracking and Recognition -6

- Back-door camera captured a person entering a building



Continuous Tracking and Recognition - 7

Front-door video

Structure from Motion:Help or Hindrance

- Flow-based method.
  - 43 FFT computations for structure recovery using two frames.
  - Fusion of depth maps using error covariance and stochastic approximation.
- Discrete feature based method using SIS.
- Tango between structure and motion – Critical motion sequences.
- How much of shape is needed for recognition ?
- Motion to recognition without explicit structure recovery?

Bayesian SfM Algorithm

- A single solution is not enough to describe the solution space of the SfM problem.
- To find the posterior distribution of the parameters is a better way.

$$\pi_t(\chi_t) = P(\chi_t \mid Y_t)$$

$$= P(\{x_\tau\}_{\tau=1}^{t} \mid \{y_\tau\}_{\tau=1}^{t})$$

- Monte Carlo methods approximations to the posteriors.

SIS for SfM

$$\{\chi_t^{(j)}, w_t^{(j)}\}_{j=1}^m \sim \pi_t(\chi_t) \xrightarrow{SIS} \{\chi_{t+1}^{(j)}, w_{t+1}^{(j)}\}_{j=1}^m \sim$$

SIS steps: for $j = 1, \cdots, m$,

(A) Draw $X_{t+1} = \mathbf{x}_{t+1}^{(j)}$ from $g_{t+1}(\mathbf{x}_{t+1} | \mathcal{X}_t^{(j)})$. Attach $\mathbf{x}_{t+1}^{(j)}$ to form $\mathcal{X}$

(B) Compute the "incremental weight" u_{t+1} by

$$u_{t+1}^{(j)} = \frac{\pi_{t+1}(\mathcal{X}_{t+1}^{(j)})}{\pi_t(\mathcal{X}_t^{(j)}) g_{t+1}(\mathbf{x}_{t+1} | \mathcal{X}_t^{(j)})}$$

and let $w_{t+1}^{(j)} = u_{t+1}^{(j)} w_t^{(j)}$.

$$g_{t+1}(x_{t+1} | \chi_t) = q_t(x_{t+1} | x_t)$$

$$u_{t+1} \quad p(y_{t+1}$$
$$= p(y_{t+1}$$

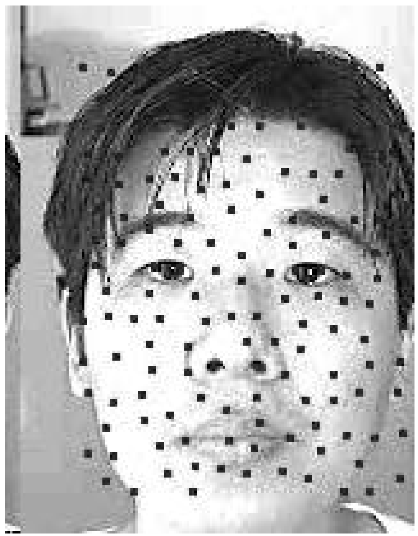
3D Modeling from Video

- A two-step strategy

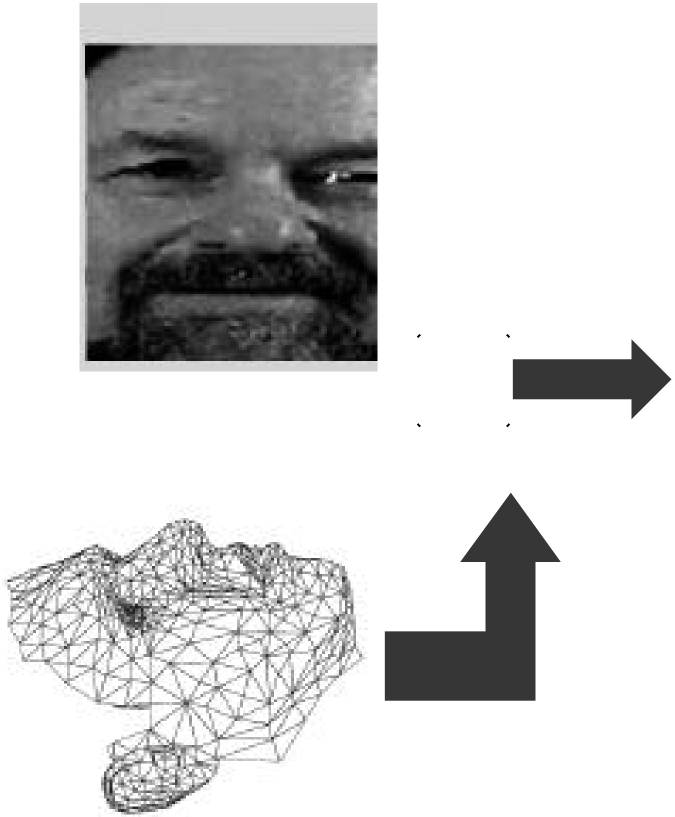
Sequential importance sampling

Feature trajectories

3D Face Modeling



Incorporating 3D Models



Quasi-invariants for Activity Recognition

- Most common human actions involve several body parts that align themselves in a plane: e.g. walking, running, sitting down, waving, jumping, etc.
- Pick phases in the action where a subset of body parts align themselves approximately in a plane.
- Pre-compute a pair of determinant form of invariants for each such ‘canonical’ pose in the action.
- Distributed sensors + overlapping quasi-invariants = Full invariants ?

Behavior-based Recognition

- Based not on who or what you are, but what you have been doing or what you intend to do.
- Knowledge about origin and destination may be useful.
- Modeling and recognition of abnormal activities.
- Dynamic belief networks (DBN)
  - Inference mechanisms understood for simple structures.
- DBN's with MCMC , a promising approach (Koller, 2001)
- How to incorporate source, sink priors and how long should one observe ?

Recognition of a Rare Event

- Normalcy distribution can be estimated.
- The rare event is a deviation from the normal hypothesis.
  - Outlier detection in an incompletely specified Markov process is hard.
- The FAA problem – monitoring the activities around an aircraft.
- Dictionary of normal activities.
  - Trash removal, catering services, fuel tank, baggage carts, a few authorized people, retraction of the jet way, etc.
- How to learn something that is rare ?

My 2 Cents!

- Representation for recognition in video is the key
  - Trade-off between shape and motion should be understood.
- Being able to learn from sparse samples is important.
- Detection of a rare event is critical.
- DBN + MCMC has potential
  - I love particle filters, but they are NOT the next best thing after the sliced bread!
- Scalability of stochastic grammar representation to video is a challenge