Recent Advances in Recognition of Humans Using Gait

Rama Chellappa

University of Maryland, College Park

Introduction

- Objective: Recognition of humans from their gait in video.
- Preprocessing to obtain an observation vector from each frame of video.
- Appearance-based modeling of Gait
  - Dynamic-time warping
- Model based approach
  - Hidden Markov Models (HMMs) for each person in the Gallery, using training video sequences.
  - Captures structural and dynamic features of the observation vector in the HMM.

Related Work

- Gait modeling using ellipse fitting – Lee & Grimson.
- Spectral partitioning framework to extract gait shapes – Tolliver & Collins.
- Dynamic Time Warping (UMD, MIT).
- USF Baseline Algorithm – Sarkar, Phillips et al.
- **HMM (UMD).**
- Visual Hull - Shakhnarovich, Lee et al.
- View-Invariant Recognition from Monocular Video – Kale, RoyChowdhury & Chellappa.

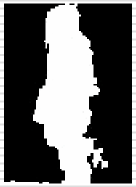
Preprocessing

Video Sequence

Background Subtraction

Actual Video Sequence

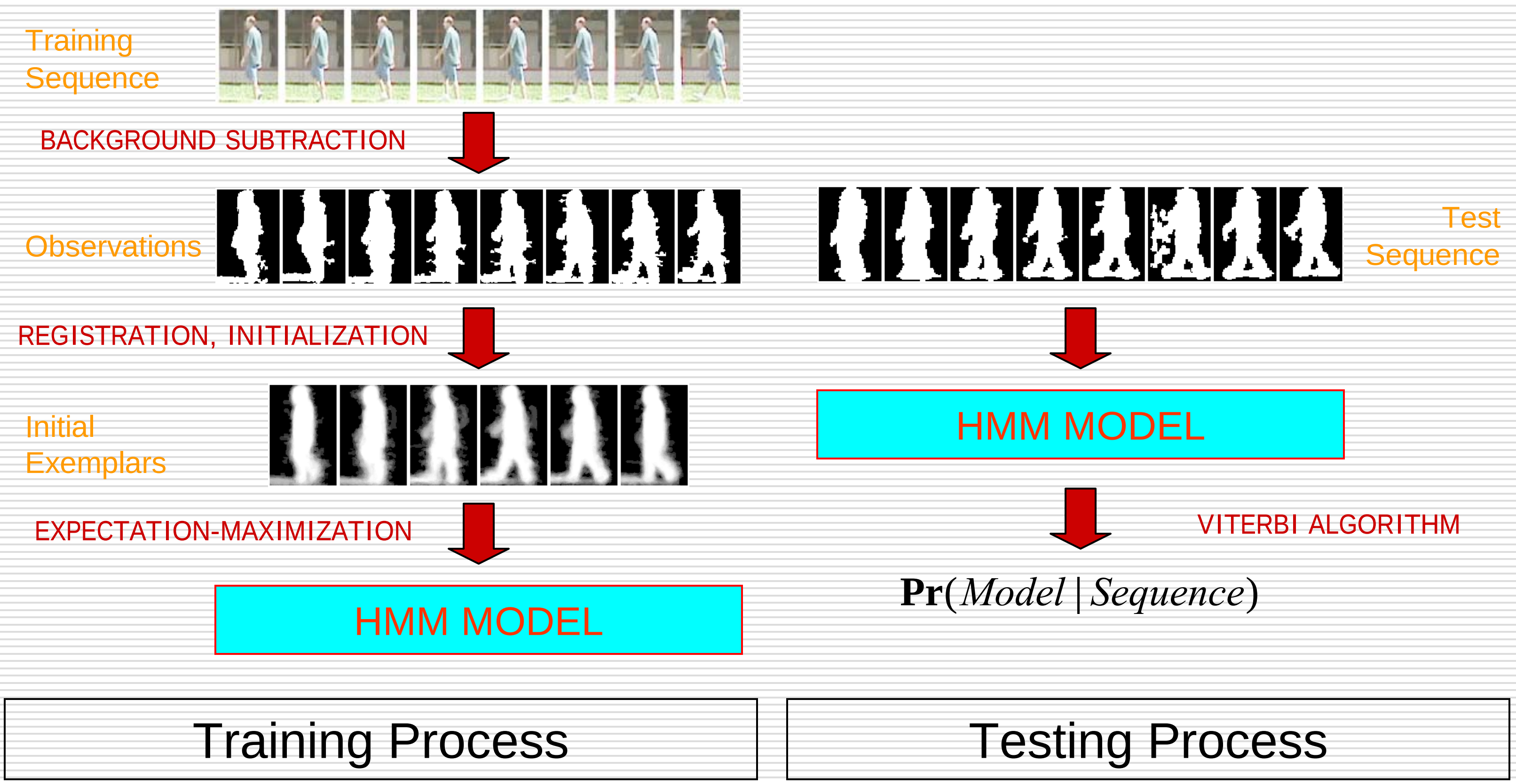
Binary Silhouettes



Background Subtracted Video Sequence

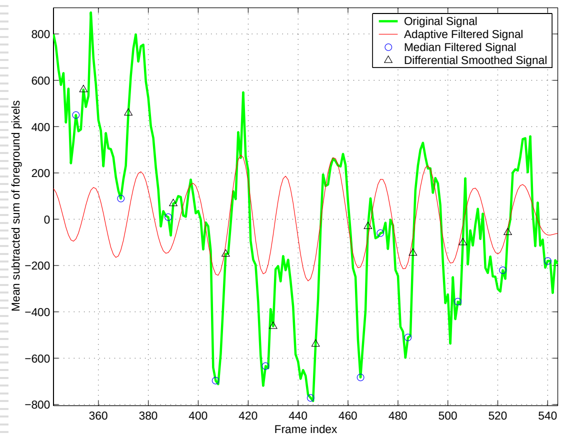
- The HMM framework is generic - we can use any feature vector as the observation.

Overview of Recognition Algorithm

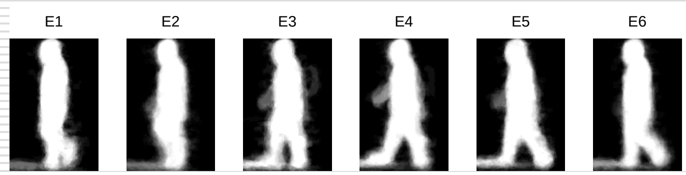


Computing Exemplars

- Plot number of foreground (white) pixels versus time.
- Identify “cycles” by using band-pass filters which admit frequencies corresponding to that of gait.
- Divide each cycle into 6 adjacent regions to obtain clusters.
- Group together observations from corresponding clusters of all cycles.
- Take an “average” to obtain an initial estimate of the exemplars for each state.



Estimation of cycle boundaries



Initial Estimate of Exemplars

The HMM Framework

- Given the observation sequence, $O=\{O_1, O_2, \ldots, O_T\}$, we would like to find a parameter set $\lambda=(A,B,\pi)$ that maximizes the posterior probability, $Pr(O|\lambda)$.
- A is the transition matrix.
- B is the probability of an observation conditional on the state index.
- π is the probability distribution of the first observation.
- The probability of the observation given the exemplar (state) is an exponential function of the distance between the observation and the exemplar.

The HMM Framework (2)

- We denote the distance between observation and Exemplar as $D(O_t, E_j)$.
- The probability of the observation given the exemplar (state), $P_j(O_t)$, is an exponential function of $D(O_t, E_j)$.
- We can use an intuitive distance measure, based on the observation vector.

$$D(O,E) = 1 - \frac{O^T E}{\sqrt{O^T O}\sqrt{E^T E}}$$

$$B_j(O_t) = \Pr(O_t \mid j) = \beta e^{-\alpha D(O_t, E_j)}$$

Training

- We start with a predefined value for A, a uniform distribution for π, and the initial estimate of the exemplars.
- The Expectation-Maximization algorithm is used to refine the estimates of the exemplars and A.
- The model parameters usually converge in a few iterations.

Updating Exemplars

$$E_j^{(i+1)} = \arg_E \max \prod_{t \in \{j^{th} group\}} P(O_t \mid E) \Rightarrow E_j^{(i+1)} = \arg_E \min \sum_{t \in \{j^{th} group\}} D(O_t, E)$$

Updating Transition Matrix, A

$$A^{(i+1)} = \arg_A \max P(O \mid (A^{(i)}, B^{(i)}, \pi)) \ \text{(Baum – Welch Algorithm)}$$

Testing

- A sequence X can be identified by finding the HMM parameters (λ_p) from the gallery that maximizes the probability of the observation sequence given λ_p.
- We use the Viterbi algorithm to compute the probability of a sequence given the model.
- The Viterbi algorithm is computationally efficient.

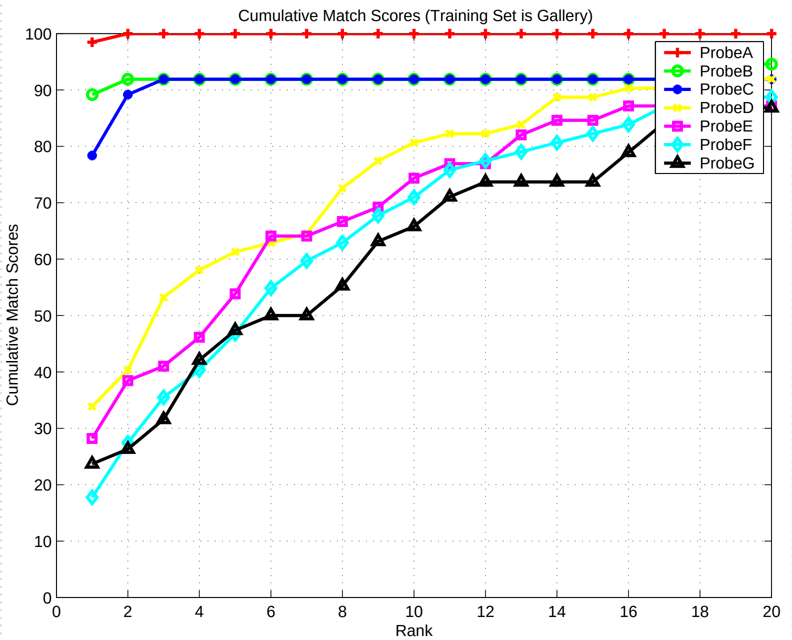
$$ID = \arg_p \max P(X \mid \lambda_p),$$

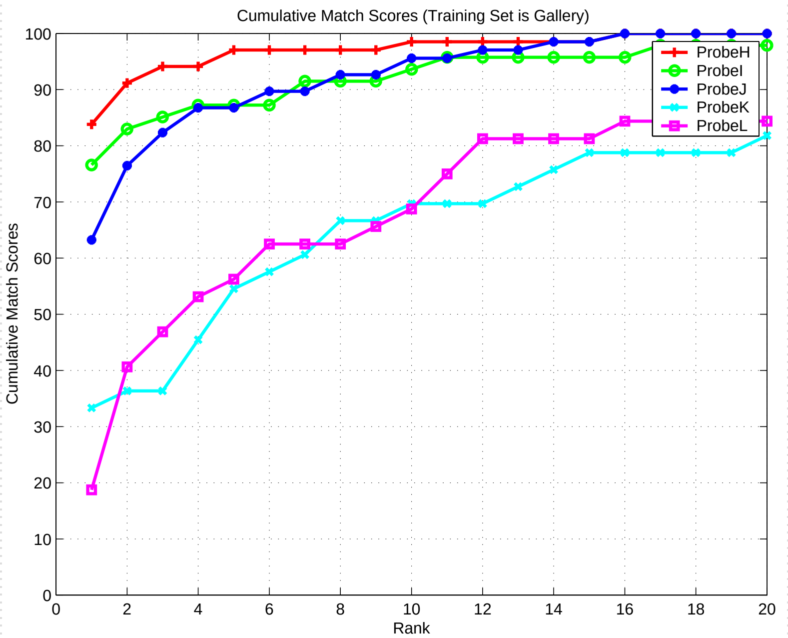
where λ_p is the HMM for p^{th} person.

Results

- The cumulative match characteristics (CMC) plot indicates the percentage of probes for which the correct match figured in the top n ranks.
- The performance of the algorithm with respect to different probe sets from the USF Database.
- Gallery
  - GAL (Grass, Shoe type A and Left camera)
- Probes
  - GAR, GBL, GBR, CAL, CAR, CBL and CBR (C is for concrete surface), with briefcase, time separated

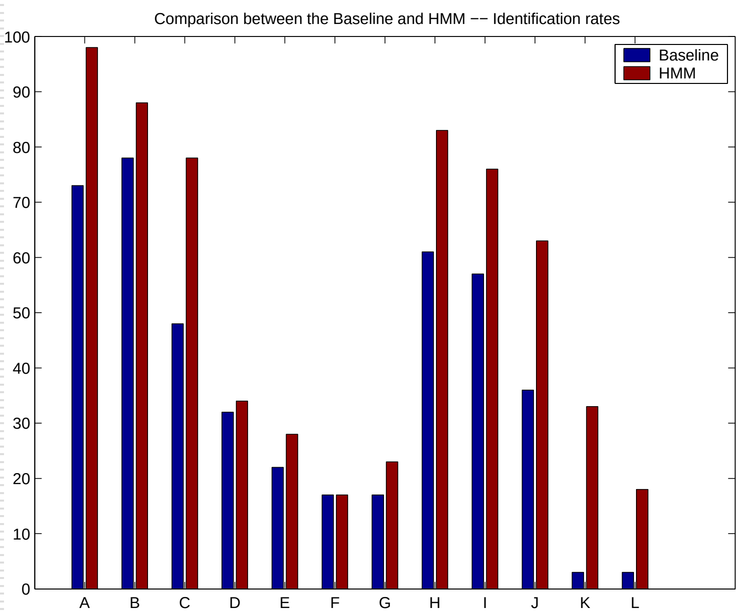
Results


Cumulative Match Scores (Training Set is Gallery)
Cumulative Match Scores
Rank
ProbeA
ProbeB
ProbeC
ProbeD
ProbeE
ProbeF
ProbeG


Cumulative Match Scores (Training Set is Gallery)
Cumulative Match Scores
Rank
ProbeH
ProbeI
ProbeJ
ProbeK
ProbeL

Results (2)

- The bar graph compares the performance of the HMM algorithm (71 persons) with that of the baseline algorithm (USF Baseline Algorithm) (122 persons).

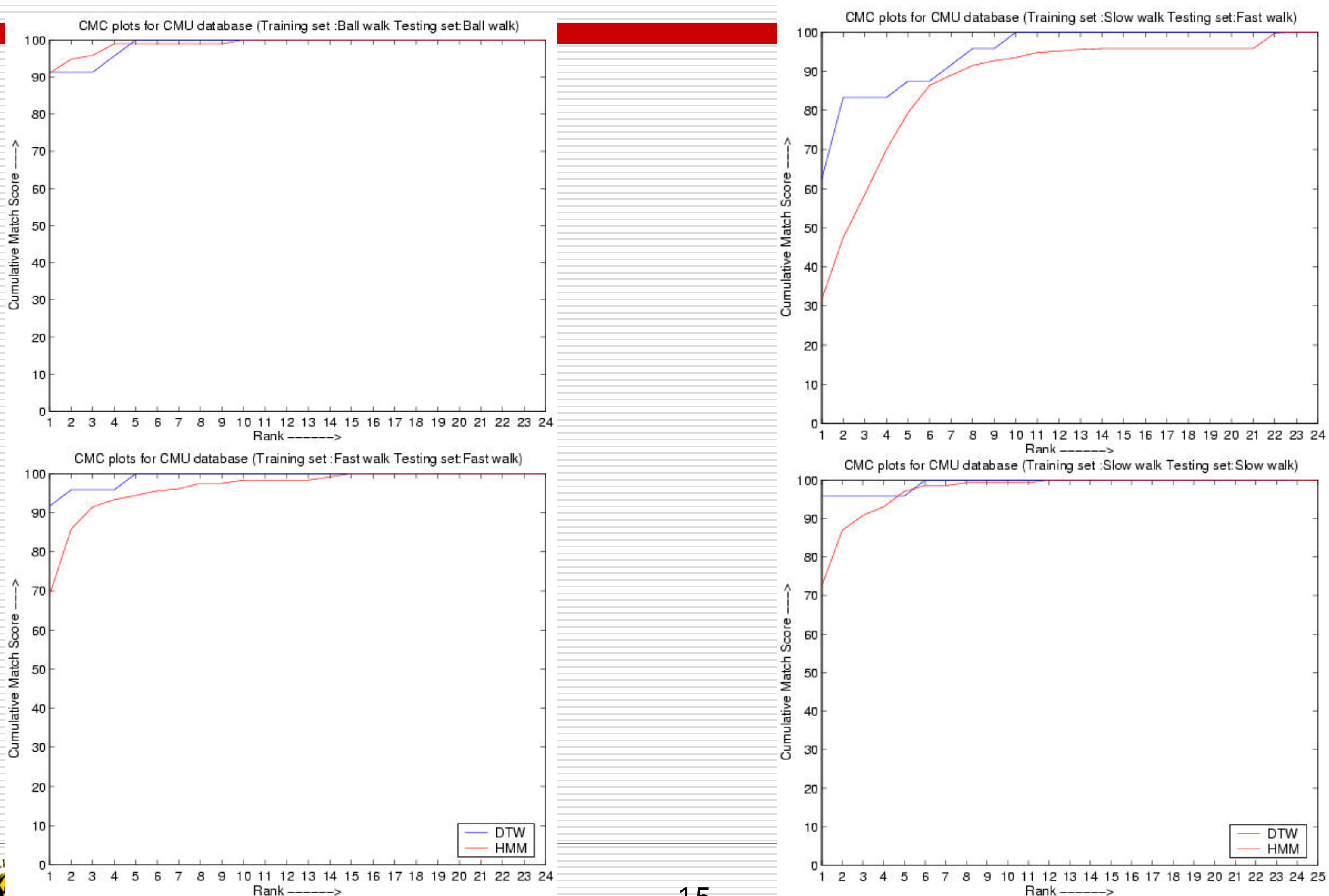


Summary of Round Robin Expts.

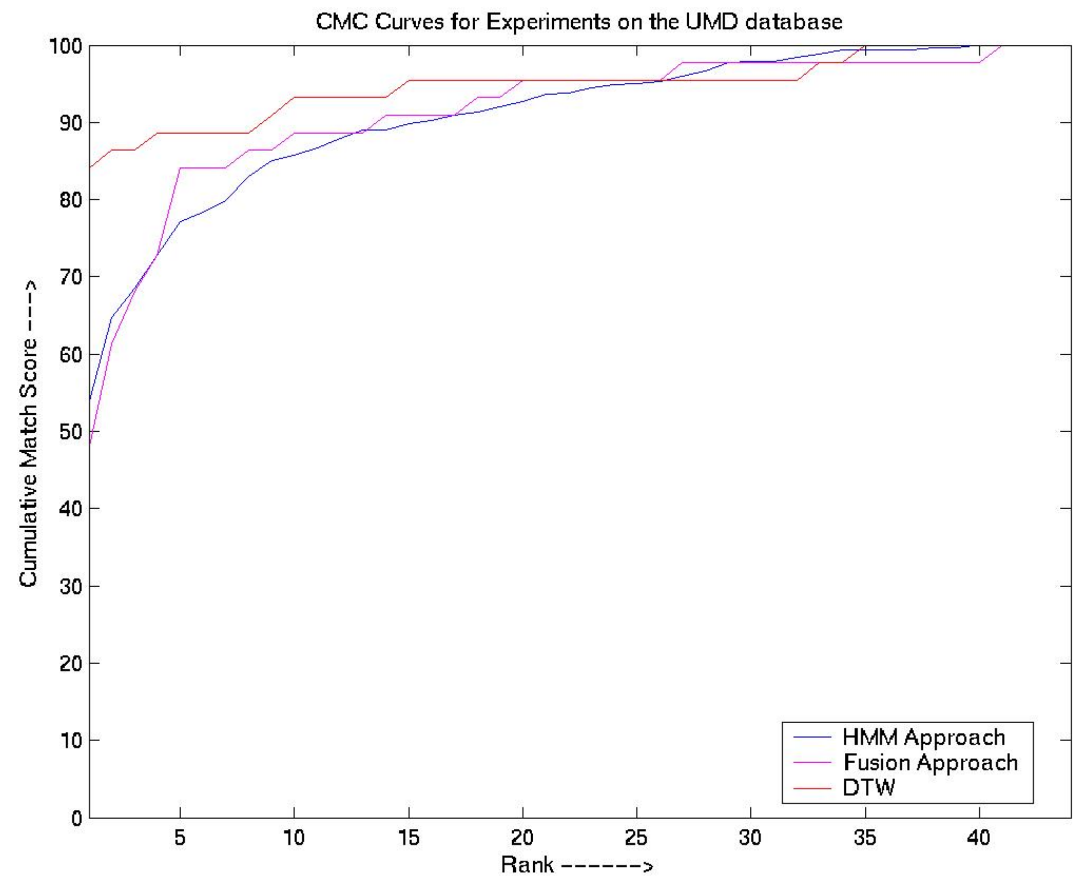
- Generalization across surface type poor

Gallery	GAL	GBR	GBL	CAR	CBR	CAL	CBL
Probe	GBL	GBL	GAL	CBR	CAR	CBL	CAL
Probe	GAR	GAR	GBR	CAL	CBL	CAR	CAR
Probe	GBR	GAL	GAR	CBL	CAL	CBR	CAR

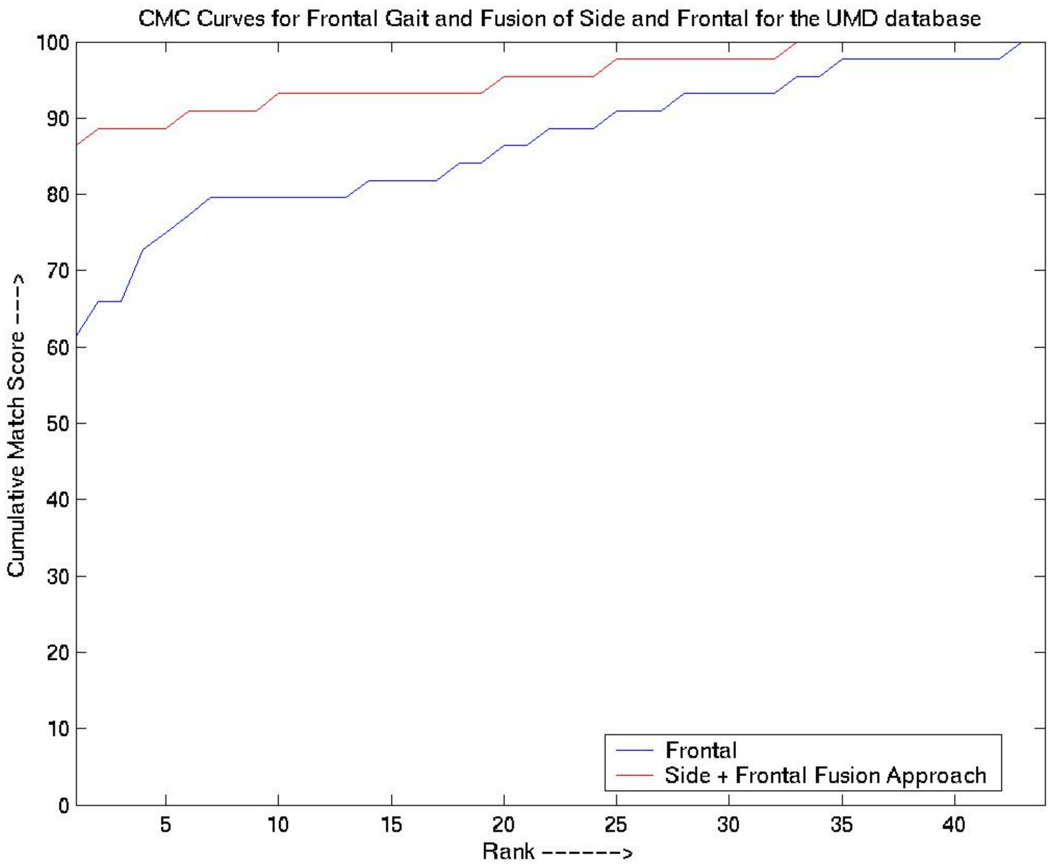
UMD Gait Recognition on CMU Dataset



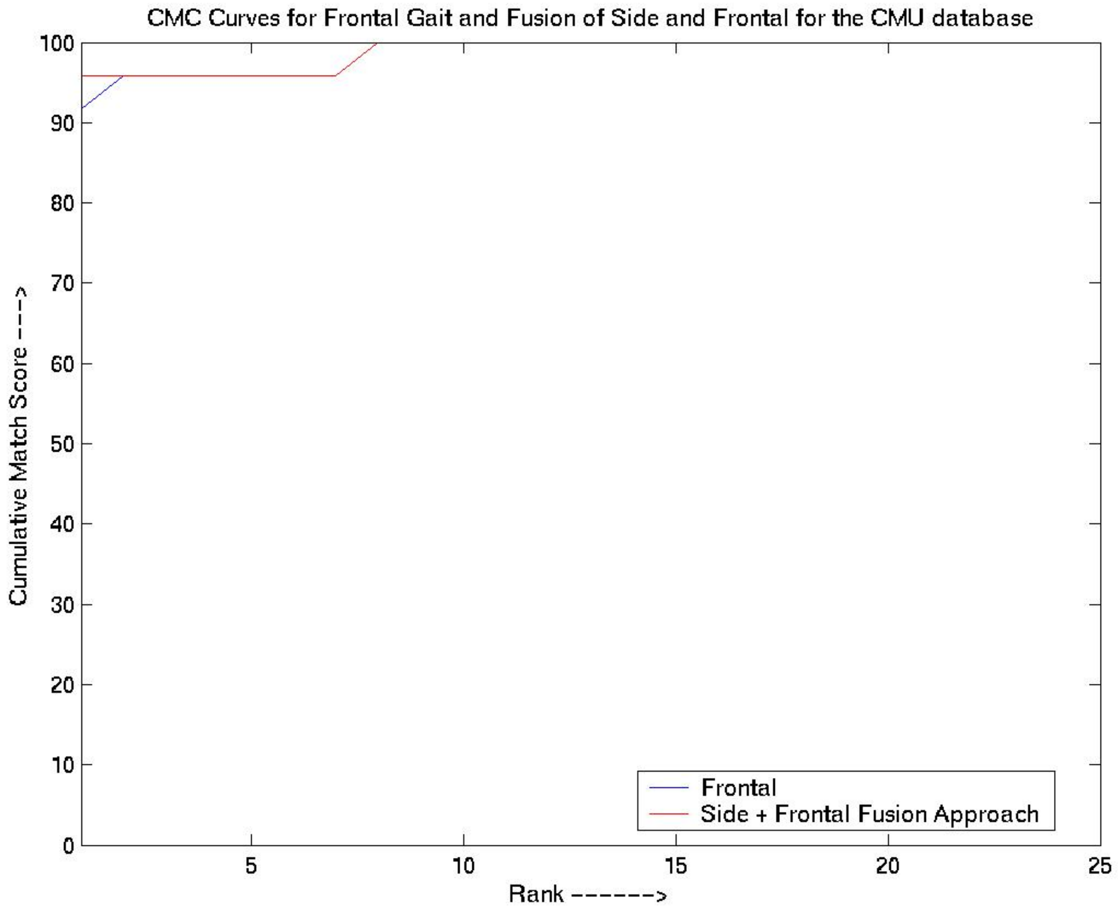
UMD Gait Recognition on UMD2 Dataset



Frontal gait and multi camera: UMD -2



Frontal gait and two-camera: CMU data

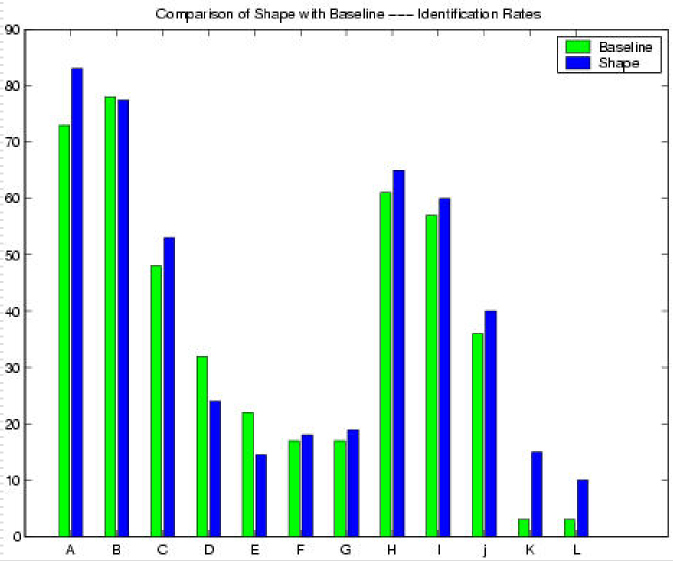


Ongoing work - Shape

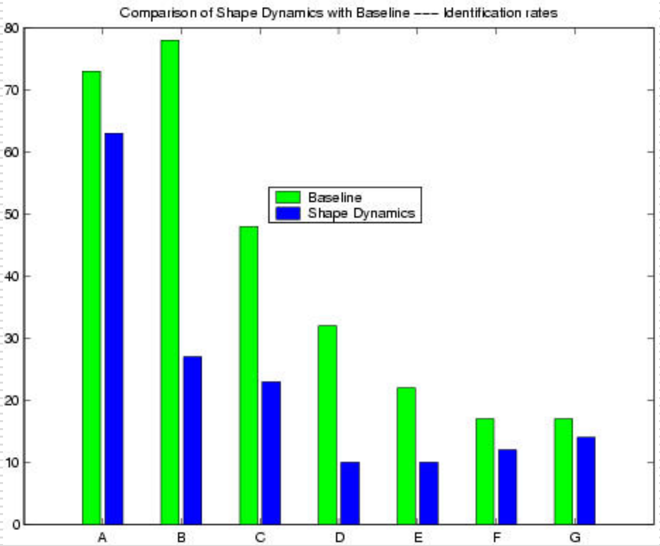
- Kendall’s statistical shape theory used to model gait.
- Procrustes distance used to compute similarity of shape - close to or better than baseline performance.
- Gauss-Markov model on shape dynamics – poor performance.
- Linear shape-dynamical model (fusion of shape and dynamics).

Gait – Shape and/or Dynamics?

Identification using Kendall's Shape



Identification using Gauss-Markov model



- Shape contributes more to gait recognition.
- Is dynamics significant for activity recognition?

Ongoing Work - View Invariant Gait Recognition

- A. Obtain 3D angle using SfM equations. $\cot(\alpha) = B - A\cot(\theta)$

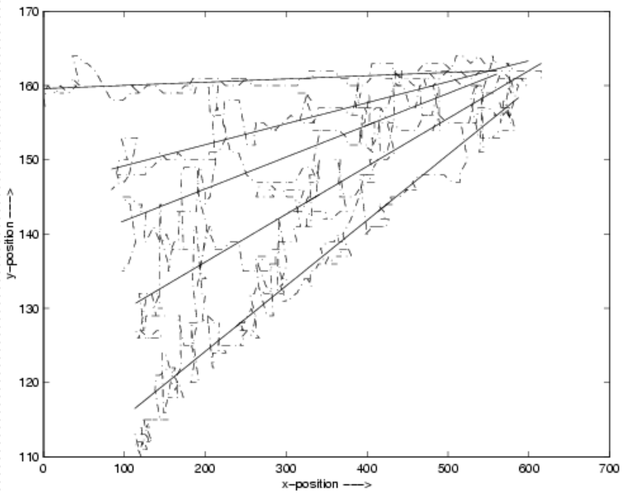
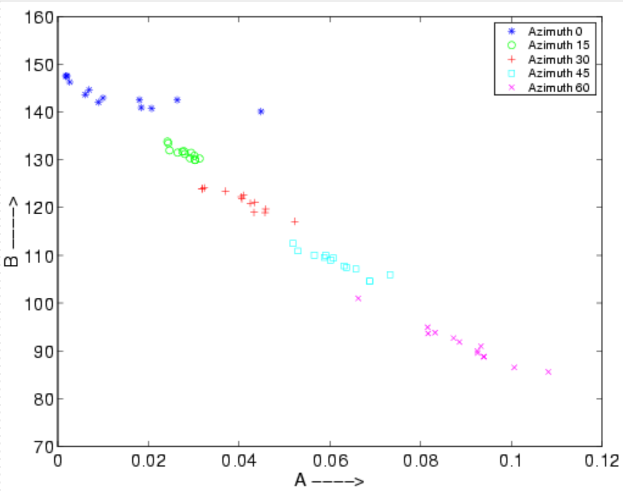


Image Plane Angle (α)

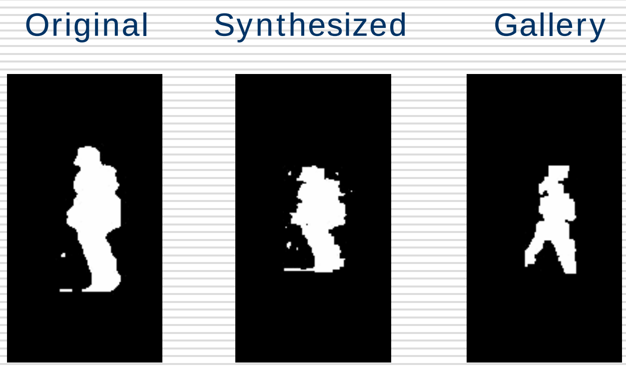


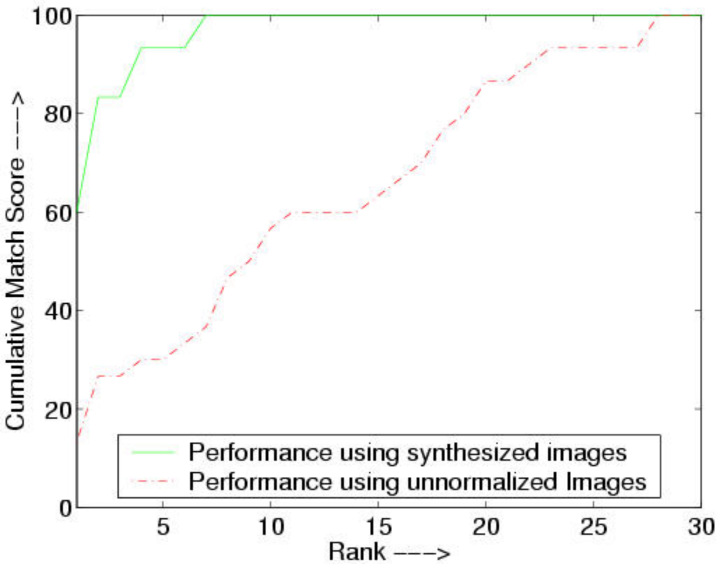
3-D Angle (θ)

- B. Synthesize new view by rotation and if required, correct for planarity assumption.

View Invariant Recognition: Results

- Results on NIST database





Future work

- Use observation vectors that provide better representation and identification.
- Study Kinematic Chain models for human motion.
- Study of the role of Shape and Dynamics in human motion modeling.