Improving Performance for Computer Vision Systems

Dr. Aditya Abhyankar

Dean, Faculty of Technology
Professor, Department of Technology
University of Pune, Pune
Homography based Hybrid Mixture Model for 3D Reconstruction

- True focal length gets converted into apparent focal length in digital camera
- In turn, the true depth information gets mapped to apparent depth information
Homography based Hybrid Mixture Model for 3D Reconstruction

- The reverse transformation needs to recover, **that is to estimate** the depth from focal length.
- The projective transformation that relates the two camera views / images

- An invertible mapping of points or lines on one projective plane ($\mathbb{P}^2$) to another plane ($\mathbb{P}^2$)

- Also called as Collineation, a projective transform or homography

\[
P^2 = H P^1
\]
Experimental Setup -

Background plane

P_1

P_2

P_3

Object

FOV More than 160°

Cam 1

Cam 2

Cam 3

Cam 4

Cam 5

Cam 6
Research: Biometrics

- Ideal and Non-Ideal
- Iris Recognition
- Biometric Security
- Fingerprint Liveness Detection
- Iris Template Aging and Update
- Wavelet Based Scalable Video Coders
- Fingerprint Interoperability
Liveness Detection

Algorithm

Part (I)
Wavelet Based Approach

Part (II)
Perspiration Pattern Characterization

Part (III)
Empirical Mode Decomposition Based Approach

Part (IV)
Fingerprint Ridge and Texture Analysis Approach
Liveness Detection

◆ Hypothesis
  ■ Live fingers demonstrate a specific changing moisture pattern due to perspiration.
  ■ Cadaver and spoof fingerprint images do not.
◆ Algorithm uses two fingerprint images over time
◆ Original algorithm: capacitive DC, 5-second time frame, small dataset
Liveness Detection

Perspiration observed in live images after applying the threshold to the first difference of wavelet-enhanced images.
Iris Recognition

Original Iris Image (a) → Segmented Iris (b) → In-band noise removal (c)

Iris Template (f) → Normalized Iris (e) → Radial Resolution (d)
Iris Recognition - Segmentation

- Segments iris region from the rest of the eye image
- Performs maxima energy extraction
- Number of wavelet coefficients retained = 10,000
- If iris region is not segmented correctly, it cannot be used for doing personal identification
Iris Recognition- Segmentation

- Traditional iris segmentation uses model based transforms like Hough or circular canny
- Appropriate for on-angle images, but fails for off-axis images
- Goal: Develop segmentation which improve iris recognition performance, particularly for non-ideal iris images
ASM Segmentation
- The basic idea behind matching using BWNs
  - \( \text{BWN}(y,w) \) optimized for a particular class of an iris
  - \( \text{BWN}(y,w) \) specific for that particular class only
  - Any other class is not well represented by the same BWN
  - Each representation gets uniquely trained
Encryption Scheme (OTBT)

- Trained NN to generate ‘representations’
- Mixing
- Communication Channel
- Trained HMM
- Fuzzy Matcher
- No use of stolen Information
- Raw template not stored
- No privacy threats

Acts like transformation and key
Continuously updated i.e. dynamic system
Self generating dynamic ‘keys’
The key stream generated using OTBT tested using the NIST Statistical Test Suite. (Only Reports Number of Streams Failed on average in 1000 independent tests with random keys)
Image Quality - Fuzzy c-means
**Compute Unified Device Architecture**: a new hardware and software architecture for issuing and managing computations on the GPU

- Available for GeForce 8800 Series, Quadro FX 5600/4600 and beyond
- No need to go through a graphics API
- Dedicated features for general computing
Research

- Non-linear Adaptive Signal Processing
- Bio-informatics
- Biometrics
- Biomedical Signal Analysis
- Scalable Video Coders
- Statistical Analysis of Signals
- Security Systems

WORK
Multi-Target Tracking (MTT)

◆ **Multiple Human targets to be tracked**
  - Computationally complex process

◆ **Applications are real-time**

3.1. Example...

![Multi-Target Tracking Diagram](image-url)
MATLAB GPU Computing

- MATLAB GPU Computing – AcceLereYes Jacket
- Casting input data to Jacket’s GPU data-structure
- To run native codes on the GPU
1.1. The Concept

- Human motion tracking an important application for security identification
- Marker-less motion capture approaches rely on gradient based methods
- Often suffer with ‘local minima’ problems
- Particle filters rely on approximation of probability distribution
- Large particles needed in state space
- For proper sampling of high-dimensional state space
1.2. The Drive

- Human motion, like human mind, non-linear and non-predictive
- Human motion processes difficult to model using standard distributions
- Human motion dynamics difficult to capture using linear motion models
- In this we attack
  - multi-target tracking problem
  - incorporation of complex target-to-target interactions
  - particle filter framework
2. Tracking Framework

- In the Bayesian tracking framework, desired to discover human target state $X$ (position, velocity etc.) from image observations in $Z$

- We use particle filter to approximate the posterior density by a set of $M$ weighted samples of this density

- Let the particle set $\{(s_{t-1}^1, w_{t-1}^1), ..., (s_{t-1}^M, w_{t-1}^M)\}$ represent the posterior density for time $t - 1$

- $\delta$ represents the Kroneker delta function, and the weights sum to infinity ($\sum_{i=1}^{M} w_{t-1}^i = 1$)

\[
p(X_{t-1} \mid Z_{1:t-1}) \approx \sum_{i=1}^{M} w_{t-1}^i \delta(X_{t-1} - s_{t-1}^i) \quad (1)
\]
We use CONDENSATION (Conditional Density Propagation) method for generating a particle set

\[ h(X_t) \approx \frac{1}{M} \sum_{i=1}^{M} w_i^i h(s_t^i) \quad (2) \]

To estimate the state \( X_t \) by the expected value of \( X_t \), we insert \( h(X_t) = X_t \) in equation (2)

For maximum a posteriori estimation the estimated target state is \( s_t^j \) where \( j = \arg \max_j \{w_t^1, ..., w_t^M\} \)
3. Multi-Target Tracking

- Human surveillance applications often involve multiple targets
- One possible solution for tracking $N$ subjects would be replicating single-subject technique $N$ times
- This straightforward method suffers with
  1. $N$-fold increase in the computational complexity
  2. No subject-to-subject interaction considerations
Multiple target tracking (MTT) challenging problem

1. complexity of the data
2. multiple targets
3. several measurements

Appearance and disappearance of targets from the field of view necessitate for a sound framework

We use multiple hypothesis tracking (MHT) technical framework

followed by the Markov chain Monte Carlo (MCMC) method
3.1. Example...

1. Subject Disappears
2. Subject Reappears
3.2. Multiple Hypothesis Tracking

\[ h(d_i) = \begin{cases} 
  j \neq 0 & \Rightarrow d_i \text{ is associated with track } j, \\
  0 & \Rightarrow d_i \text{ is generated from clutter}
\end{cases} \quad (3) \]

MHT computes the posterior distribution of the joint target state \( X_t \) by taking into consideration all possible hypothesis configurations and their associated probabilities:

\[ p(X_t \mid Z_{1:t}) = \sum_h p(h)p(X_t \mid h, Z_{1:t}) \quad (4) \]
3.3. Markov Chain Monte Carlo Method

A widely applied MCMC sampling algorithm in the biometrics image analysis named Metropolis-Hastings (MH) sampling is used in this particular study

\[
r' = \frac{\pi(s)q(s^i; s)}{\pi(s^i)q(s; s^i)} \tag{5}
\]

\[
\int h(X)\pi(X)dX \approx \frac{1}{N} \sum_{i=1}^{N} h(s^i) \tag{6}
\]

\[
p(X_t \mid Z_{1:t}) = c_t p(Z_t \mid X_t) \sum_{i=1}^{M} p(X_t \mid X_{t-1} = s^i_{t-1}) \tag{7}
\]
MH first randomly selects $k^{th}$ component of equation (7)

$$q(X_t^j) = p(X_t^j \mid \{X_t^l : l \neq j\}, X_{t-1} = s_{t-1}^k)$$  \hspace{1cm} (8)

Here, motion model $p(X_t \mid X_{t-1})$ factors as,

$$p(X_t \mid X_{t-1} = s_{t-1}^k) = p(X_t^j \mid \{X_t^l : l \neq j\}, X_{t-1} = s_{t-1}^k)$$

$$p(\{X_t^l : l \neq j\}, X_{t-1} = s_{t-1}^k) \text{ for an}(9)$$

$$r = \frac{p(Z_t \mid X_t^j = s^{\text{new}}, \{X_t^l : l \neq j\})}{p(Z_t \mid X_t^j = s, \{X_t^l : l \neq j\})} \hspace{1cm} (10)$$
Pseudo code for implementing MTT-MCMC is as follows:

- \( \{s_t^{j,i}\}_{j=1}^{N},i=1 \) = MTT-MCMC[\( \{s_t^{j,i}\}_{j=1}^{N},i=1 \)]
- Choose initial samples \( \{s_t^{j,0}\}_{j=1}^{N} \)
- for \( i = 1 : M \)
  - Choose an index \( k \in \{1, ..., M\} \) distributed uniformly
  - for \( j = 1 : N \)
    - Draw:
      \[ p(X_t | X_{t-1} = s_{t-1}^k) = p(X_t^j | \{X_t^l : l \neq j\}, X_{t-1} = s_{t-1}^k) p(\{X_t^l : l \neq j\}, X_{t-1} = s_{t-1}^k) \]
    - Compute:
      \[ r = \frac{p(Z_t | X_t^j = s, \{X_t^l : l \neq j\}, X_{t-1} = s_{t-1}^k) q(X_t^j = s)}{p(Z_t | X_t^j = s, \{X_t^l : l \neq j\}, X_{t-1} = s_{t-1}^k) q(X_t^j = s^{new})} \]
  - Choose: \( u \in [0, 1] \), uniformly distributed
  - if \( r \geq u \)
    - Assign: \( s_t^{j,i} \leftarrow s \)
  - else
    - Assign: \( s_t^{j,i} \leftarrow s_t^{j,i-1} \)
  - end
  - end
- end
4. Data Management

<table>
<thead>
<tr>
<th>Database</th>
<th>No of samples</th>
<th>Average Time</th>
<th>fps rate</th>
<th>Average targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>moviedb</td>
<td>50</td>
<td>9.8 seconds</td>
<td>30</td>
<td>18</td>
</tr>
<tr>
<td>labdb</td>
<td>50</td>
<td>12 seconds</td>
<td>25</td>
<td>3</td>
</tr>
</tbody>
</table>
5. Multiple Tracking

- We apply the designed algorithm for detecting multiple targets.
- We first detect the human subjects within each video frame.
- For this, we employ CONDENSATION method followed by gray-level thresholding.
- For each target \( j \), we build the motion model

\[
p(r_t, \tau_t \mid r_{t-1}, \tau_{t-1}) \propto \exp\left[\frac{(r_t - r_{t-1})^2}{-2\sigma_r}\right] \exp\left[\frac{(\tau_t - \tau_{t-1})^2}{-2\sigma_r}\right]
\]

(11)

- We call this model as MULTI-HUMAN-TRACK model.
5.1. CONDENSATION method
MTT-MCMC algorithm: probability model

\[
p(Z_t \mid \{(x_t^j, y_t^j)\}_{j=1}^M) \propto \prod_{x,y \in T_{in}} \exp\{-[I_t(x,y) - \mu_{in}]^2 / 2\sigma_{in}^2\} \\
\times \prod_{x,y \in T_{out}} \exp\{-[I_t(x,y) - \mu_{out}]^2 / 2\sigma_{out}^2\}
\]

Observation/measurement density:

\[
p(Z_t \mid \{(x_t^j, y_t^j)\}_{j=1}^M) \propto \prod_{x,y \in T_{in}} \frac{\exp\{-[I_t(x,y) - \mu_{in}]^2 / 2\sigma_{in}^2\}}{\exp\{-[I_t(x,y) - \mu_{out}]^2 / 2\sigma_{out}^2\}}
\]
Efficient Multi-target Human Motion Tracking

Pseudo code for implementing MULTI-HUMAN-TRACK is as follows:

- $\{(x^j_i, y^j_i), (x^j_i, y^j_i)^{M}_{i=1}\}^N_{j=1} = \text{MULTI-HUMAN-TRACK}([\{x^j_i, y^j_i\}^M_{i=1}]^N_{j=1})$
- Choose initial samples $\{x^{j,0}_t, y^{j,0}_t\}^N_{j=1}$
- for $i = 1 : M$
  - Choose an index $k \in \{\lfloor 0.2M \rfloor, \lceil 0.2M \rceil + 1, ..., M \}$ distributed uniformly
  - Convert Cartesian $(x^{j,k}_{t-1}, y^{j,k}_{t-1})$ to Polar $(r^{j,k}_{t-1}, \phi^{j,k}_{t-1})$
  - for $j = 1 : N$
    - Determine Gate $G_j$
    - Draw: $(r, \phi) \sim \{(r^l, \phi^l), p(r^l, \phi^l | r^{j,k}_{t-1}, \phi^{j,k}_{t-1})\} \in G_j$
    - Convert Polar to Cartesian
    - Compute Metropolis-Hastings ratio
    - Choose: $u \in [0,1]$, uniformly distributed
      - if $R \geq u$
        - Assign: $(x^{j,i}_{t}, y^{j,i}_{t}) \leftarrow (x, y)$
      - else
        - Assign: $(x^{j,i}_{t}, y^{j,i}_{t}) \leftarrow (x^{j,i-1}_{t}, y^{j,i-1}_{t})$
      - end
    - end
  - end
  - for $j = 1 : N$
    - Compute $(\overline{x}^j_t, \overline{y}^j_t) \leftarrow (x^{j,0}_{t} y^{j,0}_{t})^{M}_{i=0.2M}$
  - end
Tracked paths across the frames for all the identified subjects:
6. Results

- Human multiple tracking method applied to all the available sequences
- Tracking results compared with the manual annotations (ground truth)
- Average errors across all the targets calculated for every video against every video tag
- The average error obtained is 8.3%
- For challenging non-trivial conditions justifiable average error obtained
Plot showing the error information across 50 video samples of both the databases:
Research

- Non-linear Adaptive Signal Processing
- Bio-informatics
- Biometrics
- Security Systems
- Statistical Analysis of Signals
- Biomedical Signal Analysis
- Scalable Video Coders
Scalable Video Coders (SVCs)

- Produces bit-stream which can be decoded at different bit rates
  - Encoder design crucial

- Decoding on less powerful platform possible
Wavelet based SVC

- Spatio-temporal scalability
- Performs motion compensation
- Mesh based approach
- Computationally complex
- GPU-CUDA accelerates
- Effective implementation
Mesh Based MC
MATLAB GPU Computing – AcceLereYes Jacket
Casting input data to Jacket’s GPU data-structure
To run native codes on the GPU
## Results

<table>
<thead>
<tr>
<th>CPU Specifications</th>
<th>GPU Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP xw9400, AMD Dual core processor, 2.00 GHz, RAM 2 GB</td>
<td>nVIDIA GF 9800 GX2 + AMD Dual Core</td>
</tr>
<tr>
<td><strong>CPU Time (VD-10)</strong></td>
<td><strong>GPU Time (VD-10)</strong></td>
</tr>
<tr>
<td>270.74639 sec</td>
<td>20.86976 sec</td>
</tr>
</tbody>
</table>

- Acceleration Achieved ~ 13x

**Diagram:**
- CPU Processing
- GPU Processing

**Graph:**
- Time Units
- Video Depth
Research

Non-linear Adaptive Signal Processing

Bio-informatics

Biometrics

WORK

Security Systems

Statistical Analysis of Signals

Biomedical Signal Analysis

Scalable Video Coders
Research: Biometrics

Ideal and Non-Ideal
Iris Recognition

Biometric Security
Fingerprint Interoperability

Fingerprint Liveness Detection
Iris Template Aging and Update
Wavelet Based Scalable Video Coders
Research: Biometrics

- Ideal and Non-Ideal
- Iris Recognition

Biometrics

- Fingerprint Orientations
- Fingerprint Liveness Detection
- Iris Template Aging and Update
- Wavelet Based Scalable Video Coders
- Biometric Security
- Fingerprint Interoperability
Fingerprint Orientations
### Results

<table>
<thead>
<tr>
<th>CPU Specifications</th>
<th>GPU Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP xw9400, AMD Dual core processor, 2.00 GHz, RAM 2 GB</td>
<td>nVIDIA GF 9800 GX2 + AMD Dual Core</td>
</tr>
<tr>
<td>CPU Time (dpi-1000)</td>
<td>GPU Time (dpi-1000)</td>
</tr>
<tr>
<td>35.44659 sec</td>
<td>5.06976 sec</td>
</tr>
</tbody>
</table>

**Acceleration Achieved ~ 7x**

![Diagram showing CPU and GPU processing times](image)
## Results

<table>
<thead>
<tr>
<th>CPU Specifications</th>
<th>GPU Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP xw9400, AMD Dual core processor, 2.00 GHz, RAM 2 GB</td>
<td>nVIDIA GF 9800 GX2 + AMD Dual Core</td>
</tr>
<tr>
<td><strong>CPU Time (dpi-1000)</strong></td>
<td><strong>GPU Time (dpi-1000)</strong></td>
</tr>
<tr>
<td>72.54639 sec</td>
<td>9.26976 sec</td>
</tr>
</tbody>
</table>

**Acceleration Achieved ~ 8x**

Facial Images dpi resolution

Time Units

CPU Processing

GPU Processing
## Results - CBIR

<table>
<thead>
<tr>
<th>CPU Specifications</th>
<th>GPU Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP xw9400, AMD Dual core processor, 2.00 GHz, RAM 2 GB</td>
<td>nVIDIA GF 9800 GX2 + AMD Dual Core</td>
</tr>
<tr>
<td>CPU Time (dpi-1000)</td>
<td>GPU Time (dpi-1000)</td>
</tr>
<tr>
<td>50.54639 sec</td>
<td>5.36976 sec</td>
</tr>
</tbody>
</table>

Acceleration Achieved ~ 10x

**Diagram:**
- **CPU Processing**
- **GPU Processing**
- **Search Depth**
- **Time Units**
Research

- Non-linear Adaptive Signal Processing
- Bio-informatics
- Biometrics
- Security Systems
- Statistical Analysis of Signals
- Biomedical Signal Analysis
- Scalable Video Coders
Fuzzy c-means scatter plot
## Results

<table>
<thead>
<tr>
<th>CPU Specifications</th>
<th>GPU Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP xw9400, AMD Dual core processor, 2.00 GHz, RAM 2 GB</td>
<td>nVIDIA GF 9800 GX2 + AMD Dual Core</td>
</tr>
<tr>
<td><strong>CPU Time (Epochs-200)</strong></td>
<td><strong>GPU Time (Epochs-200)</strong></td>
</tr>
<tr>
<td>37.07 sec</td>
<td>4.87 sec</td>
</tr>
</tbody>
</table>

Acceleration Achieved ~ 8x
P1: Liveness Embedded Multimodal Biometric System

Fusion:
1. Image Level
2. Feature Level
3. Score Level

Biometric System → Security, Forensics

LIVENESS
P2: Adaptive DPOAE design

Dual Pulse Oto-Acoustic Emissions
P3: Homography based 3D voxel mapping for Rehabilitation Engineering
P4: Evolutionary Framework for objective diagnosis of cancer

### Images Produced in Pharmaceutical R&D

<table>
<thead>
<tr>
<th>Target Identification</th>
<th>Lead Identification</th>
<th>Lead Optimization</th>
<th>Pre-clinical</th>
<th>Clinical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional Genomics</td>
<td>High Throughput Screening</td>
<td>Structure / Activity Analysis</td>
<td>ADME and <em>In Vitro</em> Toxicity</td>
<td>Animal Testing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Human Testing</td>
</tr>
<tr>
<td>Proteomics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Problem: No effective way to mine image data**
P4: Evolutionary Framework for objective diagnosis of Histo-paths
References (selected):

- Stephanie Schuckers, Sunil Kumar, Aditya Abhyankar, “Encryption of Biometric Templates using Biometrics as the Key”, CTIE Fall meeting, Morgantown, November 2005.

- Aditya Abhyankar and Stephanie Schuckers, “B-orthogonal Wavelet Network (BWN) systems for efficient Iris Encoding and Recognition”.
- Aditya Abhyankar, Amit Vijayn, Sunil Kumar, Sun Rose, Stephanie Schuckers, “Encryption of Biometric Templates using Biometric Information as the Key”.
- Aditya Abhyankar, Nilesh Kulkarni, Sunil Kumar, Stephanie Schuckers, “Fingerprint Image Quality and Prediction of Matching Performance”.
- Aditya Abhyankar, Stephanie Schuckers, “Fingerprint recognition Anti-spoofing issues”, Biometric Consortium Conference (BCC), (Crystal City, VA), September 2003.
- Aditya Abhyankar, Stephanie Schuckers, “Iris Recognition using bi-orthogonal wavelets”, Biometric Consortium Conference (BCC), (Crystal City, VA), September 2004.
- Aditya Abhyankar, Stephanie Schuckers, “Feature Level Fusion of Level-3 Fingerprint Fingerprint Recognition Pattern Characteristics for Robust Fingerprint Recognition”.
- Aditya Abhyankar, Stephanie Schuckers, “Combining Local Ridge Frequencies with Multiresolution Statistical Level-1 and 2 Features for Vitality Detection in Fingerprints”.
- Aditya Abhyankar and Stephanie Schuckers, “Active Shape Analysis for Effective Non-ideal Iris Recognition Issues”.

Questions??
Thank you!

UNIVERSITY OF PUNE
Ganeshkhind, Pune-411007.

Dr. Aditya Abhyankar
Dean, Faculty of Technology
Professor, Dept. of Technology

Office: Department of Technology, University of Pune,
Pune-411007
Phone: +91-20-25601270
Mobile: +91-9860119930
E-mail: aditya.abhyankar@unipune.ac.in
aditya1210@gmail.com