Different Target Prediction Algorithms for Automotive, HRI and VR Digital Twin

Pradipta Biswas, ***PhD (Cantab)***

Associate Professor
Indian Institute of Science

https://cambum.net/, *pradipta@iisc.ac.in*

Short Bio

Associate Professor at Dept of Design and Manufacturing and Associate Faculty at Centre for Cyber Physical System, Indian Institute of Science

Vice Chairman ITU SG9, Co-Chairman of IRG AVA, Vice-Chairman at FG Smart TV at International Telecommunication Union (ITU)

Convenor of BIS (ISO) Panel on Metaverse

Member, UKRI International Development Peer Review College

PhD in Computer Science from Trinity College, University of Cambridge, UK

Senior Research Associate at Cambridge University Engineering Department and Governing Body Fellow at Wolfson College

Worked with Jaguar Land Rover, Technicolor, BAE Systems from 2010 – 16

I3D Lab (2016)

Pradipta Biswas


BusinessLine

IISc develops eye gaze-controlled robotic arm for those suffering from speech and motor impairment

Pradipta Biswas talks about how his research on digital accessibility led to work on India's space programme.

₹ 1L ≈ £1K
₹ 1Cr ≈ £100K

Major Funding Agencies

- Designing VR model of Gaganayan Crew Cabin funded through **ISRO Human Space Flight Centre** (₹ 38L)
- Metaverse related projects with **Siemens**, Germany (₹ 77L) and **Collins Aerospace, USA** (₹ 71L)
- Designing Advanced HMI for next generation fighter aircraft platforms funded through Aeronautical Development Agency (**ADA**) & **ARDB, DRDO** (₹ 2 Cr)
- Developing VR model of Office Spaces funded by **British Telecom**, UK (₹90L)
- Patented AR-based Interactive Head Up Display System with **Forvia (Faurecia)**, France (₹60L)
- Developed VR based Flight Simulator as part of IISc-**Hindustan Aeronautical Ltd** (HAL) Skill Development Centre (₹ 42L)
- **Facebook** Responsible Innovation in AR/VR Award ($75K)
- **TCS Innovation Lab** Grant and PhD Fellowship (₹1 Cr) for Multiagent Robotics systems
- **Microsoft** AI 4 Accessibility Grant Award ($15K)
- Department of Science and Technology & Dept. of Bio-Technology, Govt of India (₹75L)

Graduated PhD Students

1. Dr Gowdham Prabhakar (2016-20) has joined as **Assistant Professor at IIT Kanpur** after postdoc at **UCL, UK**
2. Dr Jeevithashree DV (2018-21) is working as a PostDoc at **Purdue University, USA**
3. Dr Somnath Arjun (2016-21) has joined as Senior Engineer at **Siemens**
4. Dr Vinay Krishna Sharma (2018-23) has joined as Senior Engineer at **Siemens**
5. Dr LRD Murthy (2017-23) has joined as TechLead at **Mercedes-Benz**
6. Dr Archana Hebbar (2018-23) is Senior Principal Scientist at **CSIR-NAL**
7. Dr Abhishek Mukhopadhyay (2018-23) is Manager – Product Design at **Ashok Leyland**

HoloLens Based Assembly Demonstration

Subin Raj, Bikram Karmakar, Gyanig Kumar, Harshitha, Abhishek Mukhopadhyay, Yash Sahoo, Amrit Chatterjee, Pradipta Biswas
Indian Institute of Science, Bangalore, India | Collins Aerospace Systems

Gaze Controlled Mixed Reality Robotics Applications for Children with SSMI

Dr Pradipta Biswas, PhD (Cantab)
Associate Professor
Indian Institute of Science
https://cambum.net/

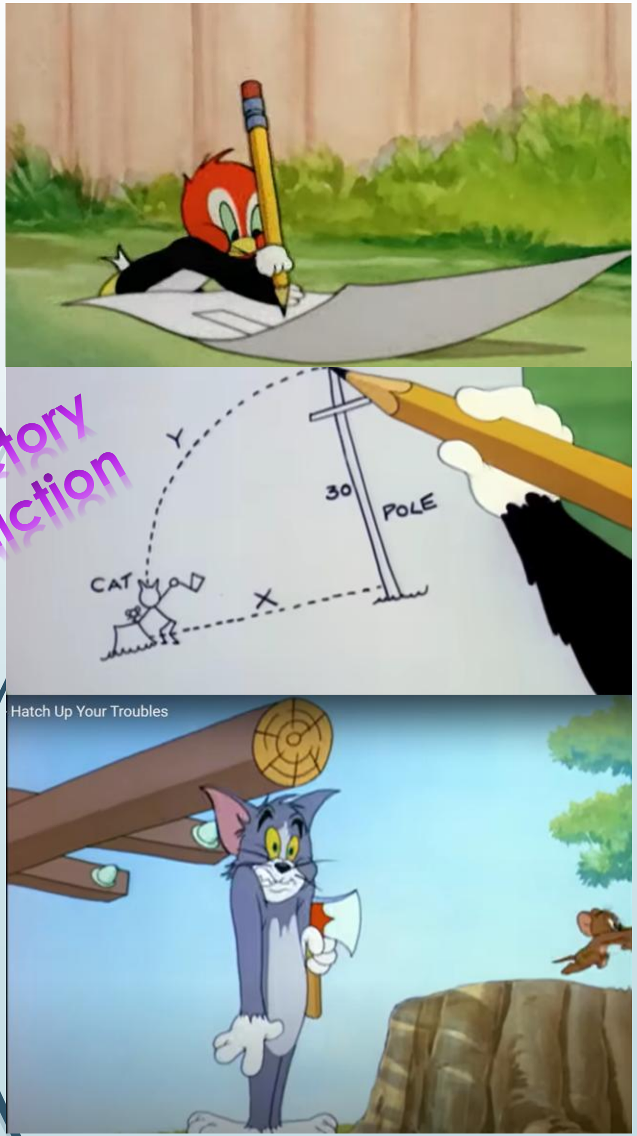
Robert Bosch Centre for
Cyber Physical Systems

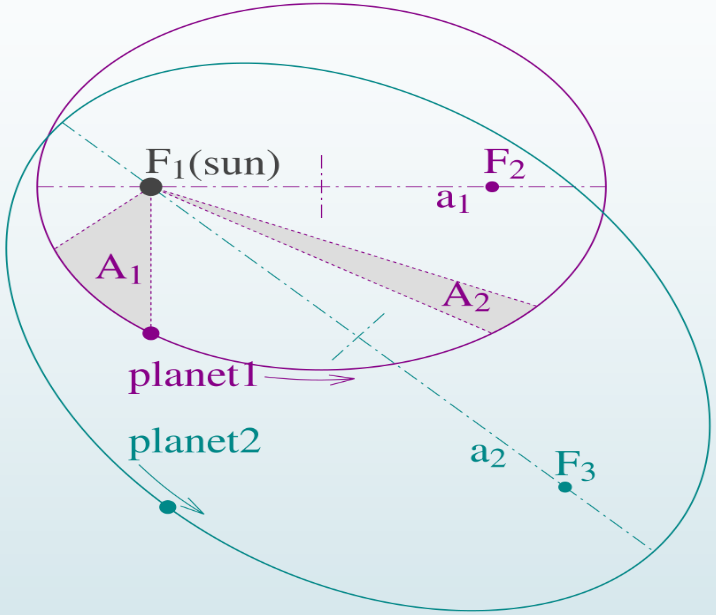
DEPARTMENT OF
DESIGN AND MANUFACTURING
INDIAN INSTITUTE OF SCIENCE

Contents

- Rapid Aiming Movements
- Artificial Neural Network based Approach
  - Automotive and Aviation Use Cases
- Imitation Learning
  - Inverse Reinforcement Learning
  - Behaviour Cloning
  - XR Interaction and Human Robot Handover
- Interactive Digital Twin
  - Objective
  - Walking Trajectory Prediction
- Concluding Remarks

Why Target Prediction



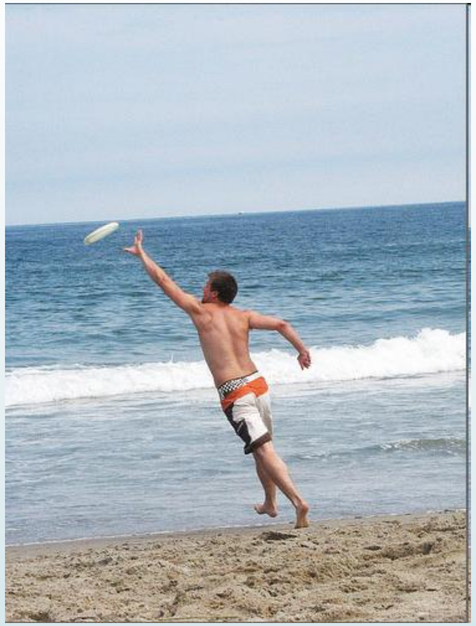


Rapid Aiming Movements

Quick

Accurate

Preprogrammed



Analysis of Rapid Aiming Movements

Woodworth's study

Eye open – visual feedback

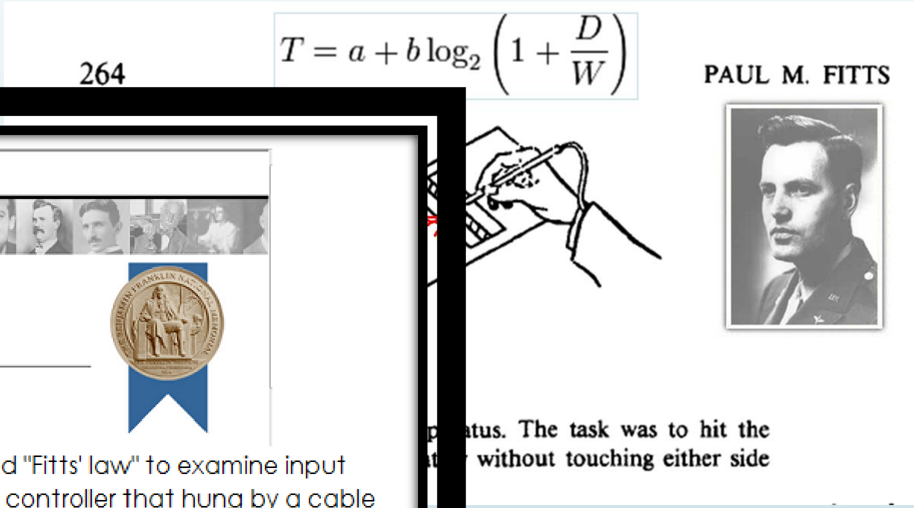
Error ∞ Speed

Eye close – no v

Accuracy d

- Phases of W
  - Initial in
  - Curre

Fitts' Law



THE FRANKLIN INSTITUTE **Awards**

MAIN SITES

The Franklin Institute

Resources for Science Learning

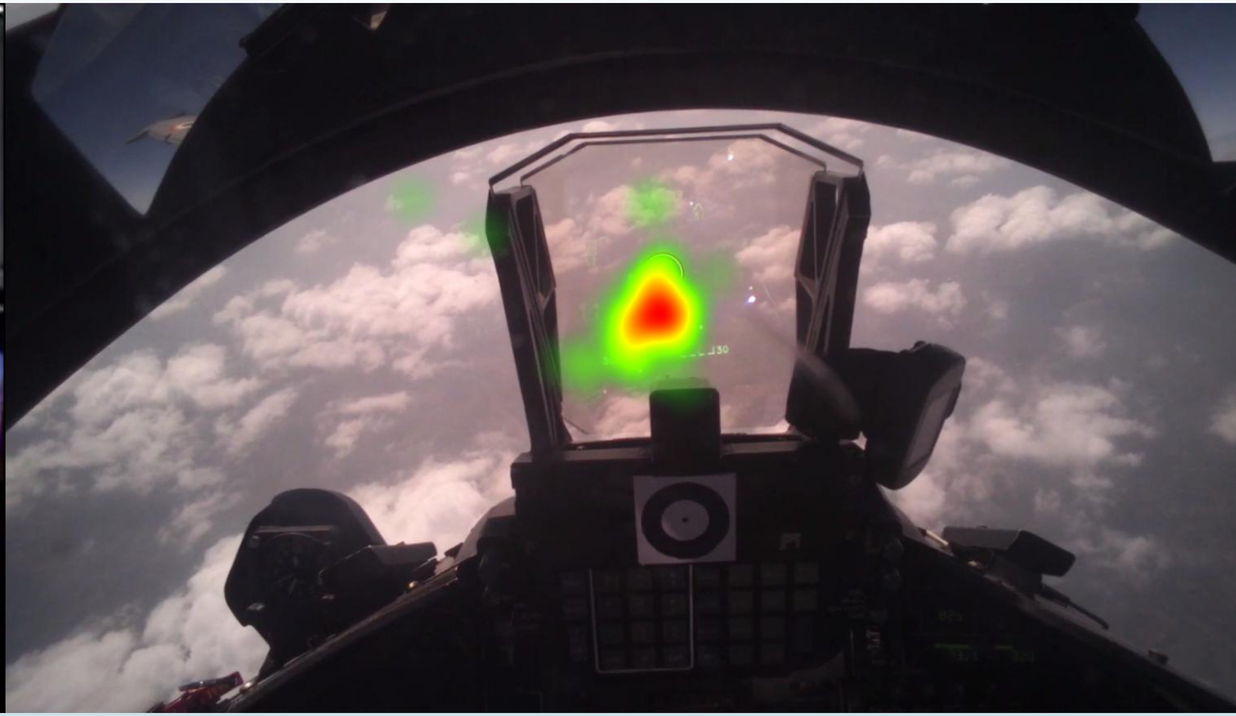
Franklin Laureate Database

Laureate Details

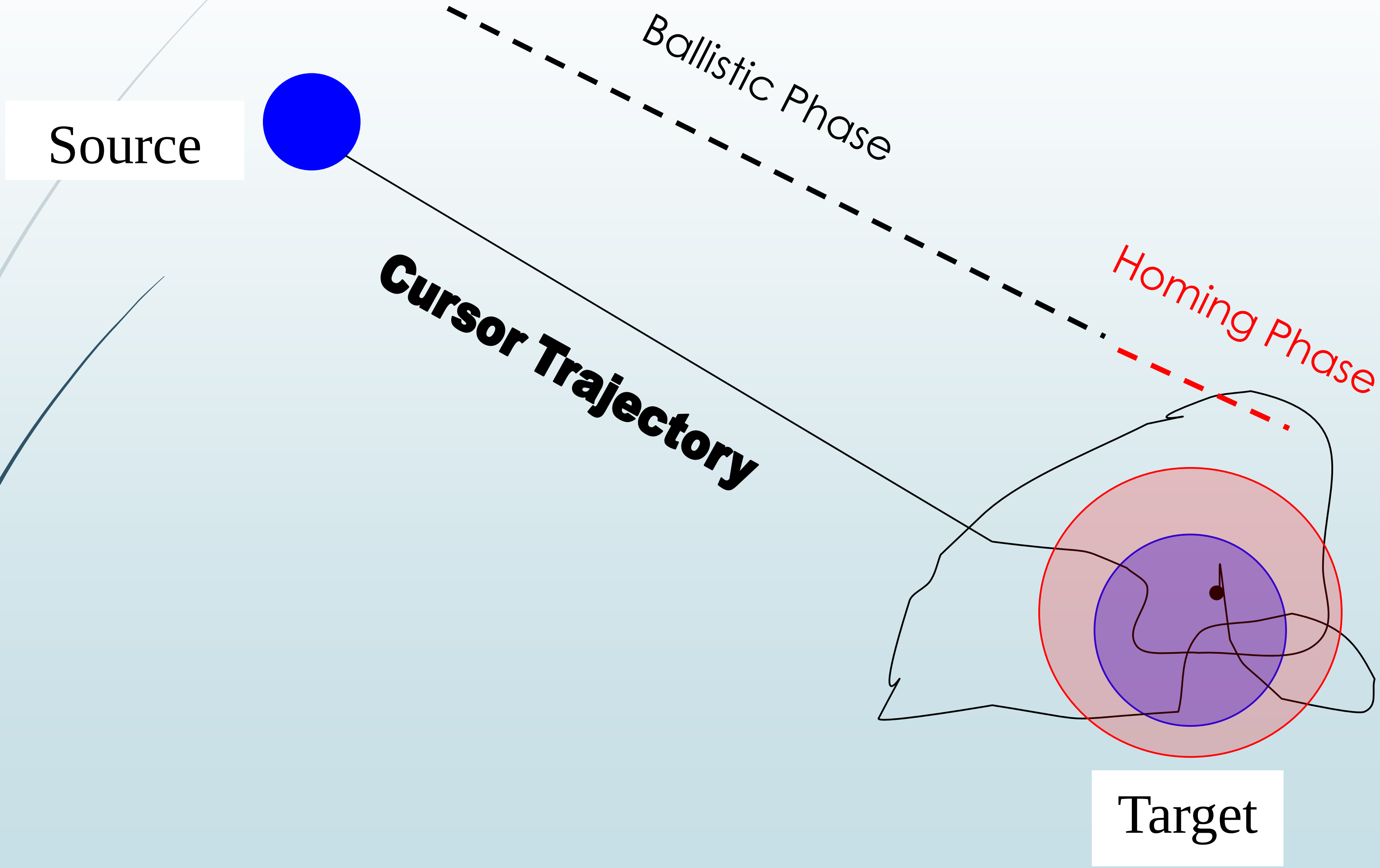
Awards Home » Franklin Laureate Search Form » ...Laureate Search Results » Laureate Details

Card's first home run came when he used a then-little-known law called "Fitts' law" to examine input devices such as joy sticks, a head-motion detector, and a newfangled controller that hung by a cable from the computer called a "mouse." Fitts' law analyzes how easily a human can hit any given target - in this case, moving a cursor to a specific point on a screen - and Card measured the mouse to be almost as simple as if one could move the cursor around on the screen with one's hand. After Card's work, Xerox began manufacturing mice, Apple soon followed, and now they're practically de rigueur on all computers.

Prospective Applications

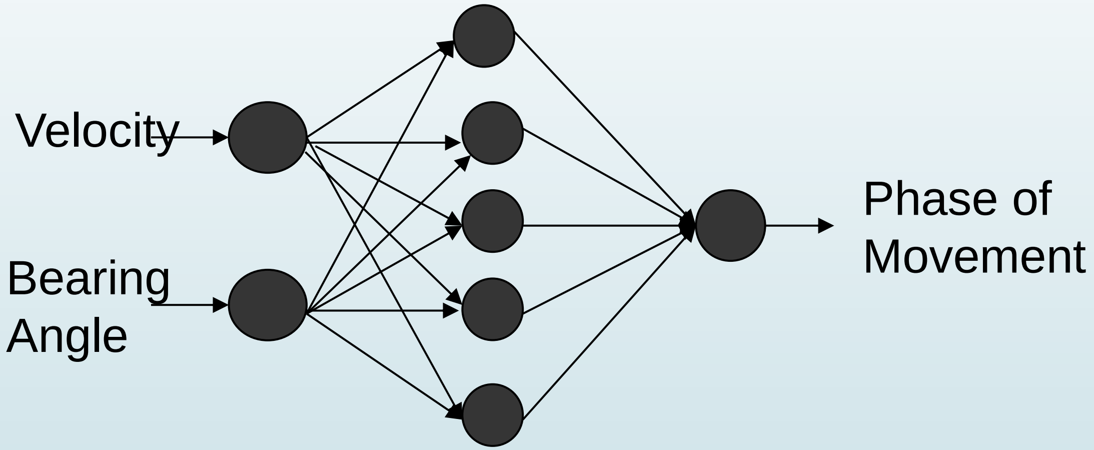


Analysing Trajectory

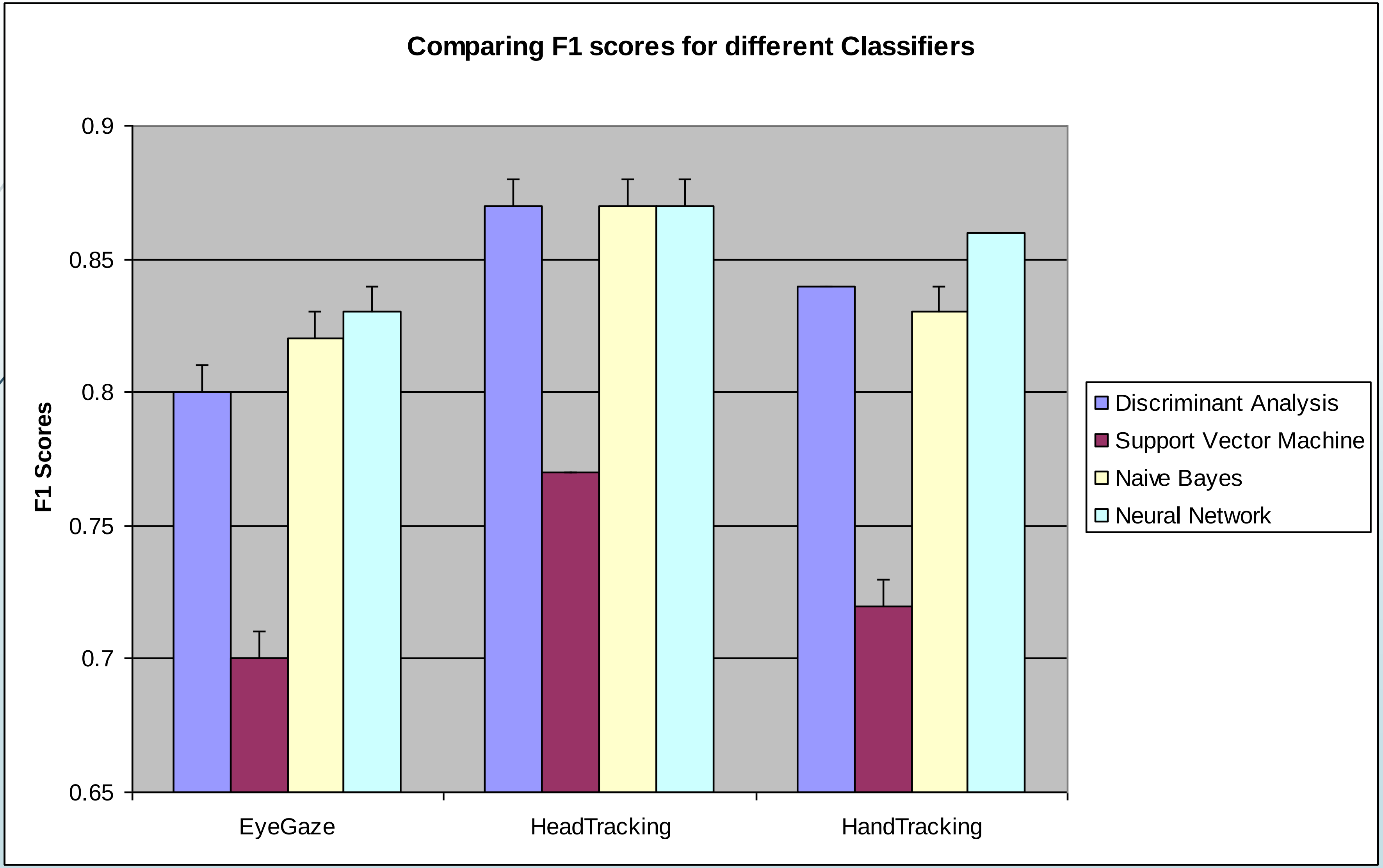


Algorithm – Neural Network

- For every change in position of pointer in screen
  - Calculate angle of movement
  - Calculate velocity of movement
  - Calculate acceleration of movement
- Run Neural Network with Angle, Velocity and Acceleration
- Check output
- If output predicts homing phase
  - Find direction of movement
  - Find nearest target from current location towards direction of movement

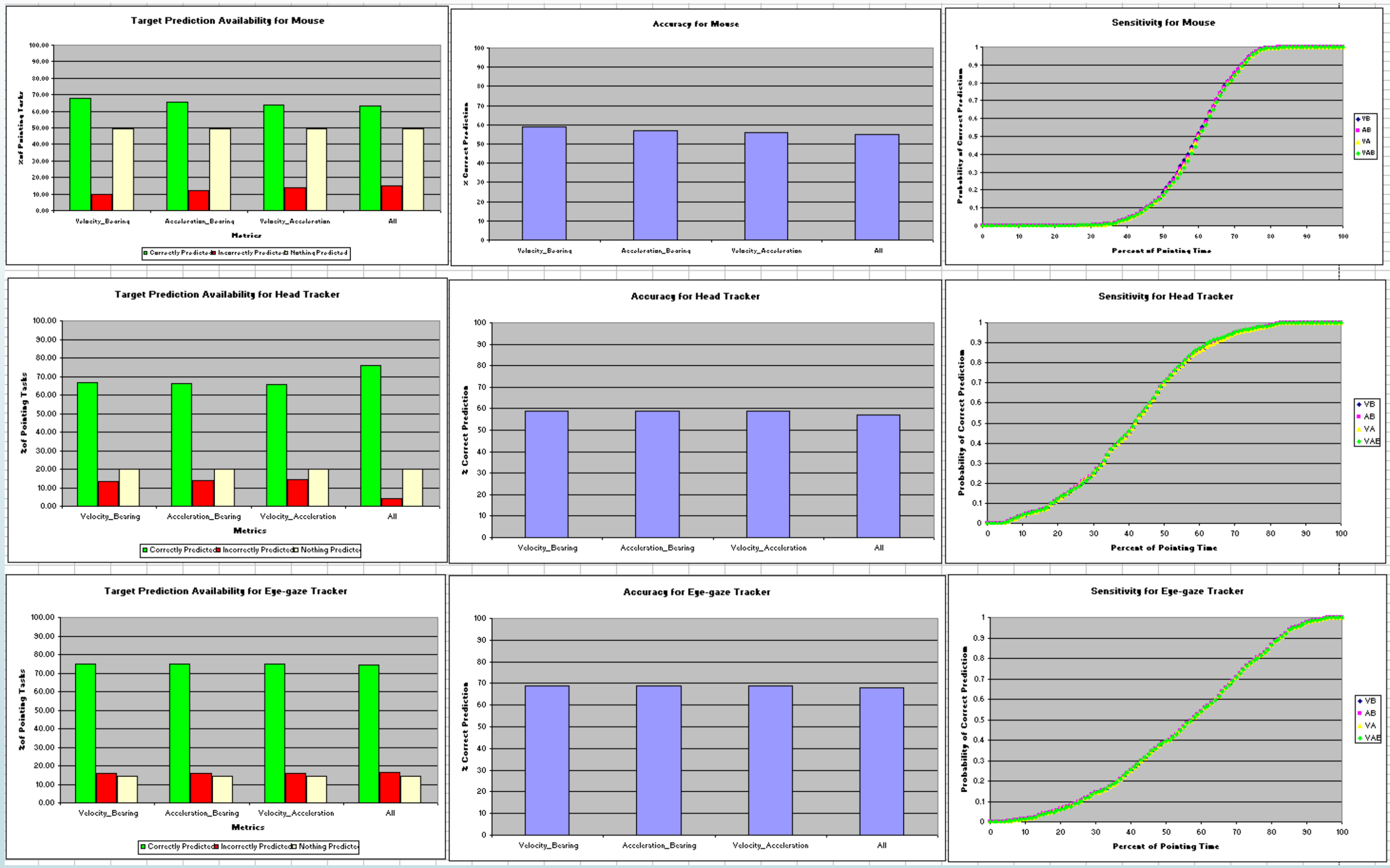


Classifier Result



Evaluation Criteria

- **Availability**: In how many pointing tasks the algorithm makes a successful prediction.
- **Accuracy**: Percentage of correct prediction among all predictions
- **Sensitivity**: How quickly an algorithm can detect intended target



Proposed Applications

- G. Prabhakar, A. Ramakrishnan, L. R. D Murthy, V. K. Sharma, M. Madan, S. Deshmukh and P. Biswas, Interactive Gaze & Finger controlled HUD for Cars, Journal on Multimodal User Interfaces, Springer, 2019
- P. Biswas, S. Deshmukh, G. Prabhakar, M. Modiksha, V. K. Sharma and A. Ramakrishnan,A System for Man- Machine Interaction in Vehicles , Indian Patent Application No.: 201941009219, PCT International Application No. PCT/IB2020/050253
- P. Biswas and P. Langdon, Mutimodal Target Prediction Model, ACM CHI 2014 Extended Abstract
- P. Biswas, and P. Langdon, Multimodal Intelligent Eye-Gaze Tracking System, International Journal of Human-Computer Interaction 31(4), Taylor & Francis, Print ISSN: 1044-7318

Learning from Demonstrations (LfD)

Why challenging?

- Variability in user demonstrations
- Real-time implementation
- Needs to be Quick and accurate
- Difficult to optimize these variations and predict future trajectory

Learning from demonstrations (LfD) → Imitation learning (IL)

- Behavior cloning (BC) – learning directly from actions, quick but less accurate
- Inverse reinforcement learning (IRL) – learning a reward function from actions, time consuming but accurate

Inverse Reinforcement Learning (IRL)

IRL → learn r under which expert demonstrations are optimal

s → state

$\mathcal{D} = \{\tau_i\}, i = 1,..,M$ be the expert dataset

f → features, captures human preferences during the task.

ω → feature weights

r → reward

The likelihood of expert demonstrations:

$$P(\mathcal{D}|\omega) = \prod_{i=1}^{M} \frac{e^{r(\tau_i,\omega)}}{\sum_{\tau_i \epsilon \mathcal{D}} e^{r(\tau_i,\omega)}} = \prod_{i=1}^{M} \frac{1}{Z} e^{r(\tau_i,\omega)}$$

Objective:

$$\omega^* = \underset{\omega}{\text{argmax}} \frac{1}{M} logP(\mathcal{D}|\omega) = \underset{\omega}{\text{argmax}} \frac{1}{M} \sum_{i=1}^{M} \{logP(\tau_i|\omega)\}$$

Features: distance, velocity, acceleration and jerk

Algorithm 1 Target Prediction

Input: Partial hand trajectory $\psi = \{s_1, ..., s_m\}$, Goal $G \epsilon \mathbb{G}$, MDP = $\{S, A, T, \gamma, r_G\}$, $r_G \epsilon R$.
Output: Predicted goal G_{pred}

1: **for** G in $\mathbb{G}$ **do**
2: $\quad r_G \leftarrow G$
3: $\quad$ Initialize $V_G = 0$
4: $\quad$ Update $V_G \leftarrow (r_G, S, A, T, \gamma)$,
5: $\quad p(\psi|G) = exp\left[\{\sum_{i=2}^{m} r_G(s_i)\} + V_G(s_m) - V_G(s_1)\right]$
6: $\quad p(G|\psi) = \frac{p(\psi|G)p(G)}{\sum_{G \epsilon \mathbb{G}} p(\psi|G)p(G)}$
7: **end for**
8: $G_{pred} \leftarrow \ max\ p(G|\psi)$

Future hand trajectory is obtained using $\pi_{G_{pred}}$

For multimodal prediction, prior $p(G)$ is obtained from the gaze data

IRL Algorithms

Sampling-based Maximum Entropy IRL (SMEIRL)

Z is approximated by summation over all sample trajectories $\psi_m, \mathrm{m} = 1, \dots, \mathrm{K}$

$$Z \approx \sum_{m=1}^{K} e^{R(\psi_m, \omega)}$$

Objective:

$$L(\omega) = \frac{1}{M} \sum_{i=1}^{M} \left\{ R(\tau_i, \omega) - log \sum_{m=1}^{K} e^{R(\psi_m^i, \omega)} \right\}$$

$$\nabla_\omega L = \frac{1}{M} \sum_{i=1}^{M} \{ f(\tau_i) - \tilde{f}(\psi_m^i, \omega) \}$$

Maximum Entropy Deep IRL (MEDIRL)

Reward r is estimated by a neural network: $r_\omega(s) = g(f(s), \omega)$

ω^* is obtained by backpropagating the gradient: $\frac{\partial L}{\partial \omega} = (\mu_D - \mathrm{E}[\mu]) \frac{\partial g(f, \omega)}{\partial \omega}$

μ_D → State visitation frequency (SVF) obtained from expert dataset

$\mathrm{E}[\mu]$ → Expected SVF from learned reward at each iteration

Approximate value iteration and policy propagation algorithm were used to estimate SVF

Algorithm 2 Target Prediction using MEDIRL

1: **function** MEDIRL
2: **Input:** Expert dataset $\mathfrak{D}$, partial trajectory φ, MDP, features(), MEDIRL(), targetPred(), valueIteration(), goal G, prediction horizon Δl ▷ Input parameters
3: f ← features($\mathfrak{D}$) ▷ Feature estimation
4: r_ω ← MEDIRL($\mathfrak{D}, MDP, G$) ▷ Reward learning
5: G_{MEDIRL} ← targetPred(r_ω, φ) ▷ Intended target prediction
6: $\varphi(l + \Delta l)$ ← valueIteration($G_{pred,MEDIRL}$) ▷ Future trajectory prediction
7: **return** $G_{MEDIRL}, \varphi(l + \Delta l)$ ▷ Return the result
8: **end function**

M Mitra, AA Patil, GVS Mothish, G Kumar, A Mukhopadhyay, Murthy LRD, PP Chakraborty, P Biswas, Multimodal Target Prediction for Rapid Human-Robot Interaction, ACM International Conference on Intelligent User Interfaces (IUI 24)

Target Prediction During Rapid-Aiming Movement

Target prediction during a pointing task in VR and MR

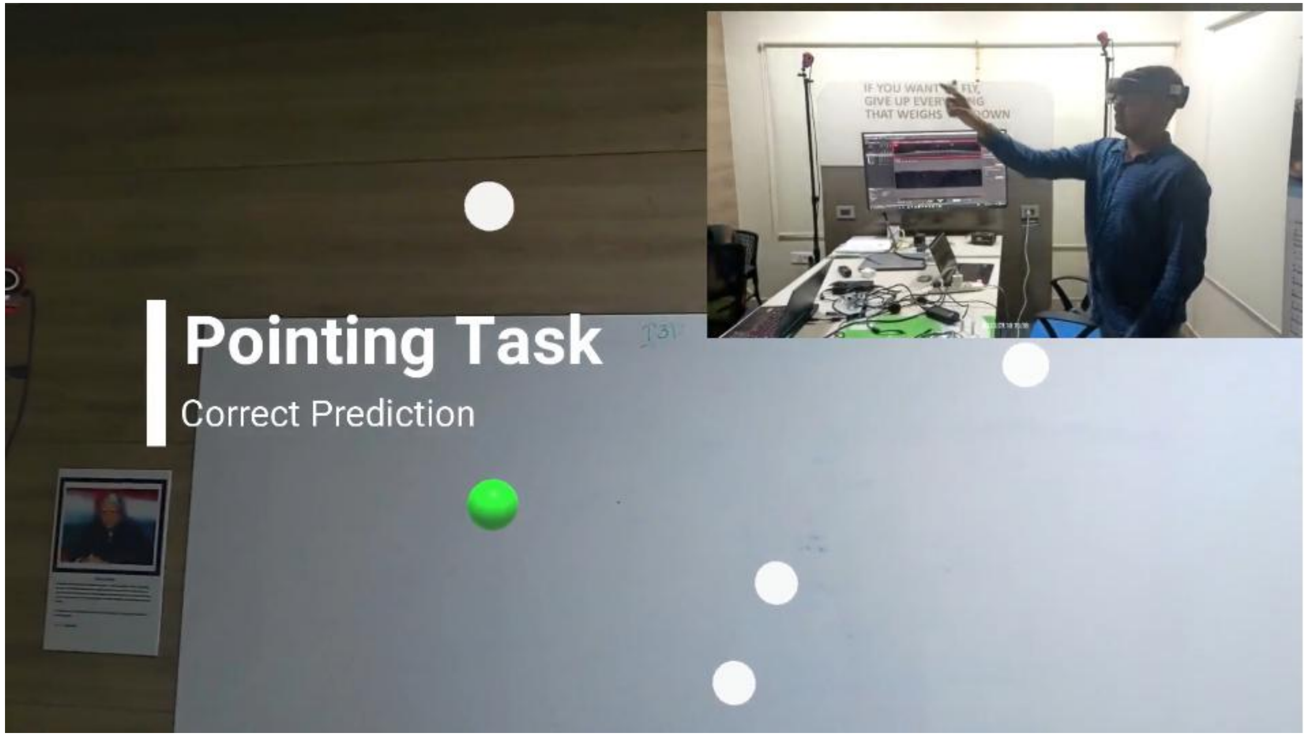
•Hand movement prediction using SMEIRL

Enhanced Human-Robot Collaboration with Intent Prediction using Deep Inverse Reinforcement Learning

•Hand movement prediction using MEDIRL

Multimodal Target Prediction for Rapid Human-Robot Interaction

•Hand movement prediction using MEDIRL + eye gaze



Behavior Cloning: Feature-based Bayesian Interaction Primitives (FBIP)

IRL accurate but slow

Hand motion representation with time-dependent basis functions and weights:

$$f(t) = \sum_{i=1}^{d} exp\left(-\frac{(f(t)-\mu_i)^2}{2\sigma_i^2}\right)\omega_i + \epsilon(t)$$

Basis matrix:

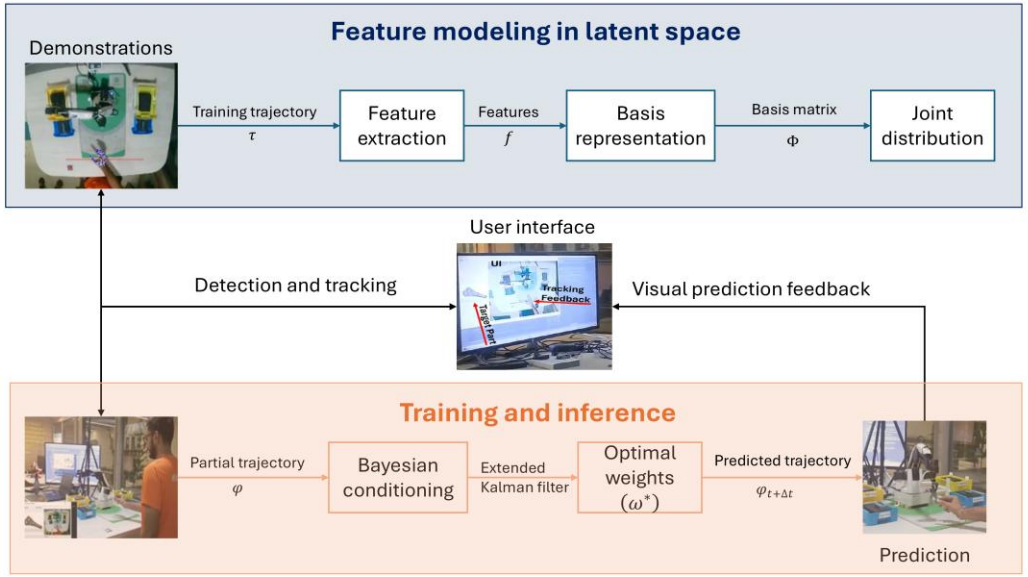
$$\Phi = \begin{bmatrix} \phi_1(\mu_1) & . & . & \phi_1(\mu_{i=d}) \\ . & & & . \\ . & & & . \\ \phi_{i=d}(\mu_1) & . & . & \phi_{i=d}(\mu_{i=d}) \end{bmatrix}$$

Objective:

$$\omega^* = \arg\max_{\omega} log(p(\varphi|\omega))$$

ω^* was obtained using EKF with state vector as the weights of the basis function. Future hand trajectory:

$$\varphi(t+\Delta t) = \Phi(t+\Delta t)^T \omega^*(t)$$



Algorithm 1 Target Prediction using FBIP

---

function FBIP
    Input: Expert dataset $\mathfrak{D}$, partial trajectory φ, features(), gaussianRBF(), EKF(), predGoal(), prediction horizon Δt ▷ Input parameters
    f ← features($\mathfrak{D}$) ▷ Feature estimation
    Φ ← gaussianRBF(f) ▷ Basis matrix formulation
    ω^* ← EKF(Φ, φ) ▷ Optimal state vector
    $\varphi(t+\Delta t) \leftarrow \Phi(t+\Delta t)^T\omega^*(t)$ ▷ Future trajectory prediction
    G_{FBIP} ← predGoal($\varphi(t+\Delta t)$) ▷ Intended target prediction
    return $G_{FBIP}, \varphi(t+\Delta t)$ ▷ Return the result
end function

Fast and Accurate: Rapid Human-Robot Handover for Collaborative Assembly and Disassembly

M Mitra, G Kumar, PP Chakraborty, P Biswas, Enhanced Human-Robot Collaboration with Intent Prediction using Deep Inverse Reinforcement Learning, IEEE International Conference on Robotics and Automation (ICRA 24)

BT Group

BTIRC
Indian Institute of Science

Interactive Digital Twins

THE NATIONAL AI AWARDS
FINALIST
The AI Innovation Award
2024

Team

- Prof Pradipta Biswas, IISc PI (+ research students)
- Dr Anasol Pena-Rios, BT Research & Networks Strategy
- Shakti Srivastava, BTIRC Technical Delivery

Interactive Digital Twins PoC

BTIRC
Indian Institute of Science

Office-based Workplaces

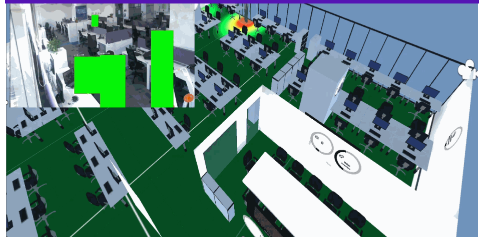
Solutions to Key Challenges

Know how much energy is used and what is used for operations

Energy Consumption Estimation

Know how much space is used to make decisions

Real-time Activity-based Analysis

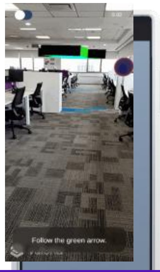


Maximise capacity considering foot traffic

Occupancy Simulation

Plan an effective workplace strategy, identifying workflows and considering employees' health, safety and welfare

Special Workflows (e.g., Social Distancing Analysis)



AR Guided Navigation

Our Scalable Solution

Privacy-by-design approach
- Anonymised real-time person and posture detection

Energy consumption estimation using AI
- Remote asset monitoring
- Real time alerts as per set up thresholds

AI-based real-time occupancy insights
- Insights into usage and movement in different size facilities

Low-cost devices
- Cameras & sensors

Workflow tracking
- Real-time activity mapping between physical and virtual spaces
- Occupancy monitoring
- Special workflows (auditing usage of personal protective equipment (PPE) in specific areas)

Simulation capabilities
- Optimisation and planning using a 3D virtual reality replica of the physical space
- Standard and special scenarios (e.g., social distancing and reduced office capacity)

Benefits

Optimised space management and asset utilisation

Provide detailed insights collecting and visualising (near) real-time data

Cost reductions and rebalancing

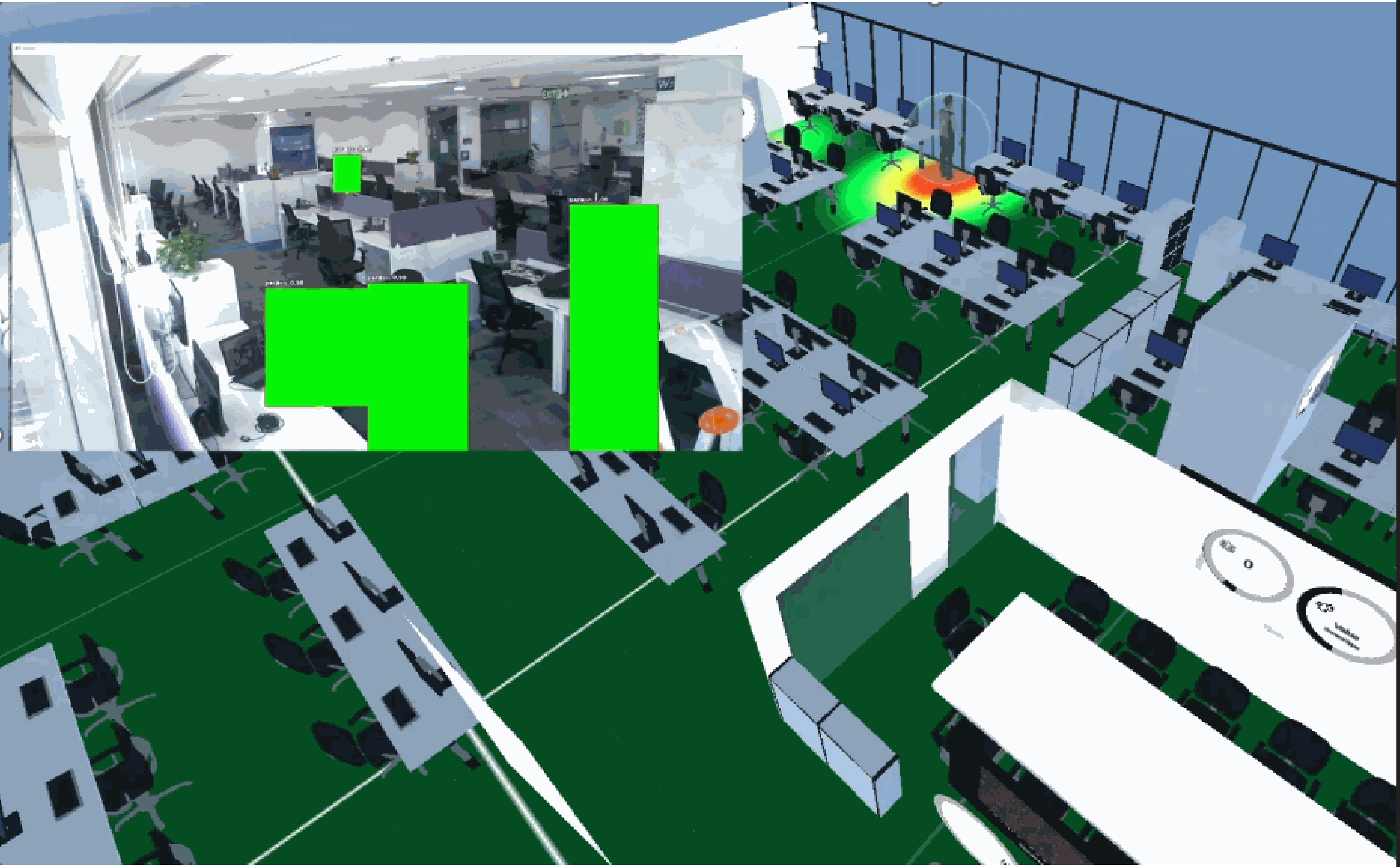
Increase operational efficiency

Stakeholder

BT India Workplace Team

- Biju Chanath, International Workplace, Property and Facilities

BT Deployment



Made with unity


eObject Component Tools ViitorCloud Window Help
Pivot
Global
x1080)
Scale
0.84x
sitting
standing
09326171875,1074.1474609375,1,p1,554.9268951416016,396.0365295410156,0

Particle filter overview

Predict the trajectory for future points

- Motion model (particles, velocity and std)

Update weights of particles

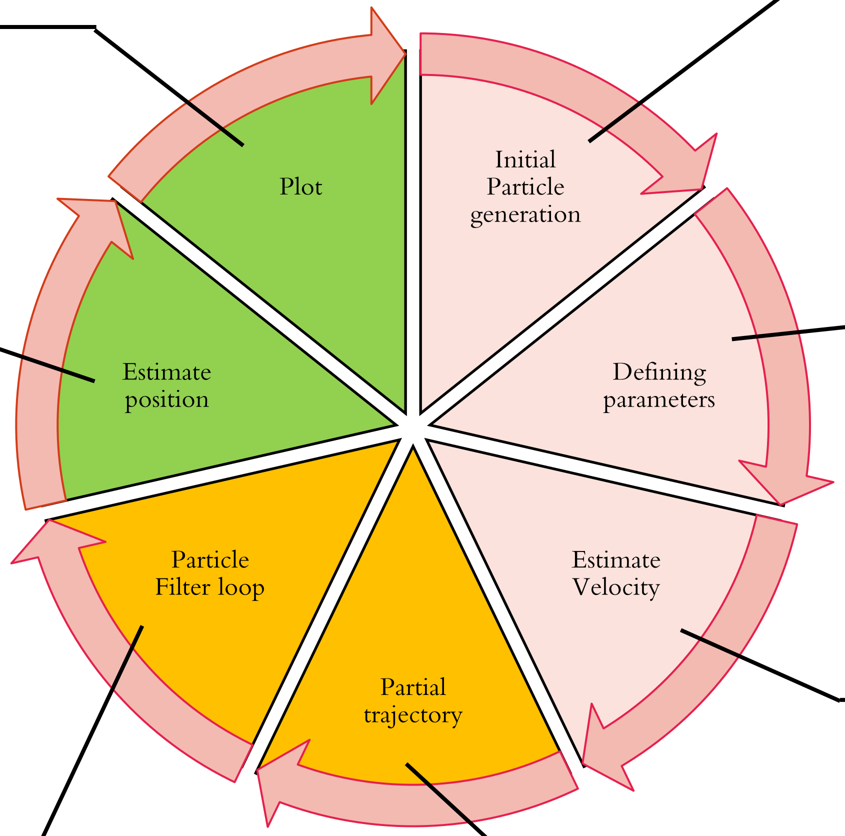
Resample particles based on their weights

Estimate the position based on particle weights

Plot actual, partial, predicted trajectories

Initializing the particles based on the trajectory

Estimating the position based on weighted average of the particles



Defining the parameters

- Number of particles: 3000
- Destination positions
- Noise and std : 0.1, 1
- % of partial trajectory

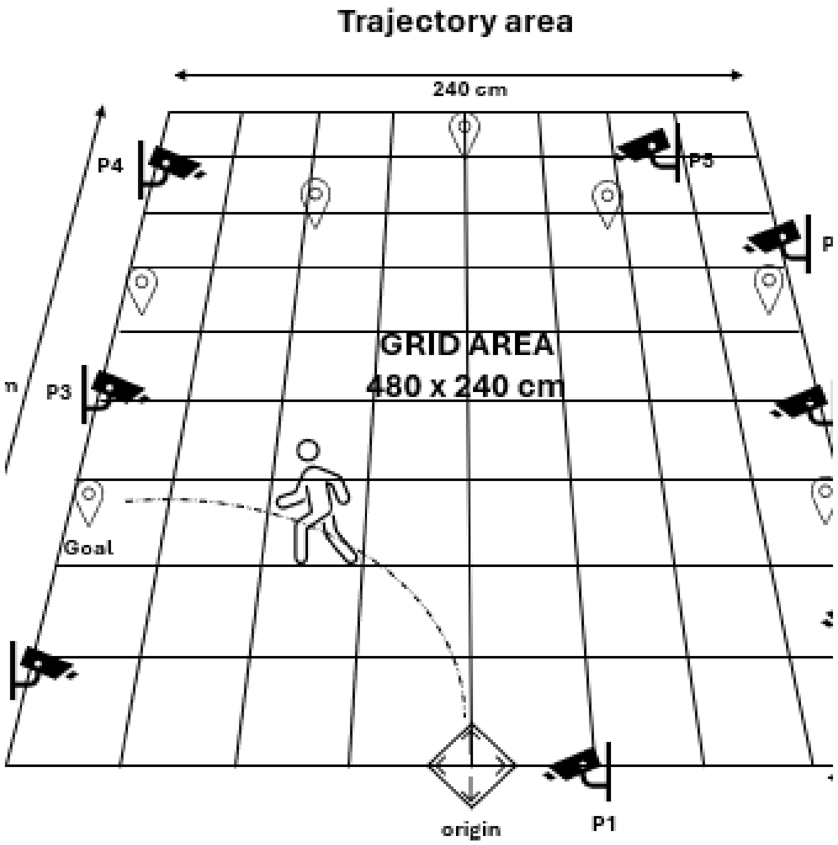
Estimating the mean velocity of human walking (Vx, Vy)

Predicting future points

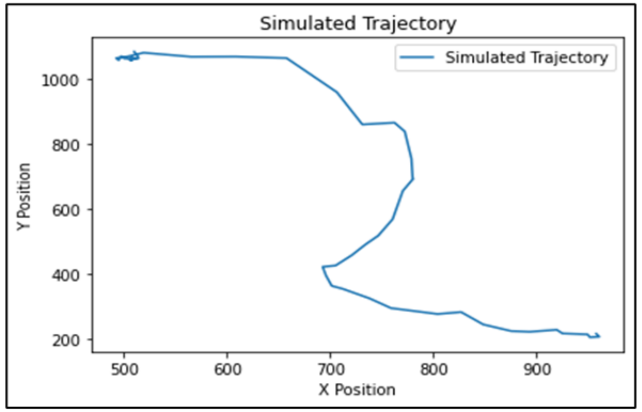
Defining the partial trajectory 0.5/0.7/0.9

Data collection

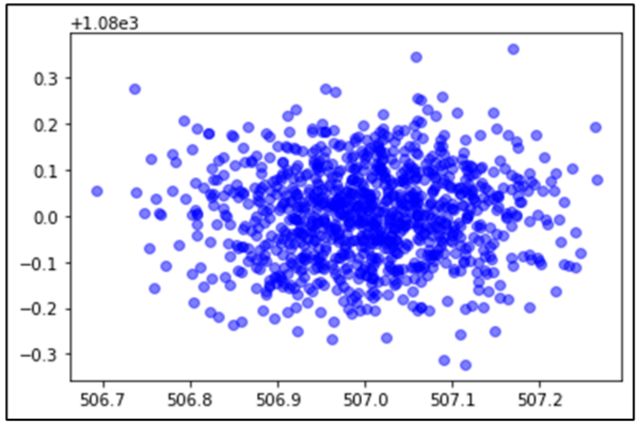
- Obtained the human trajectory points using computer vision
- Camera placed at a static position and participants walked naturally from source to destination
- Cameras placed at various positions and heights (1.5m to 2.4m) to capture diverse perspectives. Different camera angles including top view, angled view and side view are captured.
- Object detection model processed frames at 7 to 8 fps
- **Length of the path:** 6.6 meters (maximum length) with approximately 80-100 points (depends on the path length)
- **Number of participants:** 14 participants (5.7ft average height)
- **Average steps required from source to destination:** (11/8/6/4 steps depends on destination location)
- **Percentage of partial trajectory:** 50%, 70% and 80%
- **Destination points:** Predefined **nine locations** as destination points
- Participants were asked to walk straight with minor turns
- **Number of trajectory :** 227 trajectories



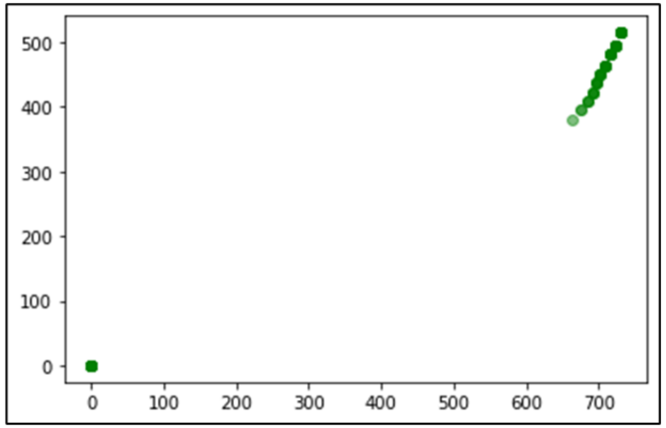
Overall particle filter working



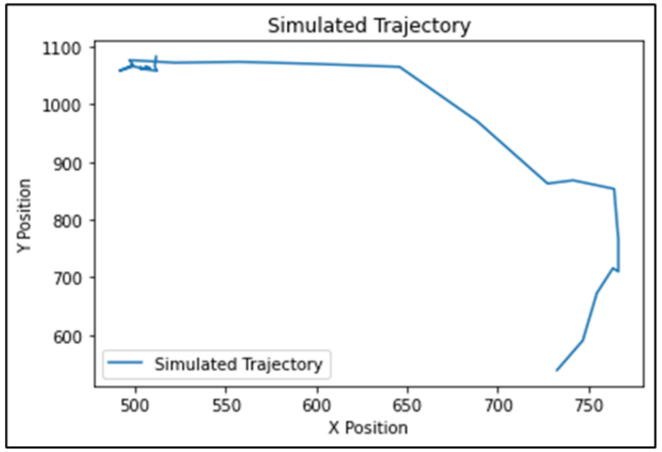
Actual trajectory from raw data



Particle initialization around initial position

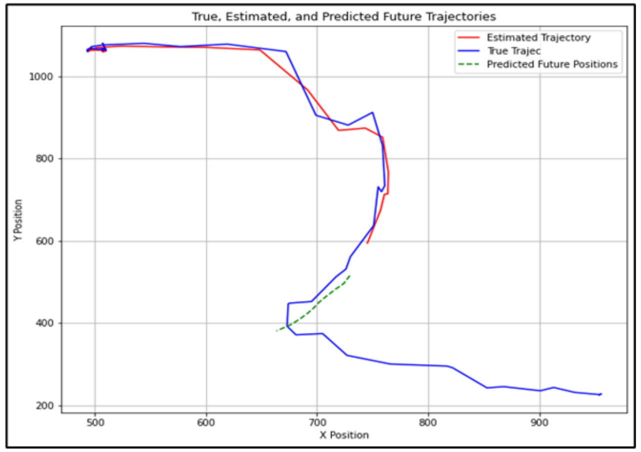


Path prediction using particles



60% of actual trajectory

- Predicting the trajectory for 60% path
- The figure shows 60% of trajectory prediction
- Based on previous velocity and weights prediction takes place for 10 future steps
- **Upcoming:**
- Accuracy testing using PF for all the 227 trajectories.



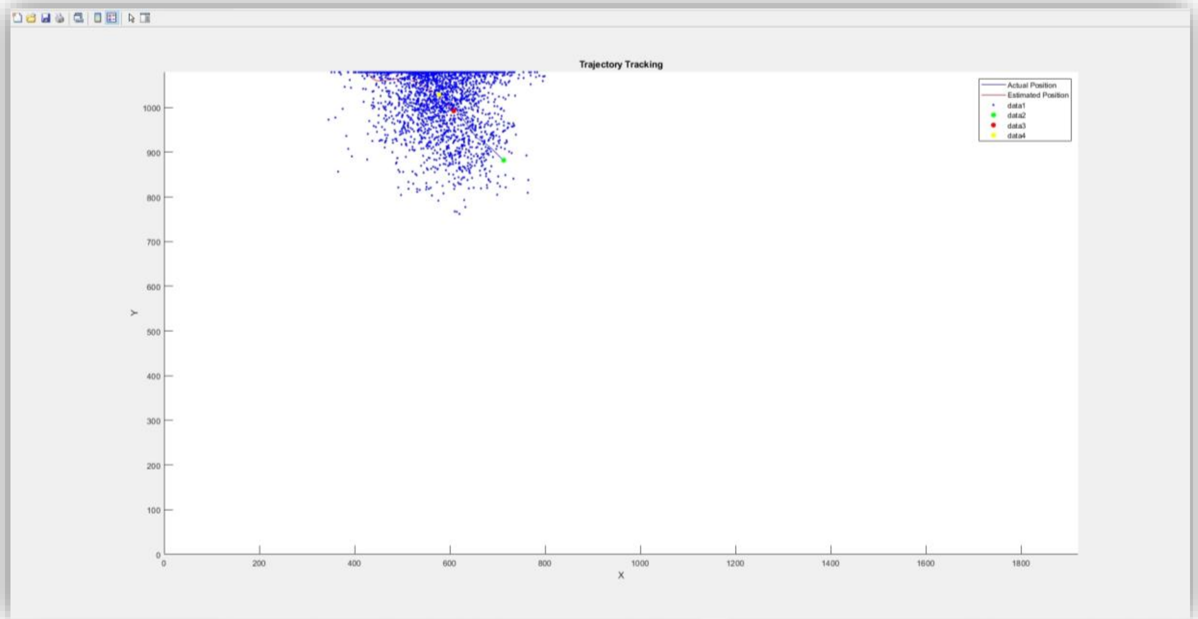
Prediction shown in green dots

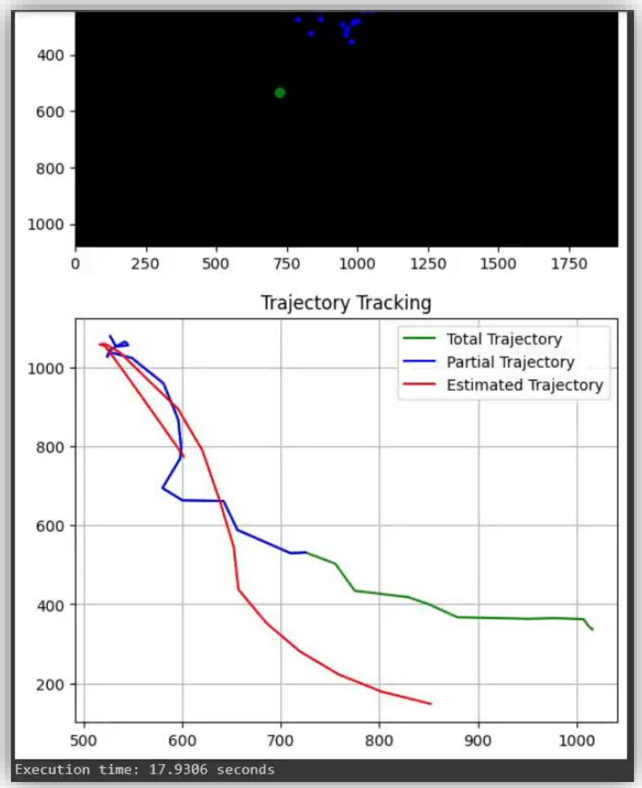
Implementation

- **Camera Integration:**
  - Object detection model to detect users
  - Coordinate values are sent to particle filter algorithm
  - Particles are uniformly distributed through out the area
  - Each particle represents possible state of the user
- **State Representation:**
  - Defining a state vector to include parameters such as position (x and y), velocity and acceleration
  - Velocity and acceleration are obtained from coordinate profile/data
- **Motion model:**
  - To predict future state of each particle "Motion model" is implemented
  - Our motion model consists of velocity and acceleration to simulate the human motion

PF test area

Simulations – Matlab & Python





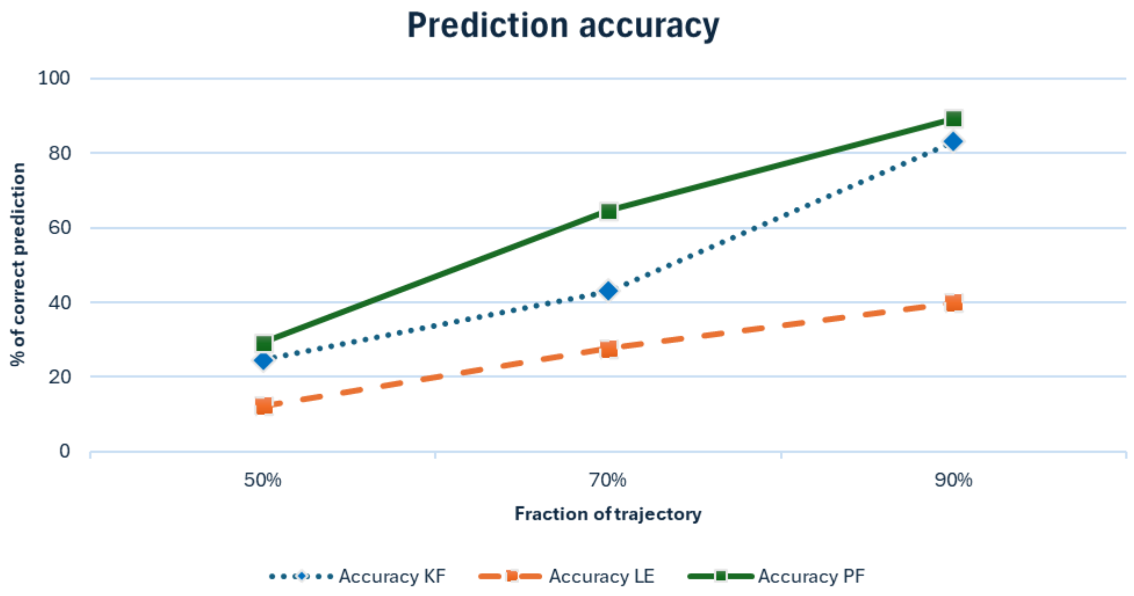
Number of particles	50% Partial Trajectory	70% Partial Trajectory	90% Partial Trajectory
1000	10.63	12.83	14.81
2000	10.83	13.97	15.91
3000	11.81	14.04	16.75

Table 1: Execution time (seconds) in MATLAB

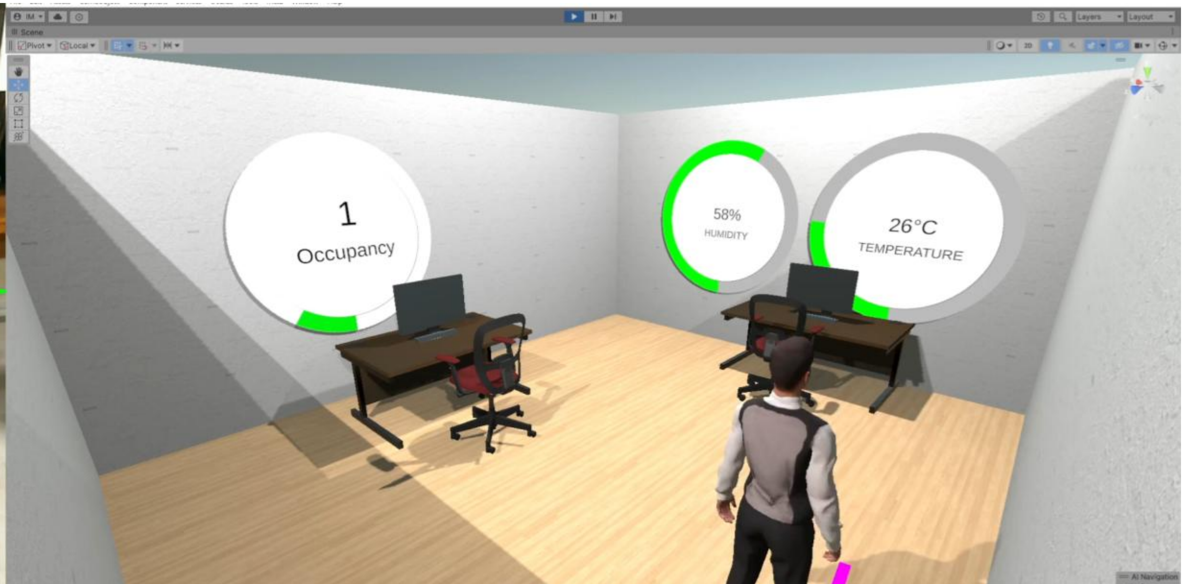
Number of particles	50% Partial Trajectory	70% Partial Trajectory	90% Partial Trajectory
1000	10.38	13.21	16.52
2000	11.36	14.33	17.86
3000	12.55	15.37	20.15

Table 2: Execution time (seconds) in python

Overall Accuracy Analysis



	Accuracy KF	Accuracy LE	Accuracy PF
50%	24.615	12.307	29.23
70%	43.076	27.692	64.615
90%	83.076	40	89.23



Concluding Remarks

- Three different ways of target / trajectory prediction
  - ANN for detecting phases of movements
  - Imitation Learning to learn trajectory from historical patterns
  - Particle Filter that does not require training data
- Standardizing evaluation criteria for trajectory prediction algorithms
  - Accuracy
  - Latency
  - Sensitivity – how quickly can we predict

Intelligent Inclusive Interaction Design

I3D

https://cambum.net pradipta@iisc.ac.in

SUNLUX
Collins Aerospace
INDIA EU
TATA CONSULTANCY SERVICES
SIEMENS
pulselabs
Microsoft
BT
DRDO
DEFENCE R&D ORGANISATION
MINISTRY OF DEFENCE
इसरो isro
faurecia
StanceBeam
facebook
ASHOK LEYLAND
wipro