

# Human Motion Tracking and Forward Position Prediction Using Optical Flow and Kalman Filtering

Gaurika Sethi  
B.Tech, Robotics & Automation  
University School of Automation & Robotics (USAR)  
Guru Gobind Singh Indraprastha University

## Abstract

This work presents a vision-based pipeline for tracking a walking human and predicting future positions using optical flow and a Kalman filter. Starting from a static-camera video, the system detects the person with background subtraction, tracks motion via Lucas–Kanade optical flow, and estimates future positions using a constant-velocity Kalman model in image coordinates. Experiments on a controlled single-subject walking sequence show that the Kalman filter smooths noisy optical-flow trajectories and produces accurate short-horizon predictions, with errors remaining within a few pixels for most frames. The approach is lightweight, does not require deep networks, and is suitable for robotics and defence applications where explainable tracking is preferred.

## 1 Introduction

Accurate motion tracking and short-horizon prediction are fundamental in applications such as autonomous robots, surveillance, and defence systems. Classical approaches based purely on optical flow often suffer from noise and jitter, especially for articulated human motion. On the other hand, probabilistic filters such as the Kalman filter can incorporate a principled motion model and fuse noisy measurements into a stable trajectory estimate.

This project implements a complete proof-of-concept pipeline for human motion tracking and prediction in a 2D image plane. The focus is on an interpretable, engineering-friendly system that can be adapted for research laboratories, for example in robotics or missile guidance testbeds.

## 2 Methodology

### 2.1 Foreground detection

The input is a static-camera video of a single person walking from left to right on a plain background. Foreground is extracted using the MOG2 background subtractor from OpenCV. After morphological opening and closing, the largest contour in the foreground mask is assumed to be the person, and a bounding box is drawn.

### 2.2 Optical flow tracking

Within the person bounding box, good feature points are detected using Shi–Tomasi corner detection. These points are tracked across frames using the pyramidal Lucas–Kanade optical flow algorithm. At each frame, the mean of the valid tracked points defines a pseudo-centroid  $(x, y)$ , which provides a smoother representation of body motion than the raw bounding-box centre. This sequence of centroids forms the measured trajectory.

### 2.3 Kalman filter model

A 4D state is used,

$$\mathbf{x}_k = [x_k, y_k, v_{x,k}, v_{y,k}]^T, \quad (1)$$

with a constant-velocity transition model:

$$\mathbf{x}_{k+1} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{w}_k, \quad (2)$$

where  $\mathbf{w}_k$  is zero-mean Gaussian process noise. The measurement is the optical-flow centroid:

$$\mathbf{z}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \mathbf{x}_k + \mathbf{v}_k. \quad (3)$$

Both the process and measurement noise covariances are tuned empirically to allow small vertical gait oscillations while damping large velocity jumps.

The filter is initialised directly from the first valid optical-flow centroid, avoiding the large transients that arise from arbitrary initialisation at the origin. For visualisation, the corrected state  $(x_k, y_k, v_{x,k}, v_{y,k})$  is used to predict a point a few frames into the future:

$$(x_k^{\text{pred}}, y_k^{\text{pred}}) = (x_k + \alpha v_{x,k}, y_k + \alpha v_{y,k}), \quad (4)$$

with  $\alpha \in [1, 2]$ .

### 3 Experimental Setup

The experiment uses a synthetic but realistic video: a single person walking sideways across a uniform background, recorded at  $1280 \times 720$  resolution. The camera remains static and lighting is uniform. All algorithms are implemented in Python using OpenCV and NumPy.

During each frame, the following steps are executed:

1. foreground mask via MOG2 and contour-based bounding box,
2. optical-flow-based centroid update,
3. Kalman prediction and correction,
4. logging of measured and predicted positions.

A CSV log of frame index, measured coordinate, and predicted coordinate is used for offline analysis.

### 4 Results

Fig. ?? shows the measured and Kalman-predicted trajectories in the image plane. The optical-flow-only trajectory exhibits small jitter due to limb motion and local tracking noise, while the Kalman trajectory is smoother and follows the global walking direction.

The Euclidean prediction error per frame is plotted in Fig. ?. For most frames, the error remains within a small number of pixels, indicating that the constant-velocity model is adequate for short-horizon prediction on this sequence.

The error histogram in Fig. ? shows that the majority of frames fall in a low-error regime, with a small tail corresponding to gait transitions and brief local tracking glitches.

### 5 Conclusion and Future Work

This project demonstrates a complete, explainable pipeline for human motion tracking and short-horizon prediction in a monocular video. Background subtraction and contour detection are used for person localisation, Lucas–Kanade optical flow provides dense local motion, and a 4D Kalman filter fuses these measurements into a smooth trajectory and forward prediction.

Future work includes extending the model to multiple targets, handling camera ego-motion, and exploring non-linear filters such as the Extended Kalman Filter and particle filters. Integration with 3D pose estimation and depth sensors would make the system directly applicable to robotics and defence scenarios requiring precise target tracking and interception.

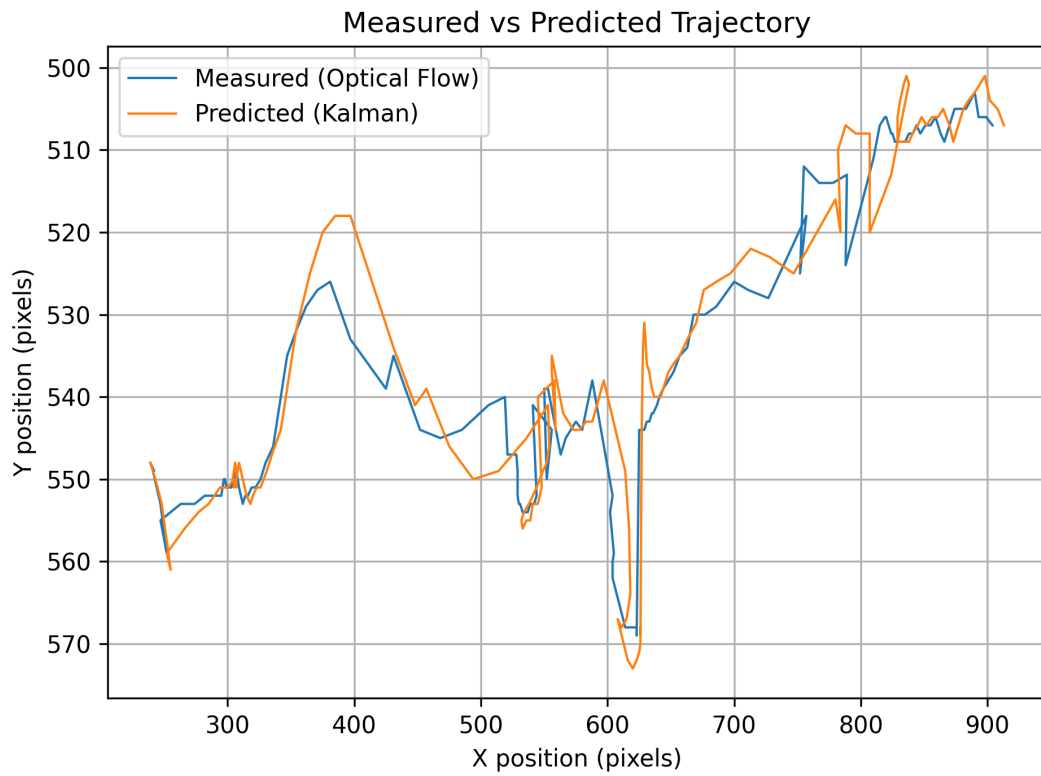


Figure 1: Measured (optical flow) vs. Kalman-predicted trajectory in image coordinates.

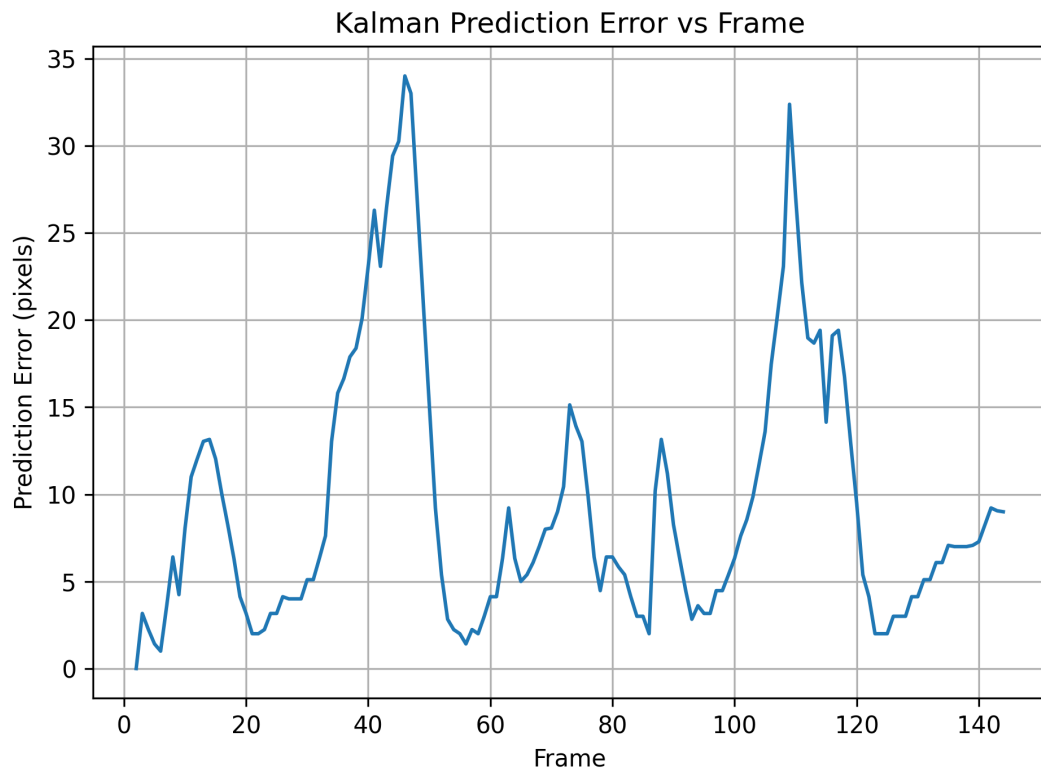


Figure 2: Prediction error (pixels) as a function of frame index.

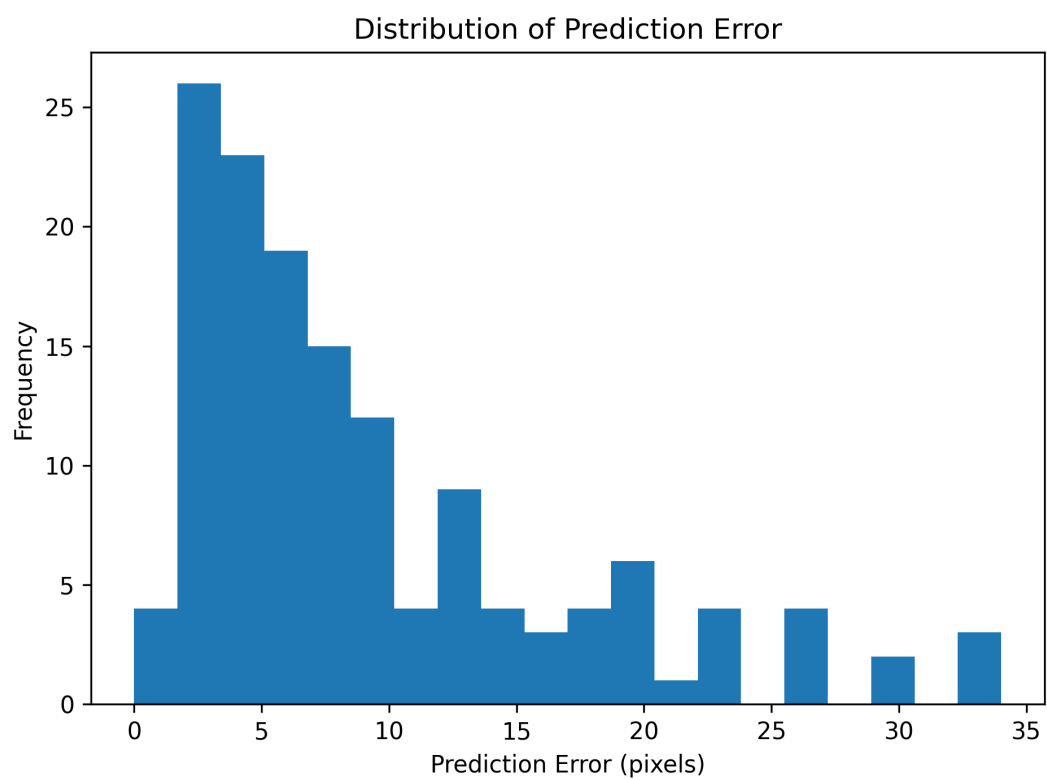


Figure 3: Distribution of Kalman prediction error across frames.