Vison Based Pedestrian Tracking for Advanced Driver Assistance System

SOUMYADEEP MUKHERJEE
11CH30022

CHIPES LAB, NTU

Abstract

This work presents an implementation of a vision-based system for tracking multiple moving and stationary pedestrians from a camera feed over a moving vehicle whose number is unknown and varies over time. Pedestrians are first detected and the detection information is used to track them and follow their trajectories in simultaneous frames. The proposed method has been tested on vehicle mounted camera feed and results are presented.

Currently, work is going on a proposed method for improving the detection from tracking information.

Executive Summary

This report gives an implementation of vision-based pedestrian tracking system for Driver Assistance Systems which includes multiple object tracking on moving camera feed mounted on a vehicle. The implementation mentioned implements the following methods for tracking:

- Kalman Based Tracking of pedestrian center with 4D State Vector with Naïve Data Association.

Kalman Filter based multiple pedestrian tracking with a state vector consisting of pedestrian position and velocity only and the cost matrix to associate pedestrians in previous frame to next is propagated using center distance only.

- Acceleration added to Kalman State vector

Acceleration Model is added to Kalman Filter state vector making it a 6-dimensional matrix and process noise term is assumed zero in Kalman-Filter.

- Motion-Agreement Tracking Based Data Association

Pedestrians in previous frame need to be associated to pedestrians in next frame for which a cost matrix is propagated using Motion-Agreement method where pedestrians are represented by multiple local patches for which a local estimate of direction of motion. This estimate along with global estimate of pedestrian motion is used to identify object stable regions to enable better cost matrix propagation. Hungarian algorithm is then used to associate data from the cost matrix.

- Ego-Motion Estimation

The video feed is from camera mounted on moving vehicle so motion due to camera needs to be removed. So Ego-motion is estimated by finding out the affine transform from good features using KLT.

- Kalman Tracking extended to Pedestrian Bounding Box

Pedestrian Tracking was extended from center point tracking to the whole boundary box by tracking top-left and bottom-right corners of every pedestrian's boundary box.

Acknowledgements

I take this opportunity to express my profound gratitude and deep regards to my guide **Wu Meiqing** for her exemplary guidance, monitoring and constant encouragement throughout the course of this project. The help and guidance given by her time to time shall carry me a long way in the field of Computer Vision.
I also take this opportunity to express a deep sense of gratitude to **Lam Siew Ki** for his cordial support, valuable information and guidance, which helped me in completing this internship.

I am obliged to staff members of CHiPES Lab, for the valuable information provided by them in their respective fields. I am grateful for their cooperation during the period of my assignment.

Lastly, my parents and friends for their constant encouragement without which this assignment would not be possible.

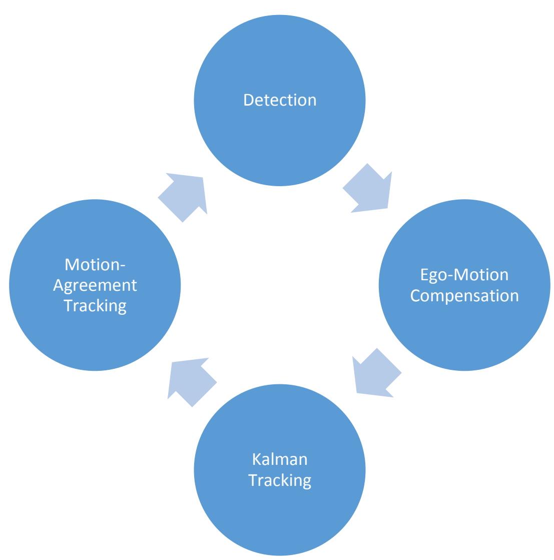
Table of Contents

Introduction

Vehicles that automatically perform safety tasks like detection of pedestrians will have an important role in the future of an intelligent transportation system. The possibility to appropriately equip a high number of vehicles will allow to reduce the casualties derived from accidents, especially in the urban environment. Pedestrian localization in outdoor scenes is a challenging task because of the variety of the environments and of the clothes. A moving vehicle has to deal with a lot of problems: noise produced by the presence of buildings and human artifacts, different illumination conditions, obstacles and so on.

The performance of pedestrian tracking systems has steadily increased during the past few years. Two factors mainly contributed to the improvement: the advance of robust pedestrian detectors and various extensions of the data association technique. Also for driver assistance systems, the ego-motion also needs to be compensated for a reliable tracker.
Developing a high-performance pedestrian detector remains an unsolved problem in computer vision. To handle the often poor quality of detections, online tracking systems have been developed. Our method proposes to track the detected pedestrians and use this tracking information to improve the detection performance.

Overall Algorithm



Ego-Motion Estimation

The goal of estimating the ego-motion of a camera is to determine the 3D motion of that camera within the environment using a sequence of images taken by the camera. The process of estimating a camera's motion within an environment involves the use of visual odometry techniques on a sequence of images captured by the moving camera. This is typically done using feature detection to construct an optical flow from two image frames in a sequence generated from a camera.

Features are detected in the first frame, and then matched in the second frame. This information is then used to make the optical flow field for the detected features in those two images. The optical flow field illustrates how features diverge from a single point, the *focus of expansion*. The focus of expansion can be detected from the optical flow field, indicating the direction of the motion of the camera, and thus providing an estimate of the camera motion.

Algorithm:

- Good features are extracted from consecutive grayscale frames and are matched using optical flow based KLT.
- Outliers are removed i.e. feature points which are not part of the background using RANSAC.
- Affine Transform is calculated from the optical flow of inlier points to get ego-motion compensation matrix.
- Obtained Matrix is used to find the corresponding points of Bounding Box corners i.e. ego-motion compensated coordinates.

Kalman Filter Based Tracking

Filtering is a very used method in engineering. A good filtering algorithm can reduce the noise from signals while retaining the useful information. The Kalman filter is a mathematical tool that can estimate the variables of a wide range of processes. It estimates the states of linear systems. The discrete Kalman filter is characterized by both a process model and a measurement equation.

The process model is characterized by the assumption that the present state, X_k , can be related to the past state, X_{k-1} , as follows:

$$X_k = F_k X_{k-1} + W_k$$

Where W_k is assumed to be a discrete, white, zero-mean process noise with known covariance matrix, Q_k ; F_k represents the state transition matrix which determines the relationship between the present state and the previous one.

In our case we try to track the state of a contact based on its last known state. Here, the state vector consists of a two two-dimensional positions expressed in Cartesian coordinates, a two-dimensional velocities and a two-dimensional accelerations.

By considering a constant acceleration, the state transition matrix can be determined from the basic kinematic equations as follows:

$$S_k = S_{k-1} + v_{k-1}t + \frac{1}{2}a_{k-1}t^2$$
$$v_k = v_{k-1} + a_{k-1}t$$
$$a_k = a_{k-1}$$

Where *s* is defined to be the contact's position, *v* is its velocity, *a is* the contact's acceleration and *t* is the sampling period. In a matrix form, the above equations can be written as:

$$\begin{pmatrix} {}^{tl}_{x,k}s \\ {}^{tl}_{y,k}s \\ {}^{br}_{x,k}s \\ {}^{br}_{y,k}s \\ {}^{tl}_{x,k}v \\ {}^{tl}_{y,k}v \\ {}^{br}_{x,k}v \\ {}^{br}_{y,k}v \\ {}^{tl}_{x,k}a \\ {}^{tl}_{y,k}a \\ {}^{br}_{x,k}a \\ {}^{br}_{y,k}a \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} {}^{tl}_{x,k-1}s \\ {}^{tl}_{y,k-1}s \\ {}^{br}_{x,k-1}s \\ {}^{br}_{y,k-1}s \\ {}^{tl}_{x,k-1}v \\ {}^{tl}_{y,k-1}v \\ {}^{br}_{x,k-1}v \\ {}^{br}_{y,k-1}v \\ {}^{tl}_{x}a \\ {}^{tl}_{y,k-1}a \\ {}^{br}_{x,k-1}a \\ {}^{br}_{y,k-1}a \end{pmatrix}$$

Here, the subscripts *x* and *y* refer to the direction of the contacts position, velocity and acceleration in the two-dimensional plane. The value of the sampling period is set to 1.

The measurement equation is defined as:

$$z_k = H_k x_k + v_k$$

Where z_k represents the measurement vector, v_k is assumed to be a discrete, white, zero-mean process noise with known covariance matrix, R_k. The matrix H_k describes the relationship between the measurement vector, z_k, and the state vector, x_k. Given the fact that the state vector is of length twelve and measurement vector is of length four, the matrix H_k must be of length twelve by four.

After the state equation and measurement equation of motion model are defined, in the next frame, Kalman filter is used to estimate the object's location, and to gain trajectories of moving objects.

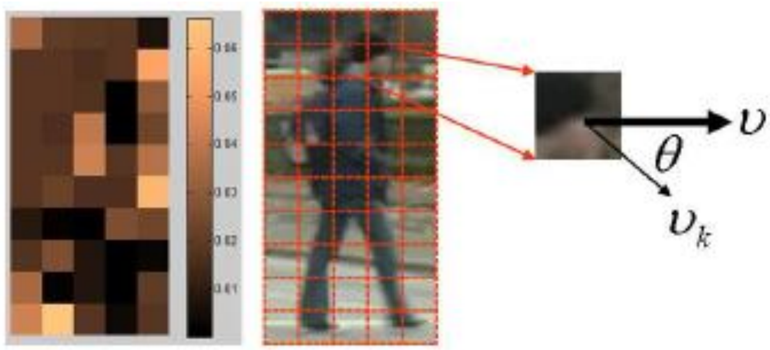
Algorithm:

- Kalman Filter is initialized with state vector, transition vector and process noise vector.
- Check if Existing Tracks > ZERO, If yes all detections added as Track.
- Cost matrix is initialized.
- Data Association from Cost Matrix is performed using Hungarian Algorithm.
- For all assigned tracks, if cost of assigned track to detection is greater than set threshold then count as frame skipped.

- For all tracks, if no of frames skipped is more than set frames skipped than remove track and its assignment.
- For all detection, find unassigned detections and make them new tracks.
- For all assignments, update Kalman with measured coordinates from detection data.
- A point is drawn at the tracked point of every track.

Motion Agreement based Cost Matrix Propagation for Data Association

Motion Agreement based Tracking works on the concept of sequentially updating the appearance model of each target by indirectly evaluating the motion consistency among its local patches. We show that a distance measure based on appropriately re-weighted local patches will successfully reduce tracker errors especially that lead to track fragmentation and track switching as mentioned in [].



Algorithm:

- Every detected pedestrian in bounding box is divided into 64 patches where each patch is represented using 64-bin histogram in HSV space.
- Local displacement of each patch for every detection is calculated by finding out the minimum histogram intersection.
- The patch velocity along with global velocity of detection from Kalman are used to calculate weight of the patch for the cost function.
- Cost function of object and candidate detection is a function of the sum of weights of each patch along with their boundary box intersection to union ratio.
- These weights are updated in every frame.

Results

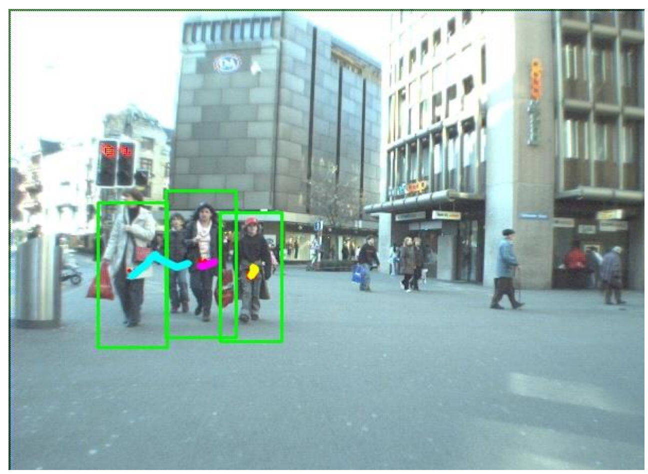
Ego-Motion Estimation

Before Implementation:

After Implementation:

Smoothening of trajectory due to compensation of camera motion.

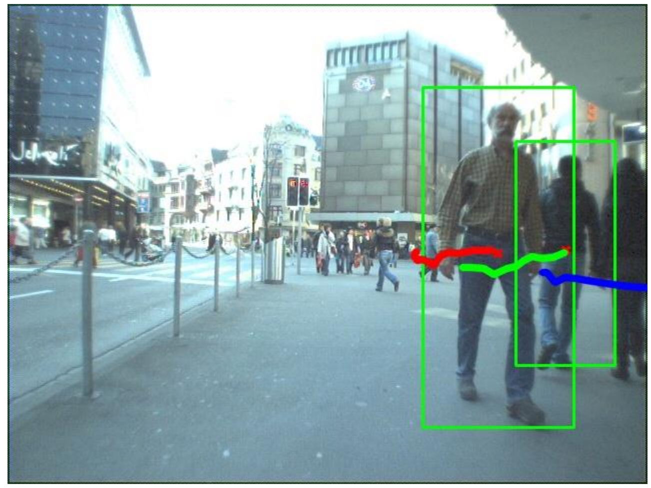
Kalman Filter Based Tracking



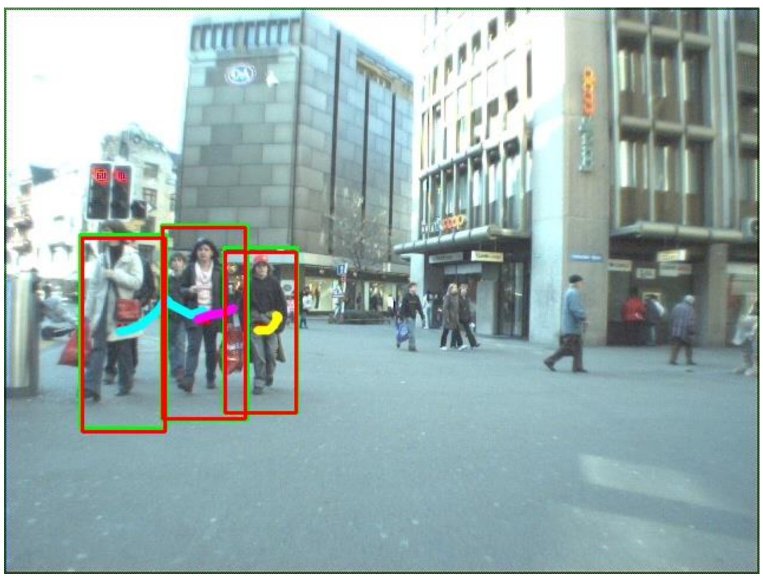
Motion Agreement based Cost Matrix Propagation for Data Association

Before Implementation:

After Implementation:



Kalman Tracking Extension

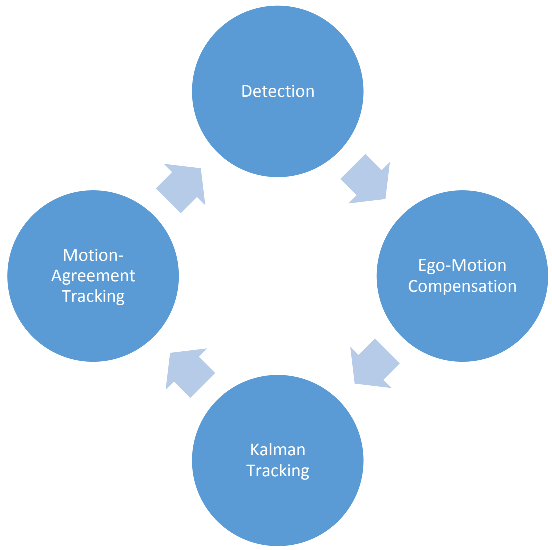


Boundary Box Tracking

Current Research – Proposed Method

Current Implementation uses Detection information directly for Tracking and no feedback from tracking is sent back to detection. On missing a detection, the tracker maintains the predicted bounding box where the pedestrian can be so this data can be used to improve detection.

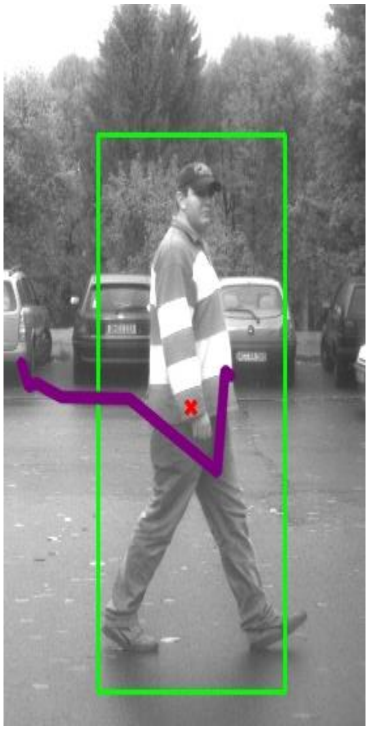
Predicted Data of Bounding box on missed detection can be used to call the detector on the specific ROI i.e. Predicted Bounding box with re-calibrated threshold to re-detect followed by associating this detection to previous detection.



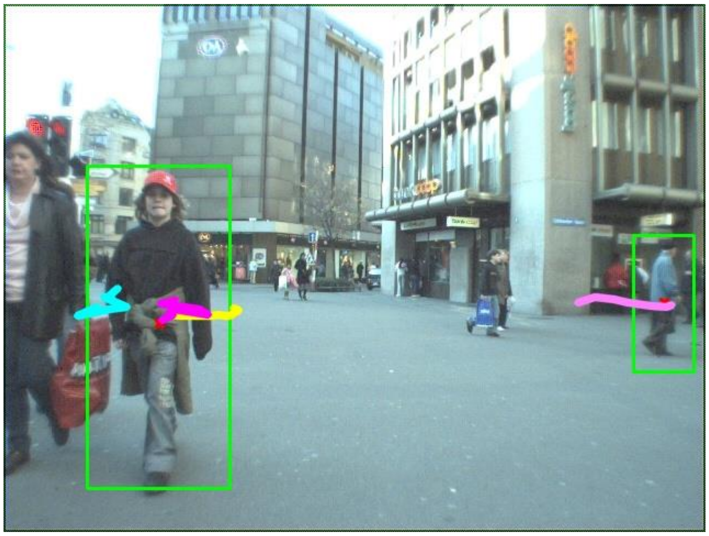
Drawbacks in Method

- Highly dependent on detection. If detection fails completely, tracking will not work. As shown, false detections are also tracked.

- Kalman filter tracking is for linear and Gaussian model only which often fails in complex scenarios and motions i.e. some situations of camera motion + pedestrian motion together.



- Parameter tuning for MAT varies in different data sets and at times data association fails causing intermingling of trackers.



- Features are extracted for HOG as well as ego-motion estimation so speed of processing is slow.

Possible Improvements

- Better data association methods can be tested. E.g. MHT (Multi-Hypothesis Tracker).
- Hysteresis can be incorporated during data association to ensure trackers do not get intermingled within themselves.
- Maintaining an Entry-Exit data of pedestrians in frame can be used to improve both detection and tracking.
- Nonlinear models like particle filters or PSO can be used which do not assume Gaussian noise only.

Conclusion

In this report, I researched on multi-pedestrian tracking for driver assistance systems based on Kalman filter. First to establish a motion model to track the centroid and then extended to the bounding box. Tracking performance is improved by compensating the camera motion and applying a Motion Agreement based cost propagation for dynamic cost matrix propagation in order to get an improved data association. Furthermore, a novel method to improve detection has been designed to improve the performance of the Driver Assistance System.

Experimental selected different scenes have been tested for effectiveness and robustness of the algorithm.

Bibliography

Caius SULIMAN, C. C. (2010). Kalman Filter Based Tracking in an Video. *10th International Conference on DEVELOPMENT AND APPLICATION SYSTEMS.* Suceava, Romania.

Xin Li Kejun Wang, W. W. (2010). A Multiple Object Tracking Method Using Kalman. *International Conference on Information and Automation.* Harbin, China.

Zheng Wu, J. Z. (n.d.). Online Motion Agreement Tracking. Boston, USA.