Object Detection and Fusion in a Camera Network

Prof.Benjamin Noack
Faculty of Autonomous Multisensor systems
Otto Von Guericke Universität
Magdeburg, Germany
benjamin.noack@ovgu.de

Marko Ristic (supervisor)
Faculty of Autonomous Multisensor systems
Otto Von Guericke Universität
Magdeburg, Germany
benjamin.noack@ovgu.de

Ranadheer Podishetti
Digital Engineering
Faculty of Computer Sciences
Otto Von Guericke Universität
Magdeburg, Germany
ranadheer.podishetti@ovgu.de

Muhammed Faizaan Aslam
Digital Engineering
Faculty of Computer Sciences
Otto Von Guericke Universität
Magdeburg, Germany
muhammed1.aslam@st.ovgu.de

Spandan Gude
Digital Engineering
Faculty of Computer Sciences
Otto Von Guericke Universität
Magdeburg, Germany
spandan.gude@ovgu.de

Siva Matta
Digital Engineering
Faculty of Computer Sciences
Otto Von Guericke Universität
Magdeburg, Germany
siva.matta@ovgu.de

***Abstract*— Distributed state estimation and localization have become an integral part of modern networked tracking systems. The use of a single measurement to estimate a state parameter is often accompanied by large uncertainty. The main focus of our project is to study the effect on uncertainties of the estimated states when multiple measurements are employed. The location of an object under a constant turn model is tracked by five raspberry pi cameras with help of an object detection algorithm and these measurements are gathered at a central node and fused together to get a location estimate more accurate than that of one which is obtained using one measurement. Kalman filter algorithm has been used to estimate the location. we have found that the use of multiple measurements has significantly decreased the uncertainties.**

Keywords—Distributed Network systems, location estimation, Kalman filter, uncertainties

INTRODUCTION

Location estimation has become one of the important applications for state estimation. The accuracy and uncertainties of those estimates are critical in location estimation systems. When using only one sensor measurement is accompanied by a large uncertainty in the estimate. Making estimations in state estimation systems often rely on imperfect information. Imperfect is referred to any information that is different or lacking is ideally required. No sensor provides perfect and complete information about the state measured. Often systems are corrupted by disturbances, which can't be controlled or even measured. This can be dealt with by obtaining as many measurements from different sensors and using them all together to estimate the state.

BACKGROUND WORK

A. Kalman Filter

The Kalman filter (Kalman, 1960), (Wald, 1998)is an optimal estimator. It infers parameters of the model from indirect, inaccurate, and uncertain observations. It is recursive so that new measurements can be processed in the next time steps. If all the noise in the measurements is Gaussian, the Kalman filter minimizes the mean square error of the estimated parameters. Kalman filter yields results due to optimality and structure. It is possible to deal with real-time processing.

METHODOLOGY

In order to study the effect of the use of multiple measurements in state estimation, we want to measure the location of a remote-controlled car moving in a circular motion with a constant turn rate. Five raspberry pi's equipped with camera modules capture the images of the car on the ground at equal intervals of time.

B. Arrangement of the experimental setup

Figure 1 : experimental setup

The experimental setup consists of an aluminum frame setup mounted with five raspberry pi cameras placed at different locations facing the ground. The cameras are located at a height of 106cm. The camera network is such that a central pi is equidistant from the other four pi modules and makes right angles with the two other modules. All the raspberry pi modules are connected to the same network through an ethernet cable.

C. Communication among the camera modules

Communication among the pi modules is necessary to capture the motion of the car in regular intervals. This is achieved by building a server-client architecture with a local machine as a central server node and all other pi's as connecting client nodes. Both Server and the clients should be in the same network to be able to communicate with each other.

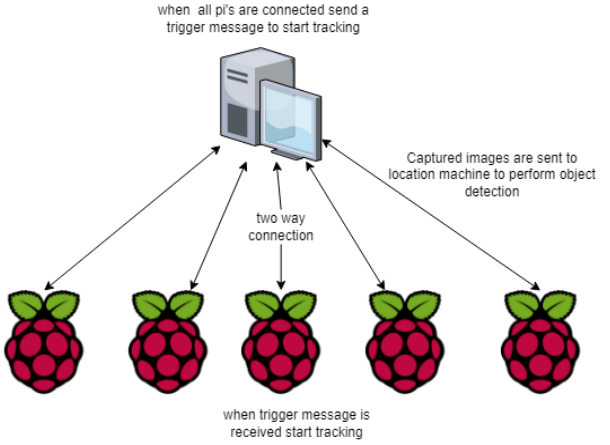

Figure 2 : Pi - Intercommunication

The communication is achieved using the Python Socketio library. The communication between the server and the clients is such that the server sends a message to all the client pi's to start capturing the motion of the car when all the clients are connected. The client nodes capture a series of images for a specified time. The images are sent to the central server where object detection is carried out.

D. Camera Calibration

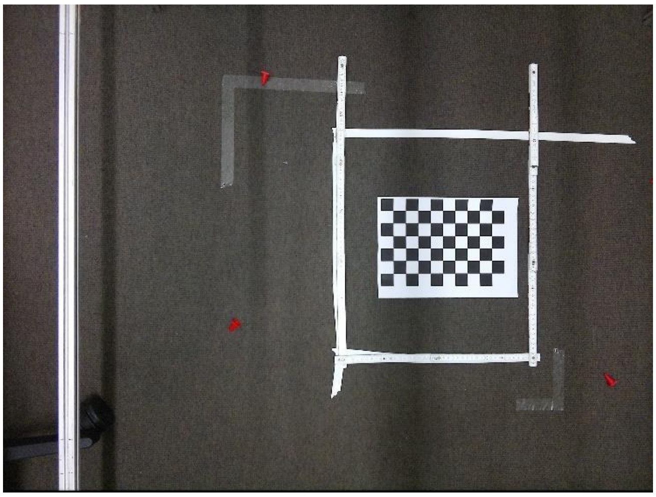

Figure 3 : Calibration

Camera calibration was done using a checkerboard pattern whose square dimensions are 2.5cm. each pixel was calibrated to about 65 pixels. The camera inclination was found out with respect to checkerboard alignment.

E. Object detection

Object detection is carried out using the Harris Corner method, (Stephens, 1988). The main idea is to find the difference in intensity for a displacement of (u,v) in a both x and y directions for the image,

$$E(u,v) = \sum_{x,y} w(x,y)[I(x+u,y+v) - I(x,y)]^2$$

We have to maximize the function $E(u,v)$ for corner detection

$$E(u,v) \sim [u \quad v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

Where

$$\mathrm{M} = \sum_{x,y} w(x,y) \begin{bmatrix} IxIx & IyIy \\ IxIy & IxIx \end{bmatrix}$$

Ix and Iy are image derivatives in x and y directions. Then comes R which determines if a window can contain a corner or not

$$\mathrm{R} = \det(\mathrm{M}) - \mathrm{k}(\mathrm{trace}(\mathrm{M}))^2$$

Where

$$\det(\mathrm{M}) = \lambda 1\, \lambda 2$$

$$\mathrm{trace}(\mathrm{M}) = \lambda 1 + \lambda 2$$

$$\lambda 1, \lambda 2 \; are\ the\ eigen\ values\ of\ M$$

The magnitude of eigenvalues decides whether a region is a corner or not. |R| is large, the region is a corner.

The object location is taken as the centroid of all the corners detected from the Harris corner detection algorithm. The location is estimated in terms of pixels.

F. Coordinate translation

As the cameras are not located at the same place, the coordinates obtained from all the cameras are not from the same coordinate system. So these coordinates need to be transformed into a world coordinate which is common to all the nodes.the camera located at the east of the central camera is assumed to in world coordinates and the location obtained from all other cameras are transformed to this world coordinates. Let X ,Y be the world coordinates and be the θ be the respective angle with the world coordinate. The Local coordinates are transformed according

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} dx \\ dy \end{bmatrix}$$

Where x,y are coordinates in local coordinates and dx and dy are translation distance from the camera at eastern node.

G. *Motion model*

A constant turn model with parameters X-coordinate x_k, Y-coordinate y_k, Velocity in X-direction Δx_k and Velocity in Y-direction Δy_k given by

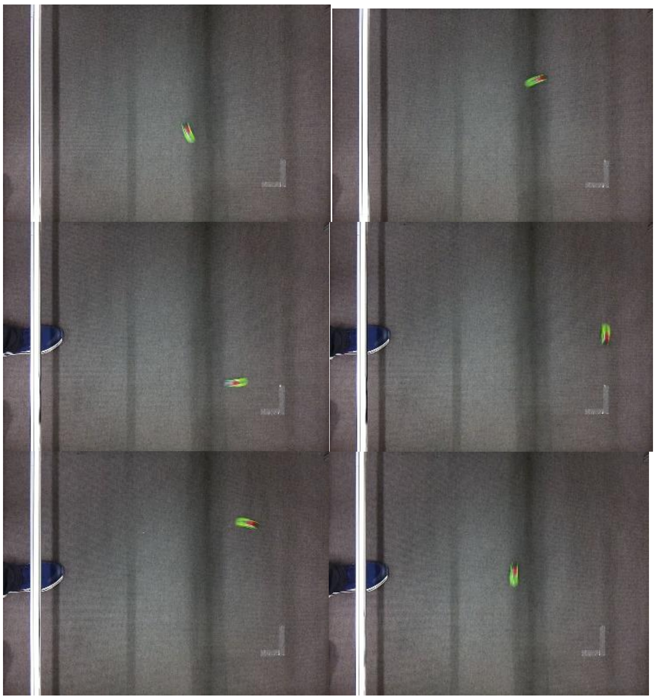

Figure 4 : constant turn motion of the remote controlled car

$$x_{k+1} = A.x_k + B.u_k + w_k$$

$$x_k = \begin{bmatrix} x_k \\ \Delta x_k \\ y_k \\ \Delta y_k \end{bmatrix}$$

The observation model is given by

$$Z_k = Hx_k + v_k$$

$$Z_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x_k$$

H. State Estimation

The Kalman filter is used in estimating the model parameters in a recursive loop where multiple measurements in each timesteps are used together.

The error covariance $w_k \sim \aleph(0, Q)$

$$Q = S \begin{bmatrix} \sigma_x^2 & 0 & \sigma_{x\Delta x} & 0 \\ 0 & \sigma_y^2 & 0 & \sigma_{y\Delta y} \\ \sigma_{\Delta xx} & 0 & \sigma_x^2 & 0 \\ 0 & \sigma_{\Delta xx} & 0 & \sigma_y^2 \end{bmatrix}$$

The Measurement Covariance is given

$$v_k \sim \aleph(0, R)$$

$$\mathrm{R} = \sigma_x I$$

$$R = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}$$

The Initial Uncertainity matrix

$$P = \begin{bmatrix} \sigma_x^2 & 0 & 0 & 0 \\ 0 & \sigma_{\Delta x}^2 & 0 & 0 \\ 0 & 0 & \sigma_x^2 & 0 \\ 0 & 0 & 0 & \sigma_{\Delta y}^2 \end{bmatrix}$$

Where σ_x^2 *is variance in x location,*
σ_y^2 *is variance in y location*
$\sigma_{\Delta x}^2$ *is variance in velocity in x direction*
$\sigma_{\Delta y}^2$ *is variance in velocity in y direction*

The prediction step

$$x_{k+1} = Ax_k$$

$$P_{k+1} = A_k.P_k.A^T + Q$$

The Update step is carried out as many as there are measurements in each time step

Computing the Kalman gain is given by

$$K_k = P_k.H^T(H.P_k.H^T + R)^{-1}$$

In each loop estimate is updated as follows

$$x_k = x_k.K_k(z_x - Hx_k)$$

Update in error covariance matrix

$$P_k = (1 - K_k H)P_k$$

RESULTS

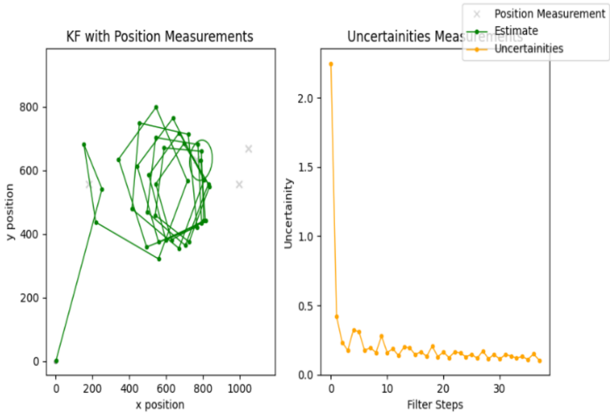

Figure 5 : Kalman filter position estimations when four measurements are used

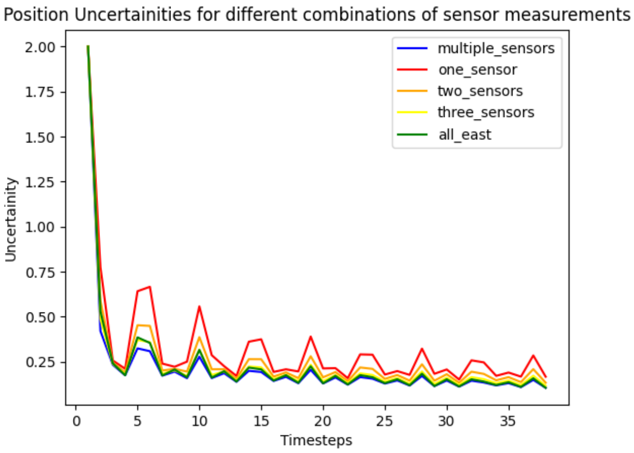

Figure 6:Uncertainties in position estimates when different combination of sensor measurements are used

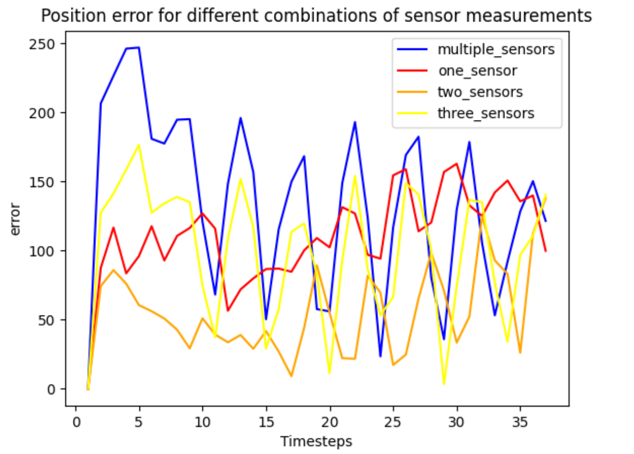

Figure 7 : Bias of estimates from different combination of sensor measurements are used. The ground truth was considered from one of the camera module measuremen

DISCUSSION

As seen from figure 5. The location estimates follow a circular path. Although the path is not smooth, if sampling rate of the measurements are increased, the path approaches a smooth circular path. The uncertainty also decreases pretty and converges smoothly. The goal of the project is to prove that employing multiple sensor readings gives us a better and certain estimate of the state variables. This can be illustrated in the figure 6. Where the uncertainty is the minimum in case of four sensor readings and maximum when a single measurement is used.

We tried to calculate the bias of the estimates, while assuming one of the sensor readings as ground truth and excluding that reading in the estimation process. In figure 7, we can see that the bias is seen to be large for multiple sensors. It is quite a opposite behaviour of what observed in terms of uncertainty. But this does not mean that the estimates are wrong instead it might be just different point than the measurement what we considered as a ground truth. We also tried to add different levels to single measurement to a single measurement and use them as multiple sensor data but that also seems to be less effective than that of measurements from four different cameras. However, it performs better than any of the single measurements.

CONCLUSION

Our project aims to find out that multiple sensor measurements would yield a better result for the estimation of a location of an object under a constant turn motion. Even with detection algorithm being noisy, Kalman filter was able to approximate the model with less uncertainty. The convergence of the model is faster than compared to single measurements. Although as long as the model assumed needs to be accurate and close to the real world for the measurements to have a effect. If the model itself is wrong, then the even ideal measurements would not lead to a perfect approximation. Kalman filter works as good approximation algorithms even with the noise data. Although multiple sensors provide a better uncertainty, the number of measurements at which the models works better is a point worth studying.

II. References

1. Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME--Journal of Basic Engineering*, 35--45.
2. Stephens, C. H. (1988). A combined corner and edge detector. 147--151.
3. Wald, I. C. (1998). Benefit of the future SPOT-5 and of data fusion to urban roads mapping. *International Journal of Remote Sensing*, 1519--1532.
4. Yufik, M. a. (294). Virtual associative networks: A new paradigm for sensor fusion. *Proceedings of the IEEE International Joint Symposia on Intelligence and Systems*, 1998.