CHAPTER 1

INTRODUCTION

1.1 MULTIPLE TARGET TRACKING

Multiple Target tracking (MTT) is a critical component of surveillance systems with one or more sensors. Its aim is to locate multiple targets, obtain their trajectories and maintain their identities. Traditional tracking algorithms are designed on the assumption that the sensors provide a set of position measurements based on which the parameters of interest about the target are estimated. A track is generally a state trajectory estimated from a set of measurements that have been associated with the same target. Depending on the environment and the application scenario, various types of sensors such as Radio Detection and Ranging (RADAR), Infrared (IR), SOund Navigation And Ranging (SONAR), Radio Frequency IDentifier (RFID) reader, GPS receiver and camera are employed. The Multiple Target Tracking (MTT) system makes use of data association algorithms and estimation filters to associate the observation of a sensor to estimate the position of multiple targets (Liggins et al 2009), (Bar-Shalom and Fortmann 1988), (Li et al 2001).

In the recent years, a significant attention has been focused on multisensor data fusion and MTT for both military applications like automated target recognition, guidance for autonomous vehicles, remote sensing, battlefield surveillance and civil applications like robotic navigation,
person/object detection and tracking, airport surveillance, health care, environmental monitoring and wildlife tracking.

The basic functions of MTT are target detection, recognition, identification, gating, data association & correlation, track initiation, deletion, filtering, prediction and decision making. The basic functional blocks of a simple MTT system are shown in Figure.1.1. In the MTT system, the data collected from the sensors are partitioned into sets of tracks. Once the tracks are formed and confirmed by successive scans, MTT estimates the number of targets and the position of the moving target in the surveillance area (Blackman 1986).

**Figure.1.1 Block diagram of a simple MTT system**

The measurement processing is performed using the input data received from the sensors. Data association combines multi-sensor data and correlates one set of sensor observations with another set of observations and produces a set of tracks for a target. Tracks refer to the discrete target states which define the target trajectory or movement at discrete moments over time. The observations of the new set of input are considered to update the existing tracks and the observations which are not assigned to existing tracks will initialize new tentative tracks. The MTT then estimates the states with an estimation filter and forms a gate around the predicted state to determine the
possible observation-to-track pairings. The state observations falling within
the gate are considered to be the update of the track. After the inclusion of
new observations, the MTT continues the initialization of new tracks and
updates the existing tracks with the next set of observations.

Traditionally, the states of the track are estimated using Kalman
Filter (KF) which involves the recursive fusion of noisy measurements to
produce an accurate estimate of the state of a system of interest. A key feature
of the KF is its representation of state estimates in terms of mean vectors and
error covariance matrices, where a covariance matrix provides an estimate of
the second moment of the error distribution associated with the mean
estimate. The square root of the estimated covariance gives an estimate of the
standard deviation. If the sequence of measurement errors is statistically
independent, the KF produces a sequence of conservative fused estimates
with diminishing error covariance (Liggins et al 2009).

Many estimation problems are non-linear, and hence, in order to
linearize the system, the KF is replaced with the Extended Kalman Filter
(EKF) (Bar- Shalom and Li 1993).

Target location detection and tracking in Wireless Sensor Network
(WSN) is a challenging research problem. The goal of the location based
tracking is to trace the paths of moving objects in the area where sensors are
deployed. One method of target tracking in WSN utilizes the Global Position
System (GPS). However, GPS cannot be used in most of the indoor
environments. Also in the non-urban outdoor settings, GPS does not yield
accurate results as it depends on the factors such as terrain, foliage and
topographical settings of the place where the target is located (Enge and Misra
1999). Since, GPS alone is not adequate for finding the location of the target,
WSN needs other sensors for finding the location of the target. Also, target
identification needs sensors. Therefore in this thesis, a scheme that uses RFID
and Passive Infra Red (PIR) sensor has been proposed to detect the target and identify the target accurately.

Though RFID based tracking of human is simple, tracking using camera network is preferred in secured zones because tracking multiple targets (people) from video sequences is an important task in surveillance. However, camera based tracking poses challenges such as occlusion and target miss association. In order to address these challenges normally KF is used (Zhao and Nevatia 2004). But in more complex scenarios, the blob-based analysis faces several problems, namely (i) a single blob may contain multiple humans due to their physical proximity or due to the camera viewing angle, (ii) a single object may be fragmented into several blobs due to low colour contrast, (iii) a blob may contain pixels corresponding to the shadows or reflections caused by the moving objects as well as noise (Zhao and Nevatia 2004). Due to these problems, the blobs may go through frequent structural changes. This thesis proposes a KF based approach using camera network, to track the target under occluded condition. The proposed scheme can be used to overcome target miss association problem.

1.2 CHALLENGES IN MULTITARGET TRACKING

Mathematical modeling of the target tracking process has been a topic of extensive study (Kishore and Pravas 1997). The KF and its variants have emerged as the preferred filters for tracking applications. In this approach the state of the tracked object consists of its position and the time derivatives of its displacement. The target motion models of the highest order include target acceleration, a second derivative of the target position. The approach presented by Bar-Shalom and Birmiwal (1982) does not rely on a statistical description of the maneuver as a random process, but introduces extra state components in the state model when the maneuver is detected.
In multiple target environments, new position reports are received at regular time intervals from ‘N’ number of targets, but the reports do not have labels indicating the targets to which they correspond. When the new positions are plotted, each report could be associated with any of the existing trajectories. With maneuvering targets, the tracking algorithm needs to be capable of identifying the beginning and the end of maneuvers. Many approaches to the maneuvering multiple target tracking problems have appeared in the literature (Mazor et al 1998, Wang and Hou 2009, Kishore and Pravas 1997, Tugnait 2004, Jeong and Tugnait 2005, Mori et al 2012, Jing et al 2012). Depending on the environment and the application scenario, various types of sensors and techniques can be employed.

In this research work, three schemes have been proposed for tracking multiple targets in Radar, WSN and camera sensor networks.

The major issues in MTT with multiple target scenarios are explained in the following sections.

1.2.1 Target Tracking in Radar Sensor Network

MTT deals with tracking multiple targets in a scenario where the number of targets may not be known and can also be varying with time, and the measurements obtained could have been originated from any target. In target tracking applications, each target is associated with a state which is represented as a vector that contains parameters defining the position of the target, its distance, velocity, azimuth elevation etc. In MTT, the observation corresponds to the measurement of a target state by a sensor at discrete moments in time. Based on the observation and the prediction, an estimate about the true target state is made. The estimate is the corrected state of the target that depends upon the variances of both observation and prediction.
The challenges associated with MTT (Blackman and Popoli 1999) include nonlinearity associated with target states, number of targets, number of sensors, the type of filter used for estimation and sensor update rate. The targets are classified as maneuvering and non maneuvering and the tracking is performed for both single and multiple target environments. For a single target case, the algorithm processes data from sensor reports and accumulates a series of positions and estimates the velocity of the target to predict its future position. It also includes certain challenges related to the uncertainty in position measurements. Singer (1970) has proposed a simple target model that closely represents the ensemble behavior of different maneuvering target and KF yields a tracking algorithm that provides optimal performance for this class of targets.

The challenges in MTT are classified into two parts: (i) data association (ii) tracking of maneuvering targets. The data association problem is due to the fact that multiple observations fall in a gate and disturb the system response when number of targets increases (Feo et al 1997). The radar sensor is mostly used as an active sensor, i.e., it emits energy and then detects the return. Hence, from the time delay and the speed of light, it is possible to calculate the distance of the target. For target tracking under radar surveillance, the radar periodically scans the volume under surveillance and provides detections that indicate a target’s presence. Moreover, the Radar provides the information on the spherical co-ordinates of each detected target in terms of range, elevation, azimuth and possibly radial speed. The sensor measurements are affected by noise which is modeled as Gaussian, zero mean, with certain standard deviation. As the sensor repeats the scan of the volume, new plots are gathered; the sequences of plots, most likely pertaining to the
same target, can be processed in a filter to smooth the measurement noise, thus providing target tracks (Liggins et al 2009).

In many situations targets are tracked by multiple sensors. The decision process involved in associating tracks belonging to the same target results in a correlation problem. Once the measurements are correlated with the targets, an algorithm is needed to provide a single target track which has uncertainty lesser than that of the individual tracks. This process is often referred as track fusion. The problem of track-to-track association arises in multisensor surveillance systems where each node has its own local information processing system and keeps its track. A major problem is in making decision on the tracks coming from different sensor systems for the same target, i.e., the problem of track-to-track correlation. When two tracks are decided to be from the same target, the next problem is to combine (fuse) the two track estimates together (Chang et al 1997).

The track-to-track association and subsequent fusion problems have been considered under the assumption that the estimation errors in two track files from different sensors, but corresponding to the same target, are independent. The errors in two such track files are correlated because the process noise of the target is the same (Bar-Shalom 1981).

In some applications, tracks are maintained for a large number of targets for surveillance. In Radar based surveillance system, there is a need to have a scheme with improved tracking accuracy, less computation complexity and less memory requirement. Hence new correlation logic is proposed in this thesis to estimate the target position quickly compared to the existing methods.
1.2.2 Target tracking in Wireless Sensor Network

Target tracking is considered to be one of the most demanding requirements in the applications of WSNs like, wild life monitoring, security applications for buildings and compounds to prevent intrusion, disaster management and international borders monitoring for illegal crossings. Most of the modern surveillance and monitoring systems deploy WSN for detecting and tracking the target (people/vehicle) and also have variety of distributed sensors established in a variety of applications.

The work done by Wang et al (2003) is one of the earliest attempts using tiny acoustic sensor devices for tracking purposes. The target position is estimated via triangulation, i.e., comparing the difference in sound propagation delays from the sound source to different acoustic sensors. Unfortunately, this study assumes that the sensor readings are free of ambient noise, which is a highly unrealistic assumption. In a typical outdoor environment, it is expected that the sensor readings would be significantly influenced by ambient noise. In addition, given that the sensor nodes are small form factor devices and of low cost, it is expected that the readings would contain inherent noise due to their low fidelity. All these solutions require the presence of a co-operative target to generate specific sequence of signals, i.e., RF and audio or a target equipped with inertial sensors as discussed by Klingbeil and Wark (2008).

Many problems of target tracking require a good state estimate, even if the system changes rapidly. The KF is suitable for the estimation of the state of targets moving with nearly constant velocity. But when the target starts maneuvering, the single KF is not suitable for tracking. The velocity model may not be suitable for maneuvering target tracking and acceleration model may not be suitable for non maneuvering target tracking. Li and Jilkov (2003) have suggested a methodology for selecting exact models precisely for
maneuvering targets. Hence multiple model (MM) filters are needed for efficient maneuvering target tracking. A modification to the Kalman filter, called the Interacting Multiple Model (IMM) (Blom and Bar-Shalom 1988, Bar-Shalom and Li 1995, Mazor et al 1998) has emerged as a solution to handle the maneuvering problem. In IMM estimator, multiple models with different structures are used to describe the motion of the target. The IMM makes use of bank of KF to accommodate various possible target trajectory patterns. The final estimate is obtained as the weighted sum of estimates from sub-filters of the different models (Chen et al 2007) and switching between models is obtained as per Markov transition probability matrix (Mazor et al 1998).

The accuracy of tracking a target in WSN is limited by the sensor detection capability and the packet loss in the wireless medium. Due to random deployment of sensor nodes, sensor devices are likely to exhibit unreliable behavior and the events might not generate samples at regular intervals of time. The target may be detected by more than one sensor or sometimes may not be sensed even by a single sensor. Moreover, the location information of the sensor is not accurate in indoors and the events are likely to be missed. In order to overcome these challenges, there is a need for a logic to predict the future position of a maneuvering target based on the past dynamics. In this thesis, IMM estimator which uses velocity and acceleration models is proposed for tracking targets in WSNs in order to improve the tracking accuracy.

1.2.3 Target Tracking in Camera Sensor Network

Automatic object tracking is the critical task in many computer vision applications such as surveillance, augmented reality, smart rooms, warehouse, etc. In single camera tracking techniques, it is a common practice to assume that distinct targets have distinct appearances with respect to colour,
texture, size, or contour features (Iketani et al 1998). These schemes face a fundamental limitation caused by occlusions. There are numerous single camera detection and tracking algorithms (Anurag and Davis 2002). All these algorithms face similar difficulties such as, occlusion, changes in lighting, background & appearance, presence of other dynamic objects, complex object interactions such as shadows & reflections, tracking of unwanted objects and target miss associations (Comaniciu et al 2003).

Krumm et al (2000) have dealt the problem of people tracking in real time for an interactive environment using depth based background subtraction model. This technique models the background as static environment, but it requires more texture information for dense stereo reconstruction. Kang et al (2004) have dealt multiview tracking of moving objects observed by multiple cameras to continuously track moving objects using Joint Probability Data Association (JPDA) filter.

Orwell et al (1999 a) have presented a tracking algorithm to track multiple colour objects using multiple cameras. They modeled the connected blobs obtained from background subtraction using colour histogram techniques and used them to match and track the objects. Orwell et al (1999), present a multi-agent framework for determining whether different agents are assigned to the same object seen from different cameras. This method faces problems in the case of partial occlusions where a connected foreground region does not correspond to one target, but has parts from several of them.

Cai and Aggarwal (1998) have extended the tracking system by starting with single camera view and switching to another camera, when the system predicts that the current camera will no longer have a good view of the object. Intille et al (1997) have presented a system which is capable of tracking multiple non-rigid objects. The system uses a top-view camera to identify the individual blobs and to adaptively select the image features used
for matching these blobs. Placing a camera on the top reduces occlusion but it is not possible in many situations. Such a camera system is unable to identify people or determine other important statistics (like height or colour distributions) and hence may not be very useful for many applications.

To overcome the existing problems such as changes in the background, occlusion, colour, texture, size etc, an alternative method is required to exactly track multiple targets. This thesis proposes a novel tracking scheme for a single camera sensor.

1.3 OBJECTIVES OF THE RESEARCH WORK

The goal of the research work is to meet the challenges occurring in MTT such as data association and tracking of multiple maneuvering targets and to estimate the target location and track the targets with improved accuracy. WSN is an emerging technology for person detection and tracking. Nowadays the modern surveillance system deploys WSN for detecting and tracking multiple targets using wireless sensor nodes covering both indoor and outdoor environment. The accuracy of tracking with WSN is limited due to the sensor’s detection capability, irregular arrival time of samples, maneuvering nature of targets, missing of events, and random deployment of sensors etc. In order to overcome the above inadequacies there is a need for a novel target tracking algorithm which can model the maneuvering target and accurately predict the future position of the target. Tracking in camera sensor network faces difficulty due to target occlusion, target miss association and background modeling for complex scenarios. In order to provide solutions for the above problems this thesis proposes a scheme Combined Gaussian Hidden Markov Model based Kalman Filter that tracks multiple people in the presence of multiple moving objects.
The main objectives of the research work are as follows.

- To propose a new data association algorithm to enhance the tracking accuracy over multiple moving targets with reduced computation time and memory utilization.

- To propose a new estimation algorithm to track the maneuvering targets using multiple filters and multiple sensors in the WSN environment.

- To propose a new algorithm for tracking multiple people from captured video in complex scenarios.

1.4 MOTIVATION FOR THE PROPOSED RESEARCH

The basic principles of MTT were first recognized by Wax (1955) and the major breakthrough came in 1964 with the publication by Sittler (1964). The papers by Bar-Shalom (1975) and Singer and Stein (1971) began the development of modern MTT techniques that combine correlation and Kalman Filtering theory. Certain applications of target tracking have become relatively more important in the last few years. Air Traffic Control (ATC) experienced huge growth due to the increase of civilian aviation and drives highway vehicle surveillance also interest in intelligent transportation systems by the creation of surveillance systems and increased sophistication. In general, these provide accurate and timely knowledge of the state of one or more moving targets based on a combination of the sensed data and the target history.

Among the existing solutions for data association such as Probabilistic Data Association (PDA), Joint Probabilistic Data Association (JPDA) and Multiple Hypothesis Tracking (MHT) do not perform well in
multiple target environments due to non-availability of track initiation (Blackman 1986). Moreover, the probability of the presence of the target is computed with reference to the previously available targets and also heavy pruning is required. To overcome such problems, this research proposes a new data association algorithm named as Zero-Backscan MHT (ZBMHT) to reduce the computation time, memory utilization and to improve the accuracy.

The existing method of target tracking in WSN uses an acoustic signal and Time Difference of arrival (TDoA) for detecting and localization (Umesh Babu et al 2006). However, the performances of this method may be inferior due to path loss and signal degradation caused by various obstacles present in the environment. This thesis proposes a scheme for detecting and tracking multiple targets using WSN. Each WSN node performs detection and identification of the target. The accuracy of tracking in WSN is limited due to the sensor detection capabilities and the sensor packets may be lost in the wireless medium. The main objective of the proposed multi-sensor and IMM based tracking scheme in a WSN is to overcome the limitations due to localization errors and errors due to missing events caused by failure of nodes.

The existing camera surveillance techniques make assumptions which greatly restrict the generality of the approach in real-world settings such as background modelling techniques. Assuming that the appearance of non-targets can be predicted well from past observations, scenes often include many other dynamic objects, fast changes in lighting, and complex object interactions like shadows and reflections that greatly influence the image (Comaniciu et al 2003). The existing techniques (Krumm et al 2000, Iketani et al 1998, Kang et al 2004) are entirely different from the proposed one and these approaches require the objects in the scene to have enough texture information for dense stereo reconstruction and build background models for static environment. Hence the main objective of the proposed Combined Gaussian Hidden Markov Model based Kalman Filter (CGHMM-KF) scheme
is to detect the target under different illumination condition and investigates a multi person tracking scheme.

The region based stereo technique overcomes with many problems of wide baseline correspondence by matching regions instead of points (Kang et al 2003). This approach requires background modelling and assumes that everyone in the scene is wearing uniquely coloured clothing to perform the region based correspondence. Multi-camera techniques perform correspondence between views and assume that the appearance of a feature in one view will be similar to its appearance in another. This assumption fails for widely separated views where the scene geometry and lighting can result in a lack of commonly observed features and different appearances of the same feature. The Mixture of Gaussian (MoG) estimates the foreground objects and the HMM modeling technique recognizes the desired object from the foreground. The tracking of multiple targets is done by KF with a bounding box, indicating the location of the person within the entire video sequence. The block diagram of proposed methodologies is shown in Figure 1.2.

Figure 1.2 Block diagram of proposed methodologies
This research provides solutions for Multiple Target Tracking in three different platforms. The problems and existing solutions are discussed in Chapter 2.

1.5 **SALIENT FEATURES OF THE PROPOSED RESEARCH WORK**

The proposed Zero-Backscan MHT (ZBMHT) algorithm has got salient features such as, ability to instantly modify each observation in the constructed hypothesis tree, simultaneous data association using the Bayes rule and Gaussian density function. Also it combines data association and position estimation techniques. Hence the proposed algorithm provides better results than the existing algorithms in terms of computation time, accuracy and memory utilization. The results of proposed method is validated with respect to the parameters namely, probability of detection and false target density.

The proposed multi-sensor and IMM based tracking scheme in a WSN uses a PIR (Passive Infra Red) sensor to detect the presence of a target, RFID reader to identify the target and Global Positioning System (GPS) receiver to get the location of sensor node. A person moving in the WSN deployment area is detected by the sensor node which produces an event in WSN and this event is sent as a wireless packet to sink. The proposed scheme uses the velocity model for a linear target and acceleration model for a maneuvering target tracking. The estimation error due to missing events caused by failure of nodes is reduced by the IMM estimator for linear and maneuvering targets. The performance of the proposed scheme is verified for different targets and sensor deployment scenarios and found to give better tracking accuracy than the existing schemes.
The proposed CGHMM-KF scheme uses the MoG to estimate the foreground objects and HMM modeling technique to recognize the desired target from the foreground and KF to track multiple targets with a bounding box indicating the location of the target within the video. The proposed scheme provides better performance even in the presence of occlusions and overcomes target miss association problem.

1.6 THESIS OUTLINE

The rest of the thesis is organized as follows. Chapter 2 summarizes the literature survey of data association techniques and target tracking in radar sensor network, target tracking in WSN and multiple targets tracking for camera surveillance. The target tracking in radar sensor network with the proposed Zero-Backscan Multiple Hypothesis Tracking scheme is described in Chapter 3. The target tracking in WSN with the proposed Interacting Multiple Model scheme with velocity and acceleration model filters is presented in Chapter 4. Chapter 5 presents the proposed Combined Gaussian Hidden Markov Model based Kalman Filter scheme for target tracking in camera sensor network. Chapter 6 concludes the thesis with the research contributions and future work.