Chapter 1

1 Introduction

The modern era of automation demands the involvement of machine to perform critical tasks efficiently and accurately for convenience and enhancement of the quality of life. These demands focus on the research to develop an intelligent system having the ability to perceive, calculate and learn from the experiences. One of the key motivations behind the use of intelligent technologies is the fact that they can deal with human cognitive limitations, i.e., human failure to monitor all information, to resolve complex and conflicting situations, to identify high-revenue opportunities or to prevent high-cost mistakes.

Intelligent systems aim to realize the real world applications targeted to Automated Visual Surveillance, Traffic Monitoring System, Intelligent Vehicle System, Robot Navigation and Animations etc [1].

- Automated Visual Surveillance System: The smart surveillance system works over security-sensitive areas such as banks, departmental stores, parking lots, and borders of the countries. In present systems, surveillance camera outputs are recorded into the video archives. These video data are currently used as forensic tools after the incidence. For real time analysis, security officers need to be alert in the control room for continuously watching the progress of the events. There is a need of alert only when the motion is detected. So it is necessary to build generalized model which can work effectively for real time applications under the normal conditions [3].

- Traffic Monitoring System: Traffic monitoring system identifies the types of the vehicles and monitors the flow of vehicles. Traffic monitoring system needs to estimate the speed and direction of the moving vehicles and pedestrians and also the traffic intensity for the prevention of the road accidents. Traffic monitoring and controlling can also be made more effective with the help of automatic toll collection system if any. Automation of toll collection system eliminates the delay on traffic roads and tolls electronically. Presently used electronics sensors like optical, electromagnetic, radar and sonic etc are very costly. These necessitate the designing of low cost system for traffic monitoring and automatic toll collection centre [1], [6].

- Intelligent Mobile Robot: Intelligent Mobile Robot has ability to determine its own position in its frame of reference and then to plan a path towards some goal location. For planning the path; object localization, speed and direction are required to avoid the collision and navigating safely in the environment [3].

- Multimedia Animation and Video Application: Automatic motion capture algorithm in Camera Tracking is used in broadcasting and cinema for visual effect creation [1].

Currently, very few algorithms exists that can perform motion detection and classification reliably and efficiently with high accuracy and execution speed for the real time applications. This thesis is an attempt to provide an effective and robust solution which is used to identify the object and also to extract the motion parameters from the moving object from image sequences.

1.1 Machine Intelligence System

The Machine Intelligence system extracts the information from image or image sequences stored in digital computer. The image data can take many forms, such as video sequences, views from multiple cameras, or multi-dimensional data from medical scanners etc. Machine Intelligence systems are linked with Artificial Intelligence (AI) which is the key computer technology applied to manage the knowledge and human resources.

Since the mid-1980s, there has been sustained development of the core ideas of artificial intelligence, e.g. representation, planning, reasoning, natural language processing, machine learning, and perception. In addition, various sub-fields have emerged, such as research into autonomous, independent systems (hardware or software), distributed or multi-agent systems, coping with uncertainty, effective computing/models of emotion, motion and manipulations, general intelligence etc. Figure 1.1 shows the traits which have received the most attention [2].

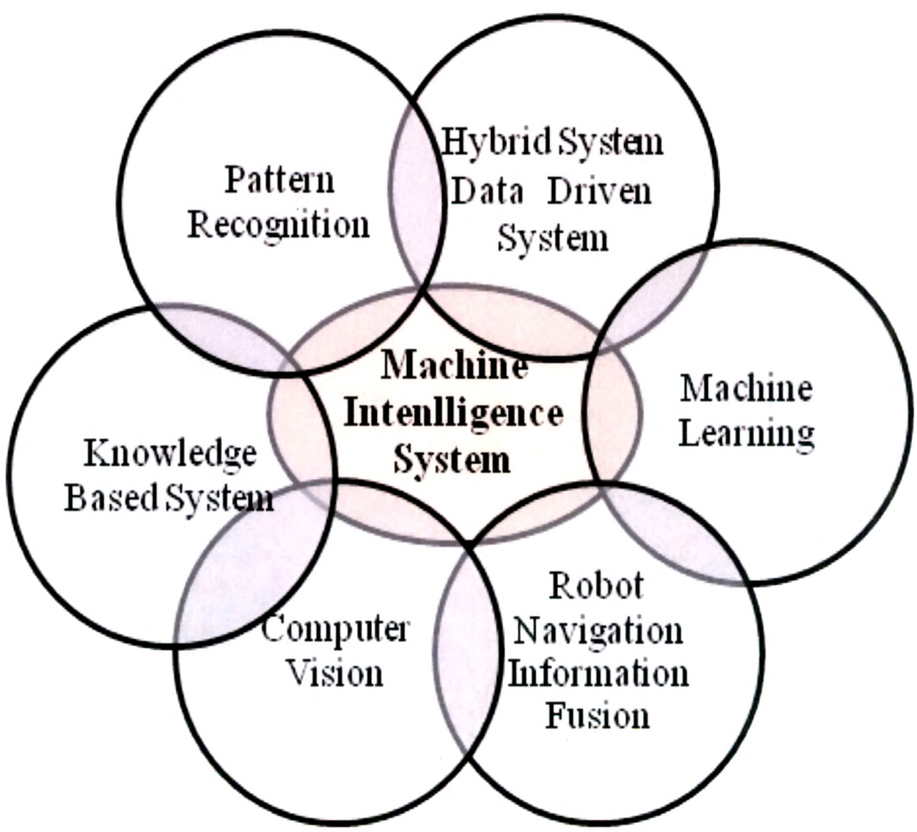

Figure 1.1 : Machine Intelligence System

The fragmentation of AI into specialized sub-fields has produced powerful component methods for standalone algorithms in almost many of the research areas. Many research papers have been published about the different algorithms in different domains, but relatively less attention has been paid to situations involving multiple domains. It is needed to combine .sub-field technologies of AI towards the construction of integrated cognitive systems that mimic broad human-level intelligence.

Challenges for developing Machine Intelligence Software System based on image analysis:

- Inferring three-dimensional structure from two-dimensional images is inherently ambiguous.

- Reflection from the object surface, inhomogeneous illumination and discontinuity of the reflecting surface at the object borders are potential causes for uncertain information due to the spatial changes in the intensity.

- Occlusions are the frequent source of the ambiguity.

Limitations:

Machine Intelligence System consists of two main components: Image capturing and Analysis of captured image. Image capturing can be done easily with 2D CCD image sensors with millions of pixels. Line cameras, logarithmic image sensors, CMOS sensors are also used for high resolution, high dynamic range, and low power consumption. However the cost involvement is more in all these devices. While analyzing of the captured image, there is a problem on two aspects: speed and quality of the processing images. Cameras and other image capturing devices produce large amounts of data. Although processing speed and storage capabilities of computers have been increased tremendously in the last decade, processing high resolution images and videos are still a challenging task. Much more work has been carried out for offline or desktop processing, but very few algorithms have been developed for real time applications. Depending on the applications, Intelligence systems try to extract different aspects of the information contained in an image or a video stream [6]. For example, moving objects are discarded to infer a structural object model from a sequence of images, whereas for the control of mobile robots, analysis may start with motion model of objects.

Two main approaches exist for interpretation of images: bottom-up and top-down. Bottom up approaches are mostly used for recognition and classification systems like character recognition system, biometric recognition system etc. The top-down approach for image analysis start with the object models rather than image. Top-down techniques are used for image registration and for tracking of objects in image sequences [6]. The hypothesis can be generated by predictions which are based on the analysis results from the preceding frames. Recognizing the objects and perceiving actions of the objects are prerequisites for Machine Intelligence System. The Machine Intelligence System involves mainly two tasks: Object Identification and Visual Tracking.

1.1.1 Object Identification

Object Identification or Object classification of moving objects in video streams is the first relevant step of information extraction in many computer vision applications. Object Recognition in machine vision is the task of finding an object in the given image or video sequence. Humans recognize an object with little effort. But for a machine, an image is a projection of 3D structure to 2D. Differing appearance of the same object with variation in the different viewpoints, viewing distance, scaling, translation, rotation, varying illumination, cluttered background, intra-category appearance variations etc. make the task difficult. Object should be recognized even when they are partially obstructed from view. This task is still a challenge in machine vision.

A complete Object recognition system consists of following Modules [4] as shown in the Figure 1.2:

1. Data Acquisition
2. Pre-processing
3. Feature Extraction and Feature Selection
4. Model Selection and Training
5. Performance Evaluation

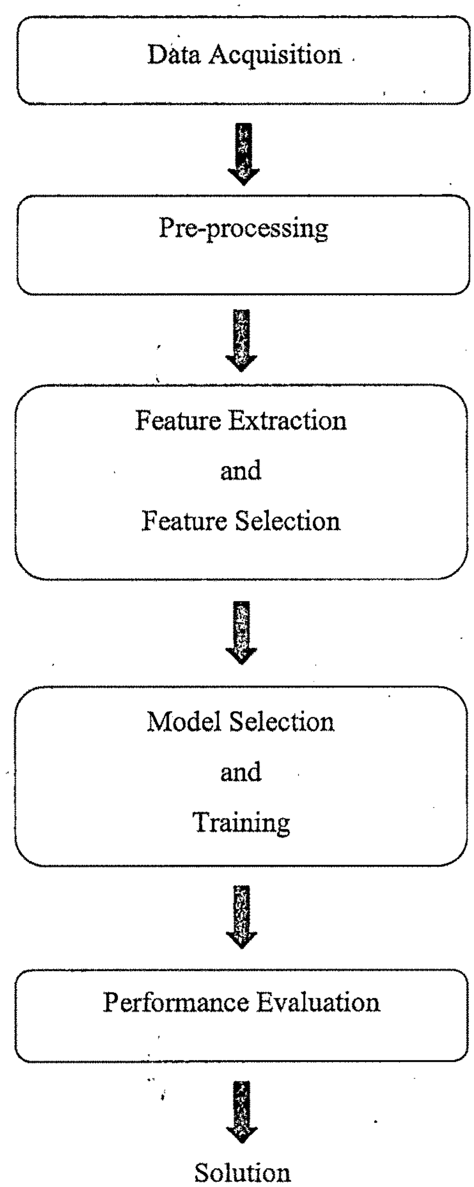

Figure 1.2 : Modules of Object Recognition System

1. Data Acquisition: One of the most important requirements for designing a successful object recognition system is to have adequate and representative training and testing datasets. A sufficient amount of training dataset required to learn a decision boundary as a functional mapping between the feature vectors and the correct class labels. There is no rule that specifies how much data is sufficient. Designer must select the types of sensors or measurement schemes that provide the data such that it should be able to design the classifier effectively [4].

2. Pre-processing: The goal is to condition the acquired data such that noises from various sources are removed as much as possible. Various filtering techniques are used if the user has prior knowledge regarding the spectrum of the noise. Conditioning may include the normalization of the data with respect to the mean and variance of the amplitude of the data normally called feature value. Pre-processing used for deblurring, image enhancement or edge detection depends upon the application domain [4], [6].

3. Feature Extraction and Feature Selection: The goal of this step is to find preferably small number of features that are particularly distinguishing or informative for the classification process and that are invariant to irrelevant transformations of the data. Better discriminating information may reside in the spectral domain or frequency domain. Feature extraction is usually obtained from a mathematical transformation of the data. In the spatial domain, feature descriptors are extracted using colors, geometric primitives like line and circles or textures. In the Frequency domain, feature extractions are performed by applying Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), Fast Fourier Transform (FFT), Discrete Wavelet Transform (DWT) etc. Some of the methods find intensity discontinuity points which are invariant to rotation, translation, scale [1], [4].

4. Model Selection and Training: After acquiring, pre-processing and extracting the most informative features of the training dataset, classifier and training algorithms are selected. Classification can be considered as function approximation problem which can use variety of mathematical tools, such as optimization algorithms. Most common object recognition algorithms use statistical approaches or Neural Network approaches. Statistical pattern recognition uses Bayes Classifier, Naïve Bayes Classifier, K-Nearest Neighbour (KNN) classifier etc. Neural Network approaches use Multi-layer Perceptron (MLP), Radial Basis Function (RBF), Support Vector Machine (SVM), Self Organizing Map (SOM) [1],[3],[4] etc.

- **Statistical Approach:**

Bayes Classifier: Bayes Classifier uses data points those are assumed to be drawn from a probability distribution, where each pattern has a certain probability of belonging to a class, determined by its class conditioned probability distribution. A given d-dimensional $x = (x_1, \ldots\ldots\ldots, x_d)$ needs to be assigned to one of the c classes $w_1, \ldots\ldots, w_c$. The feature vector x that belongs to class w_j is considered as an observation drawn randomly from a probability distribution conditioned on class w_j, $P(x|w_j)$. This distribution is called the likelihood probability. The Bayes theorem takes the prior likelihood of each class into consideration. Disadvantages of Bayes classifier is the difficulties in estimating the likelihood probabilities, particularly for high dimensional data. To overcome the problem Naïve Bayes Classifier is used. The main advantages of Naïve Bayes classifier is that it only requires univariate density to be computed, which are much easier to compute than the multivariate densities. The main disadvantage is that dependencies among the class cannot be modeled by Naïve Bayesian classifier [4].

K-Nearest Neighbour (KNN) Classifier: KNN classifier can be used as a nonparametric density estimation algorithm, and is most commonly used as a classification tool [2],[3]. The top k matches are then used to obtain a classification. Typically some sort of majority vote is used to determine the label assigned to the query object. This classification requires no training, although the value of k needs to be determined somehow. If it is too small, the classifier becomes sensitive to noise and if it is too large the computational time increases and becomes biased towards the classes with the larger number of members. A commonly used version of the k-NN is the Nearest Neighbour (NN) classifier where k is equal to one. The disadvantage of this scheme is that the quality of results depends on the training set. These algorithms have no computational cost of training but more computational cost during the testing of features.

- **Neural Network Approach:**

Multi-Layer Perceptron (MLP): The Multi-Layer Perceptron is one of the most popular classification techniques. Jones [5] showed that a MLP with just two layers using a sigmoid activation function can approximate any function to an arbitrary error. The main disadvantages that, MLP suffers from the computational cost as it increases at an exponential rate as number of dimensions increases. The MLP works by propagating an input pattern through a number of layers with varying numbers of nodes. Each node has a weight assigned to it and has some activation function assigned to it. The activation function of a MLP determines what sort of functions it can represent. If the activation function is linear, then the network is no more powerful than a single layer network. A sigmoid activation function is used which performs a non-linear mapping allowing much more powerful networks to be built. Typically MLPs are trained using the back-propagation algorithm. This algorithm takes into account the individual weightings within the network and can choose the best change to get the weights per iteration.MLP better handles the classification type problems [1],[3],[4].

Radial Basis Function Network (RBF): The Radial Basis Function network classifier [1], [4] is a technique that relies upon casting the classification problem into a much higher dimensional space than the input vector in order to increase the likelihood of creating a linearly separable problem. An RBF network consists of a number of input nodes, a hidden layer and an output layer. The hidden layer typically uses a Gaussian activation functions to perform a non-linear transform. This layer will also contain many more nodes than the input to cast into the higher dimensions. The output layer consists of a number of linear activation functions. The RBF uses a randomly initialized set of weights as it gives the different result each time it is trained. The training process should be able to reduce the effects of initial conditions if enough numbers of iterations are performed. Training a RBF network is faster than training a MLP network.. Training is split into two fast stages. The first stage uses an unsupervised method to determine the parameters of the basis functions. The final stage solves a linear problem, mapping the hidden layer to outputs. RBF performs well on function approximation problems.

Support Vector Machine (SVM): The Support Vector Machine (SVM) is a popular and powerful classification technique [4]. It is a kernel based technique which does not suffer from the curse of dimensionality like those other classification techniques. Due to the kernel nature of the support vector machine, different types of network can be built, such as polynomial learning machines, radial-basis function networks and two-layer Perceptron. An attribute particular to SVMs is that they can provide good generalization performance even though they do not incorporate problem-domain knowledge. The SVM is traditionally a Binary Classifier. This method significantly increases computation expense as the number of classes increase. The one-versus-many method trains one classifier for each class. Training examples are labeled with '1' if they are of the target class or '-1' otherwise. The label associated with the classifier returning the largest positive distance from the decision boundary is selected to make the classification. The one-versus-one method trains a classifier on every possible pairing of classes. The Final classification is made by a voting process where each classifier can vote for one of the two classes. The winning class is then used to make the classification. The third method, DAG (Directed Acyclic Graph)-SVM, is similar to the one-versus-one method except a directed acyclic graph is constructed such that each classifier is a node in the tree and the leaves are the resulting classes. This reduces the number of classifications required whilst keeping a similar level of performance.

Self Organizing Map (SOM): Self Organizing Map (SOM) transforms an input pattern of arbitrary dimension into a one- or two-dimensional discrete map and performs this transformation in a topologically ordered fashion[1], [4]. The SOM algorithm is simple to implement, however very difficult to analyze mathematically. There are three essential processes called competition, co-operation, and synaptic adaptation. During competition, neurons in the network compute their respective clause of a Discriminant function. This Discriminant function provides the basis for competition among the neurons. The neuron with the largest value of discriminant function is declared the winner. During co-operation the winning neuron determines the spatial location of a topological neighbourhood of excited neurons, providing the basis for co-operation. During synaptic adaptation the excited neurons increase their individual values of the discriminant function in relation to the input pattern through suitable adjustments to their synaptic weights. Adjustments are made so that a similar pattern returns an enhanced discriminant value.

5. Performance Evaluation: To get a good estimate of the generalization ability of a classifier, different methods like split-sample, cross-validation, Boot strapping are used. Performance evaluation can be done by splitting the entire dataset into two parts, training dataset and testing dataset where the training dataset is used for actual training and testing dataset is used for testing the true field performance of the algorithm. Since adequate and representative training dataset is highly important for successful training, no rules are fixed about the size of training and testing dataset [2], [3].

Split-Sample: A commonly used method for classifier training and validation is Split-Sample Validation [6]. The data set is split into a training and validation set (often a 50% - 50% or 75% - 25% split). The classifier is trained using the training set, and validation is performed on the validation set. The greater the number of samples, the closer to the true error the estimate will be. This means that for small numbers of samples the estimate is likely to be inaccurate.

Cross-validation: Cross-validation has been used to select classifier training parameters [7] and to estimate the generalized performance of a particular set of parameters. In this situation, training and testing data sets are produced. The training data set is further partitioned into an estimation and validation set. For each set of parameters, a classifier is trained using the estimation set and its performance evaluated using the validation set. The best performing set of parameters is then selected and its performance is evaluated using the test data set to avoid problems of over fitting the training data.

Alternatively cross-validation can be used to provide a better generalization estimate than split-sample when small data sets are available. One approach that has been accepted to provide the reasonably good estimate of the true generalized performance uses k-fold cross-validation. In k-fold cross validation, the entire available training dataset is split into k > 2 portions, creating k blocks of data. Among these k blocks, k-1 blocks are used for training. This procedure is repeated k times using different blocks for testing each case. The average performance is declared as the estimate of the generalized performance of the algorithm. It is computationally expensive for larger data sets or high values of k.

Boot Strapping: Boot strapping is similar to cross-validation except that it uses sub-samples of the data set instead of sub-sets. A sub-sample is random sampling with replacement of the original data set allowing sub-samples to be of nearly any size as required. This is useful when data sets are unbalanced or too small [2].

1.1.2 Visual Tracking

In the Visual Tracking, the motion of an object and interpretation of the object behavior are performed from image sequences or consecutive video frames. Object motion analysis has attracted a great interest due to its promising applications in the real world. Motion Parameters like location, directions and speeds are derived for the learning of Machine Intelligence System. Visual tracking task becomes complicated due to static occlusion, dynamic occlusion, varying size and shape of objects in video sequences. Also object tracking algorithm must be capable of handling trajectory part like static occlusion, dynamic occlusion and tracking complications like splitting and merging as well as object appearance and disappearance successfully. For video tracking, the task is difficult when objects are moving fast relative to the frame rate. Complexity of the problem increases when tracked object changes shape and orientation over a time. For these situations video tracking systems usually employ a motion model which describes how the image of the target might change for different possible motions of the object.

Typical motion tracking system is described as shown in the Figure 1.3. Motion tracking algorithms mainly involve Motion Segmentation, and Tracking [3].

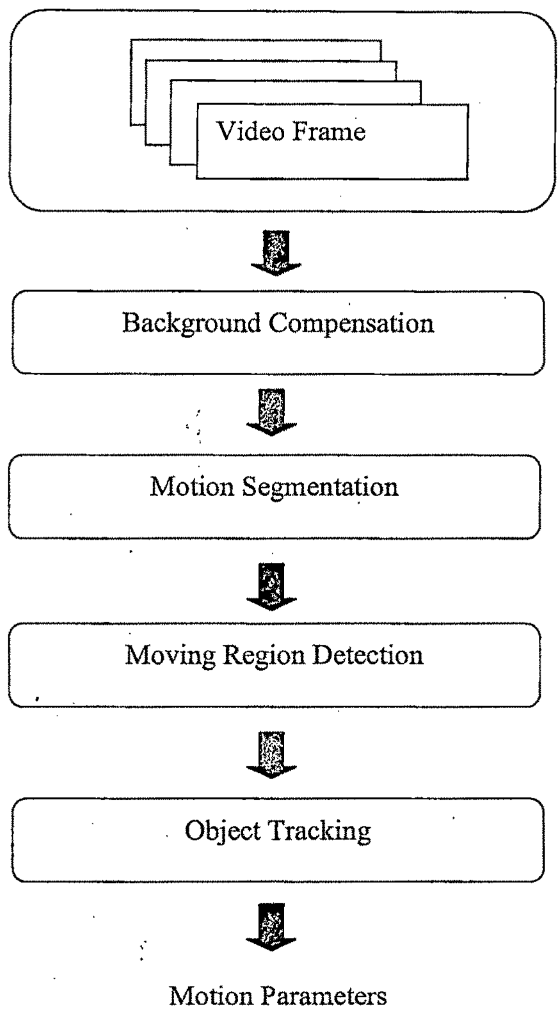

Figure 1.3 : Modules of Visual Tracking System

Motion segmentation aims at decomposing an image in objects and background to find the motion of the object. The information like textures or statistical descriptors, edges, colors can be extracted from a single image which used for object segmentation. Typical segmentation techniques such as region growing, splitting and merging, watershed, and histogram based algorithms, active contours, graph partitioning, and level sets [8], are some of the most widely used techniques.

Motion segmentation can be done in many ways:

Background Subtraction is particularly popular method with a relatively static background. It is extremely sensitive to changes of dynamic scenes due to lighting. To avoid the problem, a procedure is used to update the background model. Each image in current image can be classified into foreground and background by comparing the statistics of current background model. This approach is popular due to its robustness to noise, shadow, change of lighting conditions etc. Intensity thresholding, Gaussian mixture models are used for background model [3], [34].

An approach of temporal differencing makes use of pixel-wise difference between two or three consecutive frames in an image sequence to extract the moving regions. Temporal differencing is very adaptive to dynamic environments and having poor result of extracting features. An improved version uses three frames differencing instead of two frame differencing [8], [9].

Optical flow is generally used to describe coherent motion of points or features between image frames. Motion segmentation based on optical flow uses characteristics of flow vectors of moving objects over time to detect change of regions in an image sequence. For Most flow computation, methods are computationally complex and very sensitive to noise, and cannot be applied to video streams in real-time without specialized hardware [9].

Tracking can be employed using two approaches: Shape based tracking in which image blob area, blob bounding box are considered for tracking. Motion based tracking considers periodicity property of the object motion or time frequency analysis of the object [1], [17].

Tracking over time typically involves matching objects in consecutive frames using features such as points, lines or blobs. That is, tracking may be considered to be equivalent to establishing coherent relations of image features between frames with respect to position, velocity, shape, texture, colour, etc [3]. Blob Tracking involves detection of blobs of the object interior using thresholding technique or block matching techniques. Model based Tracking uses the geometric structure of an object and can be represented as stick figure, 2D contour or volumetric model. For small number of objects, model based approaches are efficient. Since it depends upon the geometric model it is not able to handle the occlusion as geometric model is 2D model. Feature based Tracking method uses sub-features such as distinguishable points or lines on the object to realize the tracking task. Its benefit is that even in the presence of partial occlusion, some of the sub-features of the tracked objects remain visible. Feature-based tracking includes feature extraction and feature matching. Low-level features such as points are easier to extract. It is relatively more difficult to track higher-level features such as lines and blobs. Active Contour Based Tracking uses active contour models, or snakes, aims at directly extracting the shape of the subjects. The idea is to have a representation of the bounding contour of the object and keep dynamically updating it over time. It reduces the computational complexity compared to the region based tracking. Initializations of contours for moving objects are quite difficult, especially for complex articulated objects. Region based tracking is used to identify a connected region associated with each moving object in an image, and then track it over time using a cross-correlation measure. Region-based tracking approach has been widely used today. Video tracking has been difficult in congested situation.

1.2 Overview of the Proposed Work

This thesis is an attempt to develop effective solutions for object tracking with recognition. It has been observed that the existing methods offer scope for improvement. The objective is to propose a new and efficient approach suitable for Machine Intelligence System which identifies the object and extract the motion parameters by analysis of visual tracking of the object. Development of the software model is targeted to the applications such as automated Visual surveillance systems, Traffic monitoring system; self guided vehicles, automatic guided machines and different robotics application under the normal conditions. The thesis presents the software model with the following components:

- Perform the motion segmentation.
- Tracking of the moving objects by locating them in each frame of the sequence.
- Classification of the object.
- Extraction of motion parameters such as location, direction and speed.

Research is conducted in each area of motion segmentation, motion tracking and object classification. A suitable approach for each of these components is chosen by analyzing the strengths and weaknesses of the reviewed techniques. Each process has been implemented and tested with real image sequences, and combined to produce an efficient machine intelligence system which can automatically detect, classify and track the object.

1.2.1 Contribution

To develop the software model for object identification and estimation of motion parameters, two different tasks are considered: Object Classification and Visual Tracking. After an exhaustive comparative study of available alternatives for each method, several enhancements to achieve efficient results for both the tasks have been proposed. Basically these include choice of best approaches for the task and proposed modifications to these approaches for improving their results.

Object Classification:

Context understanding is the key element when developing an information retrieval model for machine intelligence application. An efficient model is required that encodes and classifies video objects such as humans, vehicles, buildings, etc. Thus, it is required to design a generalized object classifier to identify or classify the objects of interest based on the application.

For designing the general Classifier system, a novel feature based object classification using Discrete Contourlet Transform and Principal Component Analysis (PCA) has been proposed. Feature extraction has been carried out in the frequency domain. For better result compared to the conventional discrete Contourlet Transform, pre processing and filtering stages are applied which enhance the edge details of the images. Feature extractions are performed with the preprocessed images that give more efficient result than the discrete Contourlet transform method. For efficient edge point feature extraction, Unsharp Filter is used before feature extraction. Unsharp filter amplifies the high frequency components which enhances the edges of an image.

Feature extraction coefficients are extracted by applying Curvelet transform that overcomes the problem of representing an image with smooth contours in different directions by providing two additional properties: directionality and anisotropy [28] as compared to the Discrete Wavelet Transform (DWT). The Curvelet Transform was developed initially [28] in the continuous domain via Multiscale filtering followed by a block based Ridgelet transforms applied to the subband images. Since it is a block based transform, either the approximated images have blocking effects or overlapping windows are required for calculations that increase the redundancy. Also the use of the Ridgelet transform, which is defined on a polar coordinate, makes the implementation of the Curvelet transform for discrete images on a rectangular coordinate to be very challenging. The second generation Curvelet transform [71] was defined directly via frequency partitioning without using the Ridgelet transform. Both Curvelet transforms require a rotation operation and correspond to a 2D frequency partition based on polar coordinate. This makes the Curvelet construction simple in continuous domain but caused the implementation for discrete images sampled on a rectangular grid to be very challenging [75]. This makes algorithm implementation difficult. This fact leads the development of a directional multiresolution transform like Curvelet, but directly in the discrete domain. Contourlet, as proposed by Do and

Vetterli [23], form a discrete filter bank structure that can deal effectively with piecewise smooth images with smooth contours. This discrete transform can be seen as a discrete form of a Curvelet Transform. In the discrete Contourlet transform, the image is decomposed by a double filter bank structure where the first filter bank captures the edges and second filter bank link the edge points into the contour segments.

Eigenvalue (Principal Component Analysis) of feature matrix has been calculated from the feature matrix that helps for dimensionality reduction for feature matching which increases the execution speed of algorithm. The results with discrete Contourlet with PCA are compared with Discrete Curvelet Transform with PCA. For feature matching, Euclidian Distance Measure and Neural network classifier is used to match the test feature vector with the trained feature vector and compared for analysis.

Visual Tracking:

An efficient method is to be developed to track all the moving objects with high accuracy. The method should be adaptable using ordinary camera for designing the cost effective application system. Further the motion parameters like direction; speed etc. should be extracted from data obtained. Two different algorithms proposed are: Single object visual tracking and multiple object visual tracking.

(1) Single object visual tracking: User selected object has been tracked in the video sequences. 3D Colour histogram and Euclidean distance measures are used for object tracking. Novel Block matching algorithm has been proposed to find the location of the object being tracked in the video sequence frames. Object without motion and object motions out of boundary conditions are also included in the algorithm.

(2) Multiple object visual tracking: Multiple object tracking has been performed using the statistics from data obtained with Blob analysis. Blob segmentation has been carried out by background subtraction and Thresholding. As Blob analysis includes domain independent information, for establishing temporal relationship between the block, colour segmentation is used. Template matching is implemented

using 3D histogram and Hu's seven Invariant moments to track the object. Different termination and decision conditions are included for making algorithm fast and efficient.

1.2.2 GUI for the Proposed Method

GUI for the proposed method has been designed with user friendly features. Each task is designed and implemented in a such a way, that it can be used for many applications like Object Recognition, Finger Print recognition and similar other applications also. Different Parameters for different tasks can be selected individually. The Object Identification task can be used for other object classifier applications also. In the GUI of the Object Classification task, user can obtain the training images and testing images by selecting the folder which contains the images. For training the dataset, user has the options for choosing the Curvelet or Contourlet transform approaches. For visual tracking, the Contourlet transform is implemented, which is faster than the discrete Curvelet transform. For traffic monitoring application, vehicle classifier has been designed using three class structures to improve the efficiency.

For tracking task, user can select the single object tracking method or multiple object tracking. For single object tracking, user needs to select the object of interest using mouse. The Tracking of the object is visualized using rectangular bounding boundary. The multiple object tracking algorithms tracks the moving objects with boundary. It also displays the name of object, object speed in terms of pixels, and direction of the object. Camera calibration has been done for actual speed measurement of vehicle and implemented.

1.3 Layout of the Thesis

The content of the thesis is summarized as:

Chapter 1: This chapter describes the brief history and overview of the problems. It also introduces the objective of the research work and scope of the improvement in the existing methods.

Chapter 2: Various approaches and methods are discussed related to the research problem in the object classification and visual tracking tasks for machine intelligence application. It also covers the suitability conditions, merits and demerits of the existing methods in Literature Survey.

Chapter 3: Describes the Proposed Methods with algorithms in each task: Object Classification and Visual Tracking. Two different approaches have been proposed for Visual Tracking of moving objects: Single object tracking and multiple object tracking.

Chapter 4: Discusses the detailed results of the proposed method and its comparison with the results of the conventional method.

Chapter 5: Concludes with the remarks regarding proposed solutions and their applicability under various situations. Future work possible in the area is also suggested.