Chapter 1

INTRODUCTION

Tracking can be defined as the complexity of estimate the trajectory of an object in the image plane as it moves around a scene. Face plays an important role in tracking. As faces is something that people can usually do without any effort and much conscious mind? But it has been remained a difficult problem in the area of computer vision. Face detection and tracking system in simple words can be described as a computer application for automatically detect and track a person from a video source.

1.1 MOTIVATION

In order to behave intelligently a machine should be attentive of its surroundings. A human receive the great majority of information about their surroundings through Vision and at most of the human brain is dedicated to vision. Vision is also a key Component for building artificial systems that can perceive and understand their environment. Computer vision is likely to change society in many ways; for example, it will improve the safety and security of people and it will make human-computer interaction more natural. With computer vision it is possible to provide machines with an ability to understand their surroundings control the quality of products in industrial processes, help diagnose diseases in medicine, recognize humans and their actions, and search for information from databases using image or video content.

Target tracking in a cluttered environment remains one of the challenging problems of video surveillance. Visual detection, classification and tracking are of the utmost importance in several applications. Detection and Tracking of moving object in computer vision is key concern now days. The task of target tracking provides input to high-level processing such as motion-based recognition ,Automated Video surveillance, Automatic annotation and retrieval of the videos in multimedia databases, Gesture recognition, eye gaze tracking for data input to computers, Traffic monitoring, Vehicle navigation. Intelligent and automated security surveillance systems have become an active research area in recent time due to an increasing demand for such systems in public areas such as airports, underground stations and mass events.

Many applications have need of tracking of multifarious 3D objects. These include visual serving of robotic arms on specific target objects, Augmented Reality systems that require real-time registration of the object to be augmented, and head tracking systems that sophisticated interfaces can use. Computer Vision offers solutions that are cheap, practical and non-invasive.

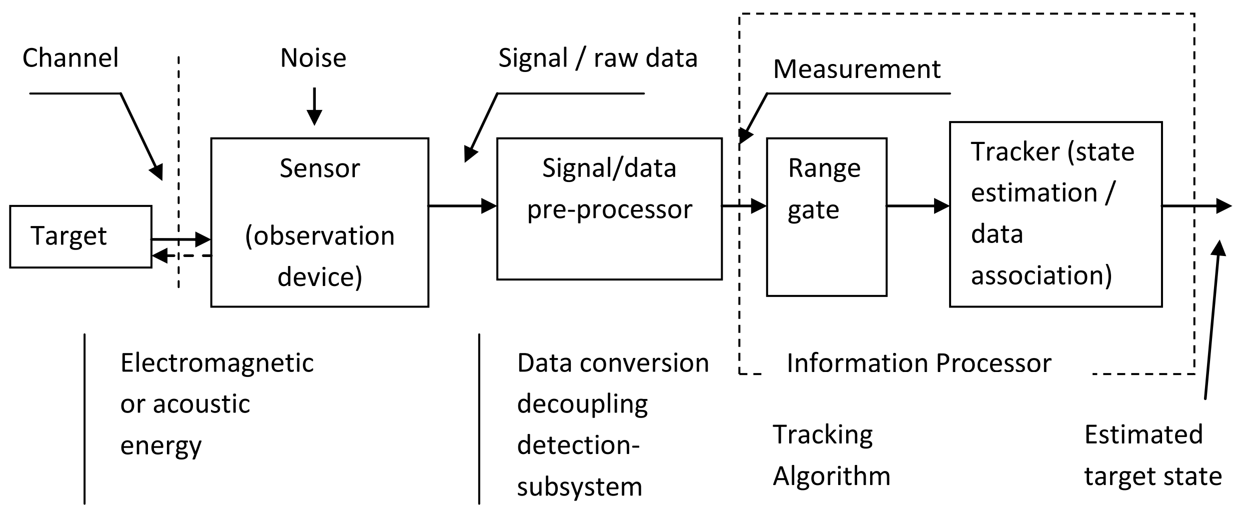


Figure 1.1 Typical object tracking system

The most popular object detection and tracking algorithms include distance transform, deformable template, probabilistic Data Association methods, K-means Algorithm, Active Contour and Appearance Models, AdaBoost based detection, particle swarm optimization techniques, ICA model, motion segmentation, Kernel and Active Contour, Partial Filter, Kalman Filter and Features, mean shift and Camshift etc.

1.2 <u>Various approaches for object detection and tracking[2]</u>

1.2.1 Object detection

The prime goal for further analysis of a video is object detection. It becomes essential due to present and future aspects of it for e.g. for video surveillance, biomedical surveillance, to detect frauds etc. object detection is the basic step. A familiar approach for object detection is to use information through first frame if background is static, otherwise make use of the temporal information computed from a sequence of frames in dynamic background. This temporal information is usually in the form of frame differencing, which things to see changing background in consecutive frames.

Some of the common object detection methods are as follows:

1.2.1.1 **Point Detectors** are used to trace the points in images which have a significant feature in their respective localities. Interest points are being used in the context of motion, stereo, and tracking problems. A desirable quality of an interest point is its invariance to changes in illumination and camera viewpoint.

1.2.1.2 **Background subtraction** Background subtraction method can be used by building a background model by illustration of the scene in the upcoming frame. Any considerable variation from the background model indicates the salient object, which is recorded for further processing.

1.2.1.3 **Segmentation** The goal of segmentation algorithm is to divide the entire image into no. of comparable regions. All segmentation algorithms follows two situations, first what must be the criteria for good partitioning and second, the way to achieve superior partitioning.

1.2.1.4 **Supervised learning** Object detection can be achieved by learning different object views automatically from a set of training database by means of a supervised learning mechanism.

1.2.1.5 **Texture** object can be detected by measuring texture features of a video frame respectively. It helps in illumination variation, occlusion etc.

1.2.2 Object tracking

1.2.2.1 **Point tracking** in point tracking algorithms objects are detected in successive frames and represented by points and the points are associated on the basis of previous state and object motion. This approach requires an external mechanism to detect the objects in every frame.

1.2.2.2 **Kernel tracking** Kernel refers to the shape and appearance of an object.it is performed by computing the motion of object which is shown by basic object region, from first frame to next. Objects are tracked by computing the motion of the kernel in consecutive frames. For example, the kernel can be a rectangular template or an elliptical shape with an associated histogram.

1.2.2.3 **Silhouette tracking** is performed by approximating the object region in every frame. This form of tracking makes the use of the information encoded inside the object region. This information can be in the form of appearance density and shape models which are usually in the form of edge maps. Silhouette tracking based methods provide an accurate shape description for objects having complex shapes.

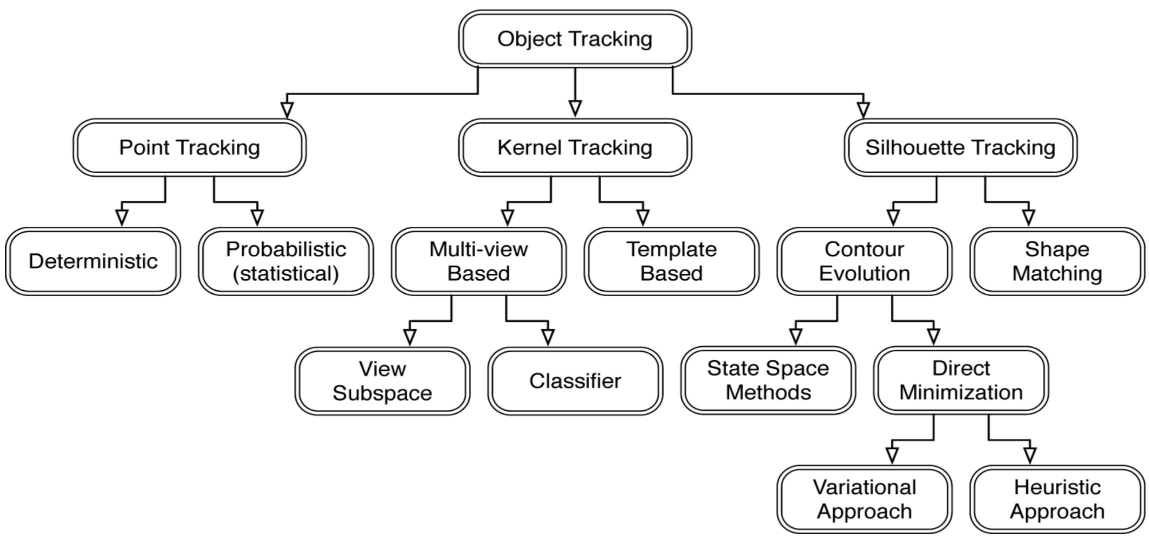


Figure 1.2 Typical object tracking methods

1.3 <u>Applications of object detection and tracking</u>

1.3.1 Air space monitoring

An important application of tracking is aircraft tracking using radar. It is used for air traffic control. Military surveillance also makes use of radar. That associates consecutive radar observations of the same target. A classical rotating air surveillance radar system detects target echoes. A radar tracker is a component of a radar system, and control system, against a background of noise. The role of the radar tracker is to monitor consecutive updates from the radar system and to determine those sequences of plots belonging to the same target, whilst rejecting any plots believed to be false alarms.

A radar track will typically contain the following information

- Position (in two or three dimensions)
- Heading
- Speed
- Unique track number
- Call sign information
- Track reliability or uncertainty information

1.3.2 Video surveillance

The use of tracking is increasing continuously in automated video surveillance i.e. monitoring a scene to detect suspicious activities or unlikely events. In all sectors of society, Video surveillance is used as a means to increase public safety and security and to determine criminal acts. It is now instrumental in implementing security at airports, buildings, banks, department stores, casinos, railway stations, highways, streets, stadiums, and crowd gathering places. Now a day's CCTV cameras are used to detect and track a person or vehicle in video images, and furthermore to infer their behavior, such as unusual, loitering or even criminal behavior, one must solve non-trivial problems.

1.3.3 Weather monitoring

In order to provide weather forecasts, weather bureaus use several techniques. One technique is to track weather balloons, which provide information on high-altitude wind velocities, pressure, humidity and temperature. Weather bureaus release 50 to 70 balloons each day with the release time spaced throughout the day. In extreme weather conditions, the number of releases increases because more information is needed for accurate weather prediction. In order to get weather-related parameters at different levels of atmosphere, each weather balloon needs to be tracked. A simple approach is to track the balloon using ground-based radars.

1.3.4 Motion-based recognition

Motion based recognition based on human identification centered on gait. Gait is a behavioral biometric source that can be acquired at a distance. Gait recognition is the term typically used in the computer community to refer to the automatic extraction of visual cues that characterize the motion of a walking person in video and is used for identification purposes in surveillance systems. If in surveillance applications, it is difficult to get face or iris information at the resolution required for recognition. Studies in psychophysics indicate that humans have the capability of recognizing people from even impoverished displays of gait, indicating the presence of identity information in gait.

1.3.5 Video indexing

It is automatic annotation and retrieval of the videos in multimedia databases.as we sort a video by name, date, time etc. Monte carlo, Bayesian network are used for video indexing.

1.3.6 Robot vision

In robot navigation, the steering system needs to identify different obstacles in the path to avoid collision. If the obstacles themselves are other moving objects then it calls for a real-time visual object tracking system.

1.3.7 human-computer interaction

It is gesture recognition, eye gaze tracking for data input to computers, etc. for security of data protection is required.

1.3.8 Animation

Visual object tracking algorithm is used for animation.

1.3.9 Vehicle navigation

It is, video-based path planning and obstacle avoidance capabilities.so that traffic can be controlled and chances for misfortunes decrease.

1.4 <u>Challenges in object tracking</u>

There are many variations in the object images that are considered for object tracking. These are due to changes in illumination conditions, viewing direction, occlusion and noise etc.

1.4.1 Illumination variations

The illumination problem is illustrated in **Figure 1.3**, where the same face appears differently due to the change in lighting. More specifically, the changes induced by illumination could be larger than the differences between individuals, causing systems based on comparing images to misclassify the identity of the input image.

Figure 1.3: Variation in facial images due to illumination

In the above figure, we see that due to different lighting conditions the same face appears different. This creates a problem for face tracking systems.

1.4.2 **Projection from a 3D world on a 2D image**

Information is lost when we project a 3D space on a 2D space .The essence of an image is a projection from a 3D scene onto a 2D plane, during which process the depth is lost. The 3D point corresponding to a specific image point is constrained to be on the line of sight. From a single image, it is impossible to determine which point on this line corresponds to the image point.

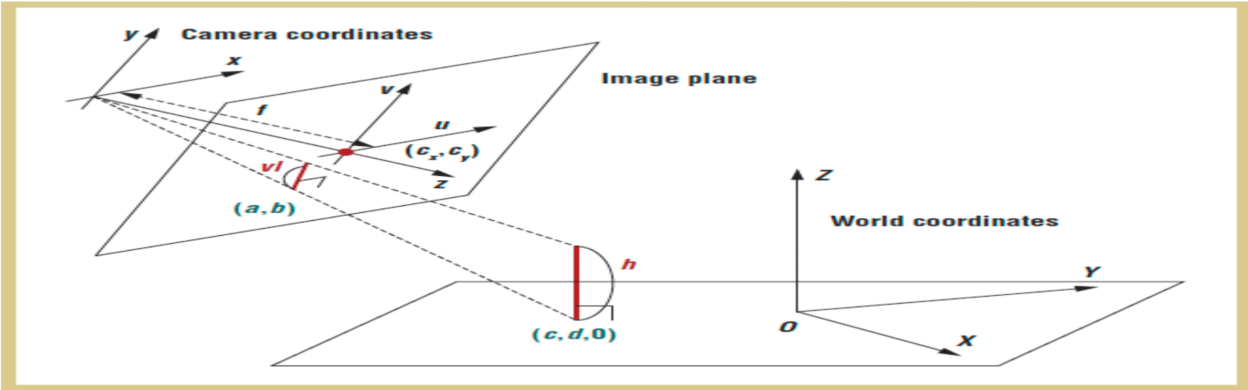


Figure 1.4: 3D to 2D projection

1.4.3 Partial and full occlusion

If there is obstructed line of sight for the object we want to track then it is called occlusion. It can be partial occlusion or full occlusion. It is very difficult to track an object which is occluded.

Figure 1.4: partial occlusion

1.4.4 Complex object motion

It is difficult to track an object whose motion is completely unpredictable. For e.g. to track the motion of balloon for weather forecasting in hilly areas. The principle of the tracking task is to control the camera pan and tilt such that the estimated center of the object appears at the center of the image.

1.4.5 Non-rigid or articulated nature of objects

Human tracking is used as an example of a non-rigid body tracking. The objective is to track the targeted person in between humans with similar outlines in a series of image frames. The targeted person is identified in the initial frame by the user. This framework consists of three stages: generation of panoramic images for a wider range, detection stage, and tracking stage.

1.4.6 Complex object shapes

Objects which exhibit non-linear deformations give rise to models which are not compact and not specific in the process of capturing the range of valid shapes; invalid shapes also become incorporated into the model. This effect is particularly pronounced when building models from automatically-gathered training data. Also, in tracking, smooth movement and deformation is generally assumed, but is not always the case the apparent shape of an object can change discontinuously over time due to, for example, rotations in 3D.

1.4.7 Real-time processing requirements

Real time processing require large amount of memory. And an algorithm which can track the object fast.

To deal with all these challenges an algorithm is required which should be Robust, adaptive and have high processing speed. Means that even under complicated conditions, the tracking algorithms should be able to follow the interested object. Additional to various changes of the environment that an object is located in, the object itself also undergoes changes. This requires a steady adaptation mechanism of the tracking system to the actual object appearance. A system that needs to deal with live video streams must have high processing speed. Thus, a fast and optimized implementation as well as the selection of high performance algorithms is required. The processing speed depends on the speed of the observed object.

1.5 Scope of the work

In this work we have focused on object detection and its tracking. Although there are many other ways to deal with identification but face is most important as human being identify a person by his/her face at his first glance. So it becomes important to make artificially intelligent computer to detecting a face.

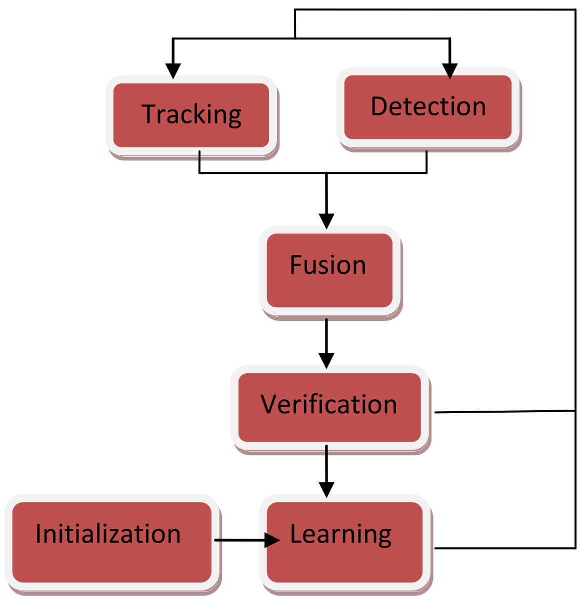


Figure 1.5: Tracking Flowchart

In the proposed paper CamShift algorithm is used, it has been recently proven to be a powerful and reliable algorithm for tracking moving objects. It is a robust non-parametric kernel based tracker. CamShift algorithm which uses a one dimensional histogram to track an object with known hue in color images sequences. Camshift algorithm cleverly exploits mean shift by changing the size of the window when it came to convergence. The Camshift coupled is adapted to color image sequences, and is used to track in real time system. And a textured area in an image can be characterized by a nonuniform or varying spatial distribution of intensity or color. In general, when a target and its corresponding background have similar colors, they have different textures. In this paper we refer to Local binary pattern (LBP), proposed by Ojala [45] [46] as local texture operator. Its computational complexities are less, and it is easy to satisfy real-time in moving objects tracking. It is a non-parametric kernel. It is invariant to monotonic grey- scale transformations and rotation invariant which is very important for texture analysis. LBP operator labels the pixels of an image with decimal numbers, which are called *LBPs* or *LBP codes* that encode the local structure around each pixel. Each pixel is compared with its eight neighbors in a 3×3 neighborhood by subtracting the center pixel value; the resulting strictly negative values are encoded with 0, and the others with 1. For each given pixel, a binary number is obtained by concatenating all these binary values in a clockwise direction, which starts from the one of its top-left neighbor. The corresponding decimal value of the generated binary number is then used for labeling the given pixel. The derived binary numbers are referred to be the LBPs or LBP codes. The LBP method can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis.

1.6 Organization of thesis

The remainder part of this thesis is organized in the following chapters:

Chapter 2: Literature survey

In this section, the critical points of current knowledge on object detection and tracking have been reviewed. Substantive findings as well as theoretical and methodological contributions in the field of object detection and tracking have been included. This section summarizes the work done in the field of object detection and tracking using various techniques in recent times.

Chapter 3: Proposed face detection and tracking system

In this section, image preprocessing has been discussed. Also, detailed description about the methods of feature extraction used has been provided. These methods include local binary pattern. Along with this, the methods of tracking also have been discussed. This includes camshift for multi object tracking.

Chapter 4: Experiments and results

In this section, initially the dataset used for the work has been mentioned. Thereafter, all the experiments conducted along with the analysis of the results have been explained in detailed. The section is divided into two parts, face detection and tracking. Figures and tables have been provided wherever necessary.

Chapter 5: Conclusion and Future Scope

In this section the conclusion of the thesis work and the future scope of the work are presented.

References: This section gives the reference details of the thesis.

Appendix : Abbreviations

Chapter 2

LITERATURE SURVEY

Tracking becomes very renowned from last few years due to its vast scope and application. To track objects, different methods are used earlier. Many global approaches that use more information have been explored to overcome errors of detections.

HUEY-MIN SUN, SHU-MIN CHIANG, in 1992[3] Tracks Multitarget in Cluttered Environment. The algorithm is performed by taking a sliding widow. A set of measurements is received and a new set of data association hypotheses for all measurements lying in validation gates. Target measurement updated with the adaptive state estimator. Instead of making a large problem, the entire set of targets and measurements is divided into several clusters. Meanwhile, the covariance-matching technique is adopted such that the accuracy of the adaptive state estimator is improved. But some limitation exists such as Matching remains dependent on a reasonable contour segmentation, shape variation become larger, so it is not the appropriate method.

Martin Armstrong Andrew Zisserman in 1995[45] describe a robust object tracker .In which object is described by a set of related geometric primitives (here after called primitives). At a low level the primitives are associated with a set of high contrast edges, and are used to reject outlying edges measured in the image. At a high level, the primitives are associated with the object pose, and are used to reject outlying model-image associations robust to a number of ambient conditions which often severely degrade performance, for example partial occlusion. The robustness is achieved by describing the object as a set of related geometric primitives (lines, conics, etc.), and using redundant measurements to facilitate the detection of outliers. This improves the overall tracking performance. Results are given for frame rate tracking on image sequences. Proposed method have many flaws and it require an improvement.

Gary R. Bradski in 1998[1] deal with Real time face and object tracking as a component of perceptual user interface. They have proposed camshift algorithm to track the human faces.

Mean shift find the peak of probability distribution. But distribution can change dynamically with time so system must be adaptive, which is provided by camshift. But problem arises when both background and foreground are same, illumination variation etc. occurs.

Yu Zhong, Anil k. Jain, Fellow IEEE, M.P.Dubuisson-Jolly, Member IEEE in 2000 [4] Track an object using Deformable Templates. Prototype-based application is used in deformable template for tracking an object in image sequences from different sources. They combines to terms for achieve their goal: the frame-to-frame deviations of the object shape and the dependability of the modeled shape to the input image. The deformable template uses prior shape information which is extracted from the previous frames along with a systematic shape deformation scheme to model the object shape in a new frame. proposed framework have many flaws and there is need to improve the tracking result.

Dorin Comaniciu, Visvanathan Ramesh, and Peter Meer [17] in2003 apply the mean shift method to feature space analysis for object tracking. Their semi-automatic method works by searching in each frame for the location of the target region, where the color histogram is similar to the reference color histogram of the tracked target. The feature histogram-based target representations are regularized by spatial masking with an isotropic kernel. The masking induces spatially-smooth similarity functions suitable for gradient-based optimization. Hence, the target localization problem can be formulated using the basin of attraction of the local maxima. They employ a metric derived from the Bhattacharyya coefficient as similarity measure, and use the mean shift procedure to perform the optimization.

Mei Han, Wei Xu, Hai Tao, Yihong Gong, in 2004[52] proposed an Algorithm for multiple object tracking that seeks the optimal state sequence which maximizes the joint state-observation probability. An observation model which is composed of the original image, the foreground mask given by background subtraction and the object detection map generated by an object detector. The image provides the object appearance information. The foreground mask enables the likelihood computation to consider the multi-object configuration in its entirety. The detection map consists of pixel wise object detection scores, which drives the tracking algorithm to perform joint inference on both the number of objects and their configurations efficiently.

Mei Han Amit Sethiy Yihong Gong in 2004[48] tracks Multiple Object based on Detection method. A multiple hypothesis method is used to track multiple objects based on object detection. We apply a neural network based object detection module to detect pedestrians. The tracking algorithm accumulates the detection results in a graph-like structure and maintains multiple hypotheses of objects trajectories. At the same time, the multiple objects tracking method gives feedbacks which are predictions of object locations to the object detection module.

Hieu T. Nguyen and Arnold W.M. Smeulders in 2004 [51], tracks Fast Occluded Object by a Robust Appearance Filter. object tracking in image sequences using template matching. To update the template, appearance features are smoothed temporally by robust Kalman filters, one to each pixel. The resistance of the resulting template to partial occlusions enables the accurate detection and handling of more severe occlusions. Abrupt changes of lighting conditions can also be handled, especially when photometric invariant color features are used. But it cannot be applied for non-Gaussian objects.

Bing Han, William Roberts & Group [5] In 2005 Track using Robust Feature based algorithm. An unsupervised object tracking, i.e., without prior knowledge about the object to be tracked. To address this problem, they using feature points (or landmark points) to represent objects. Feature-based object tracking consists of feature extraction and correspondence. Feature correspondence is particularly challenging since a feature point in one image may have many similar points in another image, resulting in ambiguity in feature correspondence. To resolve the ambiguity, algorithms, which use exhaustive search and correlation over a large neighborhood, have been proposed. However, these algorithms incur high computational complexity, which is not suitable for real-time tracking.. It provides superior performance over the existing methods; it has proven computationally inexpensive and robust to various types of object motion. But it does not implemented for Real time.

Ning song Peng, jie yang, zhi liu [10] in 2005 proposed a target model updating method in the mean shift algorithm, in which authors propose to integrate an adaptive KALMAN filter into the mean shift algorithm to update the target model and to handle temporal appearance changes. However, to update the target model in each frame makes the tracker computationally expensive and sensitive to occlusions and noise. They propose a new adaptive model update mechanism for the real-time mean shift blob tracking. Kalman filter has been used for filtering object kernel histogram to obtain the optimal estimate of the object model because of its popularity in smoothing the object trajectory in the tracking system. The acceptance of the object estimate for the next frame tracking is determined by a robust criterion, i.e. the result of hypothesis testing with the samples from the filtering residuals. Therefore, the tracker can not only update object model in time but also handle severe occlusion and dramatic appearance changes to avoid over model update. They have applied the proposed method to track real object under the changes of scale and appearance with encouraging results.

Richard, John and Jesse in 2006 [28] Track a Robust Real Time non-rigid objects based on color Thresholding. Tracking is performed on non-rigid objects in a sequence of video frames based on a user-selected region of the initial frame. user selected region of the initial frames using K-means Algorithm. But they were not combining other features of the video such as edges and texture together with color information.

Yizheng Cai, Nando de Freitas, and James J. Little in 2006[53], Robust Visual Tracking for Multiple Targets, The particle filter technique is adopted and modified to fit into the multi-target tracking framework. A rectification technique is employed to find the correspondence between the video frames coordinates and the standard hockey rink coordinates so that the system can compensate for camera motion and improve the dynamics of the players. A global nearest neighbor data association algorithm is introduced to assign boosting detections to the existing tracks for the proposal distribution in particle filters. The meanshift algorithm is embedded into the particle filter framework to stabilize the trajectories of the targets for robust tracking during mutual occlusion.

Bastian Leibe,Konrad Schindler and Luc Van Gool in 2007[9] Present a Novel approach for multi-object tracking which consider an object detection and space time trajectory estimation as a coupled optimization problem, Track multiple objects i.e. Complex objects. It is formulated in a hypothesis selection framework and builds upon a state-of-the-art pedestrian detector But have following problems trajectory initialization and target following. This get reduces only but not removed.

Jaideep Jeyakar, R. Venkatesh Babu, K. R. Ramakrishnan in 2007 [12] proposed a robust tracking algorithm which overcomes the drawbacks of global color histogram based tracking. We incorporate tracking based only on reliable colors by separating the object from its background. A fast yet robust model updation is employed to overcome illumination changes. This algorithm is computationally simple enough to be executed real time and was tested on several complex video sequences.

Yuhua Zheng and and Yan Meng in 2008 [11] Track and detect face using Bayes constrained particle swarm optimization (BC-PSO), which is a population based searching algorithm. A cascade of boosted classifiers based on Haar-like features is trained and employed for object detection. It is having several issues such as more object features need to be embedded to train the object model under different environment and light conditions. There is needed to make structure of weak classifiers to more adaptive by extending those classifiers.

J´ er ˆome Berclaz, Franc¸ois Fleuret, Engin T¨ uretken, and Pascal Fua in 2009[47] proposed Multiple Object Tracking using K-Shortest Paths Optimization. Multi-object tracking can be achieved by detecting objects in individual frames and then linking detections across frames. In this paper, we show that reformulating that step as a constrained flow optimization results in a convex problem. We take advantage of its particular structure to solve it using the k-shortest paths algorithm, which is very fast. This new approach is far simpler formally and algorithmically than existing techniques and lets us demonstrate excellent performance in two very different contexts.

Du-Ming Tsai and Shing-ChinLai in 2009[13] were subtract background using ICA for indoor Surveillance. Background subtraction concatenate two stages, one for training and the other for detection. The proposed ICA model can well separate two highly-correlated images. In the detection stage, the trained de-mixing vector is used to separate the foreground in a scene image with respect to the reference background image. Two sets of indoor examples that involve switching on/off room lights and opening/closing a door are demonstrated in the experiments, but, are not for outdoor surveillance.

J. Sullivan, P. Nillius and Stefan Carlsson [15] in 2009Tracks Multi-targets by locating the targets and labeling their identities. The latter is a challenge when many targets, with indistinct appearances, frequently occlude one another, as in football and surveillance tracking. We present an approach to solving this labeling problem. When isolated, a target can be tracked and its identity maintained. While, if targets interact this is not always the case. We build a track graph which denotes when targets are isolated and describes how they interact. Measures of similarity between isolated tracks are defined. The goal is to associate the identities of the isolated tracks, by exploiting the graph constraints and similarity measures. But time taken by it is large.

Ming-Yi et al. [6] in 2010 presented an improved mean shift tracking algorithm using a fuzzy color histogram. A fuzzy color histogram generated by a self-constructing fuzzy cluster is proposed to reduce the interference from lighting changes for the mean shift tracking algorithm. Although it is a good idea to avoid the pre-defined color bins. The tracking therefore may fail if the appearance of the object varies substantially. The experimental results show that the proposed tracking approach is more robust than the conventional mean shift tacking algorithm and the cost of increasing computation time is also moderate.

Vasilis Papadourakis , Antonis Argyros in 2010[49] Multiple objects tracking in the presence of long-term occlusions. This algorithm handles spatially extended and temporally long object occlusions. It is based on the concept of object permanence which suggests that a totally occluded object will re-emerge near its occluder. This method does not require prior training to account for differences in the shape, size, color or motion of the objects to be tracked. Instead, the method automatically and dynamically builds appropriate object representations that enable robust and effective tracking and occlusion reasoning. This method fails when background and foreground have same color.

Michael, Fabian Bastian and Group of IEEE member [14] in 2010, Track and Detect "Online Multi-Person from a single Uncalibrated Camera using Partial Filter."The algorithm detects and tracks a large number of dynamically moving persons in complex scenes with occlusion. Their algorithm uses the continuous confidence of pedestrian detectors and online trained, instance-specific classifiers as a graded observation model. The algorithm detects and tracks a large number of dynamically moving persons in complex scenes with occlusions, does not rely on background modeling, requires no camera or ground plane calibration, and only makes use of information from the past. Thus, generic object category knowledge is complemented by instance-specific information. They analyze the influence of different algorithm components on the robustness It requires a more sophisticated framework than Partial Filtering.

Leichter et al. [16] in 2010 presented an improved mean shift tracking algorithm based on multiple reference color histograms in which authors proposed to update the target model at tracking over time with multiple histograms. This strategy still computationally expensive and insufficient in several real world conditions, especially if the target color is similar to the background components' color at tracking over time. In contexts where multiple views of the target are available prior to the tracking, this paper enhances the Mean Shift tracker to use multiple reference histograms obtained from these different target views. This is done while preserving both the convergence and the speed properties of the original tracker. They first suggest a simple method to use multiple reference histograms for producing a single histogram that is more appropriate for tracking the target. Then, to enhance the tracking further, they propose an extension to the Mean Shift tracker where the convex hull of these histograms is used as the target model. Many experimental results demonstrate the successful tracking of targets whose visible colors change drastically and rapidly during the sequence, where the basic Mean Shift tracker obviously fails.

Azghani et al. [18] in 2010 suggested an intelligent modified mean shift tracking algorithm using a local search based on a genetic algorithm to improve the convergence procedure. However, their method still limited around the limitations of genetic algorithms and not robust in several conditions (e.g. light change, scale change, etc.), because no information has been used to improve the target description, which is the major need of any tracker. First, a background elimination method is used to eliminate the effects of the background on the target model. The mean shift procedure is applied only for one iteration to give a good approximate region of the target. In the next step, the genetic algorithm is used as a local search tool to exactly identify the target in a small window around the position obtained from the mean shift algorithm. The simulation results prove that the proposed method outperforms the traditional mean shift algorithm in finding the precise location of the target at the expense of slightly more complexity.

Hongwei Ying, Xuena Qiu, Jiatao Song, Xiaobo Ren [44] in 2010 tracks Object Based on Texture and Color. A particle filtering object tracking approach is proposed in this paper based on local binary pattern and color feature. Color histogram is the global description of object in color image, while local binary pattern texture contains local feature information of neighbor region texture in gray image. Target is represented by both histogram of color and local binary pattern which are combined under the frame of particle filtering. But it fails due to degeneracy and improvishment.

Juan Villalba Espinosa, José María González Linares, Julián Ramos Cózar, Nicolás Guil Mata in 2011 [7] uses Kernel-Based Object Tracking Using a Simple Fuzzy Color Histogram. Tracking is done using the HSV color space as the feature space and fuzzy color histograms as feature vectors. These histograms are more robust to illumination changes and quantization errors than common histograms. To avoid a significant increase in the computational complexity, a simple fuzzy membership function is used.

Martin Godec Peter M. Roth Horst Bischof in 2011[46] presents a Hough-based Tracking of Non-Rigid Objects. Hough Forests to the online domain and couple the voting based detection and back-projection with a rough segmentation based on GrabCut. This significantly reduces the amount of noisy training samples during online learning and thus effectively prevents the tracker from drifting. In the experiments, we demonstrate that our method successfully tracks a variety of previously unknown objects even under heavy non-rigid transformations, partial occlusions, scale changes and rotations.

Mahbub murshed, Md. Hasanul kabir and oksam chae in 2011[50] track moving object based on edge segment based approach. Canny edge detector is used for edge segmentation. Curvature based features are used for moving edge registration due to its transformation invariance nature. kalman filter is used for track individual segment. each segment is clustered using mean shift. This method fails in complex environment.

Amir Salarpour, Arezoo Salarpour, Mahmoud and MirHossein in 2011 [19] Track “A Vehicle using Kalman Filter and Features.” It detects all moving objects. It works satisfactory in dynamic motion. They detect all moving objects, and for tracking of vehicle they use the Kalman filter and color feature and distance of it from one frame to the next. So the method can distinguish and tracking all vehicles individually. This method has tracking problem such as .Appearance; .Disappearance; Occlusion.

Prajna , parimita dash,dipti patra in 2012 [20] work with Efficient object tracking method using LBP based Texture feature and OHTA color moment. This paper extends the classic Mean Shift tracking algorithm by combining color and texture features. In the proposed method, firstly, both the color feature and the texture feature of the target are extracted from first frame and the histogram of each feature is computed. Then the Mean Shift algorithm is run for maximizing the similarity measure of each feature independently. In last step, center of the target in the new frame is computed through the integration of the outputs of Mean Shift. Experiments show that the proposed Mean-Shift tracking algorithm combining color and texture features provides more reliable performance than single features tracking.

In spite of all attempts [26][30][42][43] to improve the mean shift tracking algorithm, the complex conditions of the real world remain the biggest challenge, which require the use of a very powerful and rich descriptor for better target representation. Until now, the proposed improvement into the mean shift tracker remains in target description by isolated pixels such as the color histogram and texture that lacks of spatial configuration of pixels. These features are insufficient and often invalid in practice, mainly in presence of noise, clutter, illumination change, and local deformation. We believe that one way to improve the tracking in complex conditions is not by using direct information from isolated pixels as the color histogram but through increasing the level of the target description. This level can be described through the exploitation of discriminating and invariant internal targets properties computed from local dependencies between a set of pixels within the target region, such as: local variation, degrees of texture organization, rate of homogeneity, disorder degrees, edge direction, spatial context, color context, etc. Contextual information plays an important role in objects description for objects recognition, classification [39].

Earlier Researchers use meanshift to track a moving face [43] but Mean Shift is based on static distributions, which are not updated unless the target experiences significant changes in shape, size or color. some reserchers focus on particle filter which can apply non-linear and non gaussian condition with a lots of probability hypothesis on visual tracking. But particle filter found a shortcoming in creating the samples due to degeneracy and improvishment. And huge computational complexity limits the use of particle filter. So we proposed modification of meanshift i.e. camshift [24] [28][29][32]. But camshift also suffers from a limitation i.e. when foreground and background colors are same then camshift is not reliable and robust. Texture feature is used to solve this problem with camshift. Histogram based method, GLCM matrix, GLRLM, LBP, Auto correlation feature etc. can be used and on the basis of following survey we have used LBP.

First Order Histogram Based Features: First Order histogram provides different statistical properties such as 4 statistical moments of the intensity histogram of an image. They depend only on individual pixel values. Four first order histogram statistics are

Mean, Variance, Skewness and Kurtosis. It's limitation that it provides no information about the relative position of pixels (such as a checkerboard and a Salt & Pepper noise pattern.Grey Level Co-occurrence Matrix Features : The gray-level co-occurrence matrix (GLCM) or gray-level spatial dependence matrix based calculations fall under the category of second-order statistics. A set of 14 textual features which can be extracted from the co-occurrence matrix, and which contain information about image textural characteristics such as homogeneity, contrast and entropy. A gray level co-occurrence matrix (GLCM) contains information about the positions of pixels having similar gray level values. It is a two-dimensional array, P, in which both the rows and the columns represent a set of possible image values. But GLCMs are recognized as sparse matrices, excessive computation is required to generate the co-occurrence texture features. Gray Level Run Length Matrix Features Galloway [38] introduced different run-length matrices as feature representatives. For a given image, a run-length matrix P (I, j) N be the maximum run length. Five features derived from the GLRLM. These features are: Short Runs Emphasis (SRE), Long Runs Emphasis (LRE), Grey Level Non-Uniformity (GLNU), Run Length Non-Uniformity (RLNU), and Run Percentage (RPERC).Limitation is the number of intensity levels in an image needs to be limited using one or more thresholds. Local binary pattern LBP operator labels the pixels of an image with decimal numbers, which are called *LBPs* or *LBP codes* that encode the local structure around each pixel. Each pixel is compared with its eight neighbors in a 3 × 3 neighborhood by subtracting the center pixel value; the resulting strictly negative values are encoded with 0, and the others with 1. For each given pixel, a binary number is obtained by concatenating all these binary values in a clockwise direction, which starts from the one of its top-left neighbor. The corresponding decimal value of the generated binary number is then used for labeling the given pixel. The derived binary numbers are referred to be the LBPs or LBP codes An autocorrelation function measures the linear spatial relationships between spatial sizes of texture primitives. Autocorrelation-based approach to texture analysis is based on the intensity value concentrations on all or part of an image represented as a feature vector. [37] [38] [39][40] Calculation of the autocorrelation matrix involves individual pixels. But it is well suited only for non-homogeneous database. Law's Texture Energy Features A set of nine 5x5 convolution masks is used to compute texture energy, which is then represented by a vector of nine numbers for each pixel of the image being analyzed. Texture features that certain operators such as Laplacian and Sobel operators accentuated the underlying microstructure of texture within an image. Gabor Filter-based Texture Features The 2-D Gabor filters have been proved to be an important tool in texture analysis. They consist of a sinusoidal plane wave of some frequency and orientation modulated by Gaussian envelope. A Gabor filter is a band-pass filter which can be used to extract a specific band of frequency components from an image. An image is convolved with a 2D Gabor function to obtain a Gabor feature image, and by varying spatial frequency and orientations, a bank of different Gabor filters can be produced. Wavelet-based Feature the region-based systems which use wavelet transform are classified into three categories: a hierarchical block, a moving window and a pixel. Since these methods are subject to some limitations, several improvements have been proposed. The method proposed in segments an image into some regions by clustering pixels. Then the texture features are calculated from wavelet coefficients of all regions (subbands). The segmented regions are indexed by the averaged features in the regions. Random Field Feature In this approach, relationships between the gray level of neighboring pixels are statistically characterized. Image textures are modeled as a Markov random field of pixels gray level. Fractal Feature: Fractal features can be considered to develop discriminative and invariant features for texture classification, especially in cases where scale changes are prominent in textures. Such variations can be handled by choosing local interest points and selecting their characteristic scales. Problems related to have suitable local interest points, it is difficult to have an optimized characteristic scale that can be suitable for different textures of varied scales .

Chapter 3

PROPOSED FACE DETECTION AND TRACKING SYSTEM

In the proposed paper CamShift algorithm is used, it has been recently proven to be a powerful and reliable algorithm for tracking moving objects. It is a robust non-parametric kernel based tracker. CamShift algorithm which uses a one dimensional histogram to track an object with known hue in color images sequences. Camshift algorithm cleverly exploits mean-shift by changing the size of the window when it came to convergence. The Camshift coupled is adapted to color image sequences, and is used to track in real time system. The mean location of a target object is found by computing zeroth, first and second order image moments. The position and dimensions of the search window are updated iteratively until convergence occurs [1] [28] [29] [32]. But problem may arise when one wishes to use CamShift to track multi-hued objects or objects where hue alone cannot allow the object to be distinguished from the background and other objects. To solve this problem we used texture information with hue value. A textured area in an image can be characterized by a nonuniform or varying spatial distribution of intensity or color. In general, when a target and its corresponding background have similar colors, they have different textures. No. of texture technique are there such as First Order Histogram Based Features, Grey Level Co-occurrence Matrix Features, Gray Level Run Length Matrix Features, local binary pattern, Autocorrelation Features, Law's Texture Energy Features, Gabor Filter-based Texture Features, Wavelet-based Feature, Fractal Feature etc. First order histogram based features can't be used as it doesn't provide any information about the relative position of pixels [23] [25] [35] [36]. In GLCM, features are recognized as sparse matrices, excessive computation is required to generate the co-occurrence texture features. In GLRLM the number of intensity levels in an image needs to be limited using one or more thresholds [37]. Autocorrelation feature doesn't well suit for non homogeneous database. In this paper we refer to Local binary pattern (LBP), proposed by Ojala [40] [41] as local texture operator. Its computational complexities are less, and it is easy to satisfy real-time in moving objects tracking. It is a non-parametric kernel. It is invariant to monotonic grey- scale transformations and rotation invariant which is very important for texture analysis. LBP operator labels the pixels of an image with decimal numbers, which are called *LBPs* or *LBP codes* that encode the local structure around each pixel. Each pixel is compared with its eight neighbors in a 3 × 3 neighborhood by subtracting the center pixel value; the resulting strictly negative values are encoded with 0, and the others with 1. For each given pixel, a binary number is obtained by concatenating all these binary values in a clockwise direction, which starts from the one of its top-left neighbor. The corresponding decimal value of the generated binary number is then used for labeling the given pixel. The derived binary numbers are referred to be the LBPs or LBP codes. The LBP method can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis.

The closest existing algorithm to CAMSHIFT is known as the mean shift algorithm [2].

Mean Shift Algorithm

1. Choose a search window size.
2. Choose the initial location of the search window.
3. Compute the mean location in the search window.
4. Center the search window at the mean location computed in Step 3.
5. Repeat Steps 3 and 4 until convergence (or until the mean location moves less than a preset threshold).

3.1 Face detection and improved Camshift tracking algorithm by unifying Uniform LBP and hue features.

As camshift proven as robust and widely used moving object tracking algorithm in last few years. Mean shift is key algorithm for camshift. Camshift adjust adaptively window size acc. to size of tracked object (face). Camshift algorithm is based on hue component of moving objects. But problem arises when tracked object and background have same color (hue) component, illumination condition varies severely. In this paper we combine Uniform LBP texture and hue component on multifarious objects to solve these tribulations.

The steps for improved CAMshift algorithm are as follows:-

1. Read the input video frames till a face is detected using uniform LBP and haar like features and find the region in which face lie down.
2. Find multifarious objects for tracking i.e. face unifying with nose or face unifying with eyes or mouth plus face etc. we have focused on face unifying with nose in our algorithm.
3. Now apply camshift on Created target object i.e initialize search window on face unifying with nose. and convert target object into HSV from RGB.
4. Now obtain a color probability distribution function through color histogram.
5. Extract binary images of color bins and Calculate back projection image
6. calculate Zero, first and second order moments and center of mass of search window
7. Update parameters of tracking window until step 6 converges i.e. score value matches.
8. Process the next frame with the search window position from the step 7.

3.2 CamShift Algorithm

Mean shift algorithm is about find the peak of a probability function by the change of the mean of the data. Image is segmented into several logical regions. It climbs the gradient of a probability distribution to find the nearest domain mode (peak). It uses kernel based distribution function to find pdf. Given a likelihood image, find the optimal location of the tracked object by finding $\frac{dp(x)}{dx} = \nabla P(x)$. and peak exists where $\nabla P(x)\ is\ found\ zero.$ But implementation of mean shift algorithm is very complex. It is found than implementation can be made easy by found zero ,first and second order moment.

Camshift algorithm can be summarized as follows:-

1. Choose initial search window in ROI i.e the region which contain the object we want to track and Convert image into HSV image
2. Obtain a color probability distribution function through color histogram
3. Extract binary images of color bins and Calculate back projection image

4. calculate Zero, first and second order moments and center of mass of search window
5. Update parameters of tracking window until step 4 converges
6. Process the next frame with the search window position from the step 5.

3.2.1 CAMshift for video systems

When we track a video, CAMSHIFT deal with color probability distribution image generated from color histogram. Camshift calculates the centroid of color pdf within its 2D present window, it re-enters the search window and then calculate the area for next window. Due to automatic updating of search window it does not requires calculating the color probability distribution over the whole image, but constrain the calculation of the distribution to a smaller image region nearby the current CAMSHIFT window. So camshift reduce lots of computational complexities.

3.2.2 How CAMSHIFT Deals with Image robustly even if some challenges are there

When we track an object with camshift, various problems occurs like illumination variation, occlusion, complex object motion, shape and noise etc. CAMSHIFT continuously update his search window, and fit to the requirement. CAMSHIFT produces very little jitter in noise and, as a result, tracking variables do not have to be smoothed or filtered. This gives us robust noise tolerance. CAMSHIFT's robust ability to ignore outliers also allows it to be robust against distractors. Once CAMSHIFT is locked onto the mode of a color distribution, it will tend to ignore other nearby but non-connected color distributions. Camshift doesn't leave the original object due to occlusion of other objects. CAMSHIFT's provable convergence to the mode of probability distributions helps it ignore partial occlusions of the colored object. CAMSHIFT will tend to stick to the mode of the color distribution that remains. A colored object's potential velocity and acceleration scale with its distance to the camera, which in turn, scales the size of its color distribution in the image plane. Thus, when objects are close, they can move rapidly in the image plane, but their probability distribution also occupies a large area. In this situation, CAMSHIFT's window size is also large and so can catch large movements. When objects are distant, the color distribution is small so CAMSHIFT's window size is small, but distant objects are slower to traverse the video scene. This natural adaptation to distribution scale and translation allows us to do without predictive filters or variables–a further computational saving–and serves as an in-built antidote to the problem of erratic object motion. This is just what is desired for tracking whole objects such as faces, hands, and colored tools. This property enables CAMSHIFT to not get stuck tracking, for example, the nose of a face, but instead to track the whole face.

3.2.3 Calculate probability distribution function and histogram back projection

Camshift is a probabilistic model based on color. To track a human face a training model is generated which is trained for face in terms of color. For every color probability is calculated. This algorithms use the model to compute the probability of the pixel in a frame that contain a face in the video sequence. The algorithm replaces each pixel in a video frame with the probability that it is obtain by a face color pixel., the brighter the pixel, the more likely it is to be part of a face. Now according to algorithm we define ROI which is represented by a bounding box corresponding to the person we want to track. Then the histogram is computed by simply counting the number of pixels in an image of the face. The most frequently occurring color is assigned the highest probability (1.0), and the probabilities of other colors are computed based on their frequency relative to that of the most frequently occurring Color. But there are several problems by using color’s stochastic model (RGB) for face tracking. Every color composed of basically two properties one is chromocity and other is luminosity. First problem is that every face doesn’t have same face color and second is that luminosity varies widely according to environmental condition. So to solve these problems model used for face tracking must be trained for a specific person or be flexible enough to accommodate individuals With varying skin colour. Worse, the skin colour of a particular individual may vary across the face (e.g., a pale-skinned man with a dark-coloured beard). And under extremely bright or extremely dim conditions, colour is very hard to accurately capture using common digital video equipment. We use Bradski’s model [1], in developing CAMSHIFT, he, observed that all humans (except albinos) are basically have the same hue. Hence, he uses the HSV colour space as the basis for the colour model in CAMSHIFT. By using only hue component we make object’s color 1D histogram. This histogram is stored to convert next frames into corresponding probability of the object.We also define a mask for the histogram calculation, which is the foreground image, to calculate the histogram only for the person, and not for the background inside the bounding box. The result is called backproject image. . With this method, we were able to find the location of the person in the whole frame.CAMSHIFT is then used to track the object based on this backproject image.

3.2.3 *Mass center, orientation and scaling calculation*

Using the previous location of the person from A, we detect the new position of the moving person and use it as starting search window for the next frame. The search window centroid of an area of high probability is computed by the mean location of the probability image using zero and first order moments. And the orientation and scale are calculated using second order moments.

Calculation of centroid by Zero and first order moment is given below:-

$$M_{00} = \sum_{x} \sum_{y} I(x, y) \quad (1)$$

$$M_{10} = \sum_{x} \sum_{y} xI(x, y) \quad (2)$$

$$M_{01} = \sum_{x} \sum_{y} yI(x, y) \quad (3)$$

$$x_c = \frac{M_{01}}{M_{00}}, y_c = \frac{M_{10}}{M_{00}} \quad (4)$$

Where M00 in eq.(1) is the zeroth moment, M10 in eq.(2) and M01 in eq.(3) are the first moments, these moments could be used to compute the next center position of the tracking window xc and yc as shown in eq.(4). Then, back to step 3, we calculate the new histogram of the person to update the previous one.

Calculation of orientation (Ɵ) and scale using second order moment:-

The orientation (θ) of the major axis and the scale of the distribution are determined by finding an equivalent rectangle that has the same moments as those measured from the 2D probability distribution image

$$M_{02} = \sum_x \sum_y y^2 \; I(x,y) \tag{5}$$

$$M_{20} = \sum_x \sum_y x^2 \, I\,(x,y) \tag{6}$$

$$M_{11} = \sum_x \sum_y xyI(x,y) \tag{7}$$

The first two eigenvalues (the length and width of the probability distribution) are calculated in closed form as follows. From the intermediate variables a, b and c

$$a = \frac{M_{20}}{M_{00}} - {x_c}^2 \tag{8}$$

$$b = 2\frac{M_{11}}{M_{00}} - x_c \; y_c \tag{9}$$

$$c = \frac{M_{02}}{M_{00}} - {y_c}^2 \tag{10}$$

$$\Theta = \frac{1}{2}\tan^{-1}\left(\frac{b}{a-c}\right) \tag{11}$$

The distances l_1 and l_2 from the distribution centroid (the dimensions of the equivalent rectangle) are given by,

$$w = \sqrt{\frac{(a+c)-\sqrt{b^2+(a-c)^2}}{2}} \tag{12}$$

$$l = \sqrt{\frac{(a+c)+\sqrt{b^2+(a-c)^2}}{2}} \tag{13}$$

Camshift fails when foreground and background both have same hue i.e. color component. To solve this problem we used texture feature with hue component.

3.3 Local binary Pattern

The Local Binary Pattern (LBP) operator is a non-parametric 3x3 kernel which summarizes the local spacial structure of an image. It was first introduced by Ojala et al. who showed the high discriminative power of this operator for texture classification. At a given pixel position (xc, yc), LBP is defined as an ordered set of binary comparisons of pixel intensities between the center pixel and its eight surrounding pixels

45	67	87
44	59	37
80	79	95

After applying LBP operator →

0	1	1
0		0
1	1	1

LBP => 01101110=110

Figure 3.1 calculation of LBP

The decimal form of the resulting 8-bit word (LBP code) can be expressed as follows:

$$LBP_{P,R}(x_c, y_c) = \sum_{P=0}^{P-1} S(i_P - i_c)\, 2^P \qquad (14)$$

Where i_c corresponds to the grey value of the center pixel (x_c,y_c), in to the grey values of the 8 surrounding pixels, and function s(x) is defined as:

$$S(x) = \begin{Bmatrix} 1\ if\ x \geq 0 \\ 0\ if\ x < 0 \end{Bmatrix} \qquad (15)$$

Note that each bit of the LBP code has the same significance level and that two successive bit values may have a totally different meaning. Actually, The LBP code may be interpreted as a kernel structure index. By definition, the LBP operator is unaffected by any monotonic gray-scale transformation which preserves the pixel intensity order in a local neighborhood.

AFTER APPLYING LBP OPERATOR

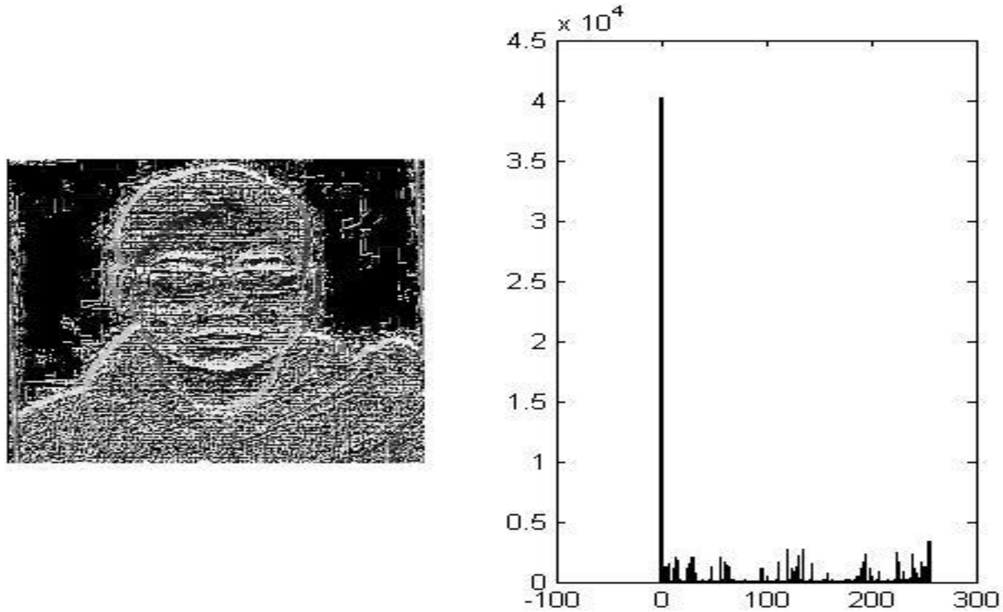


Figure3.2: Original image (up), LBP image and LBP histogram (bottom)

Due to its texture discriminative property and its very low computational cost, LBP is becoming very popular in pattern recognition. Recently, LBP has been applied for instance to face detection, face localization, face recognition image retrieval, motion detection or visual inspection.

3.3.1 Uniform LBP

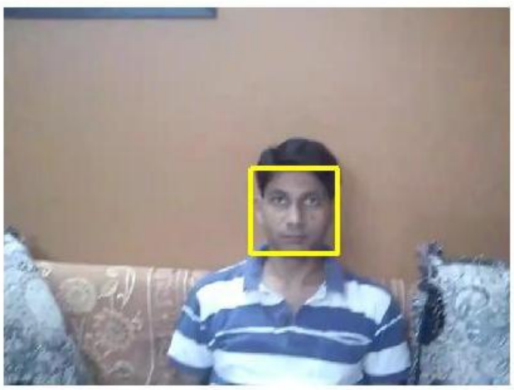
The texture model derived by has only gray-scale invariance. The grayscale and rotation invariant LBP texture model is obtained by [26]

$$LBP_{P,R}^{riu\,2} = \left\{\sum_{P=0}^{P-1} S(i_P - i_c)\ if \quad U\left(LBP_{P,R}\right) \leq 2\right\} \tag{16}$$

$$U\left(LBP_{P,R}\right) = \left|S\left(i_{p-1} - i_c\right)S\left(i_0 - i_c\right)\right| + \sum_{P=1}^{P-1}\left|S\left(i_p - i_c\right)S\left(i_{P-1} - i_c\right)\right| \tag{17}$$

By definition, the P + 1 “uniform” binary patterns occur in a circularly symmetric neighbor set of P pixels. Equation assigns a unique label to each of them corresponding to the number of “1” bits in the pattern (0 to P), while the “nonuniform” patterns are grouped under the “miscellaneous” label P+1.

The model has nine uniform texture patterns, which are shown in Figure Each of the uniform patterns is called as a micro-texton. The micro-texon includes spots, flat areas, edges, line ends and corners, etc. In Fig. 3, the white circles represent “1” and the black circles represent “0”.



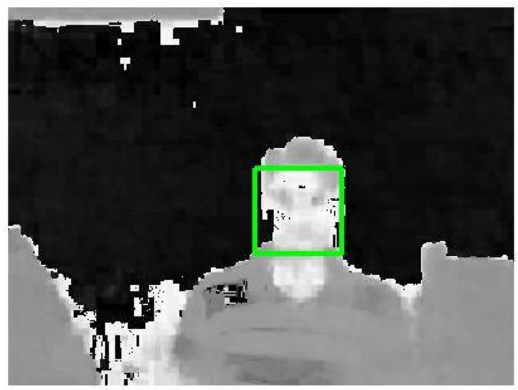


Figure 3.3: uniform LBP pattern For Face

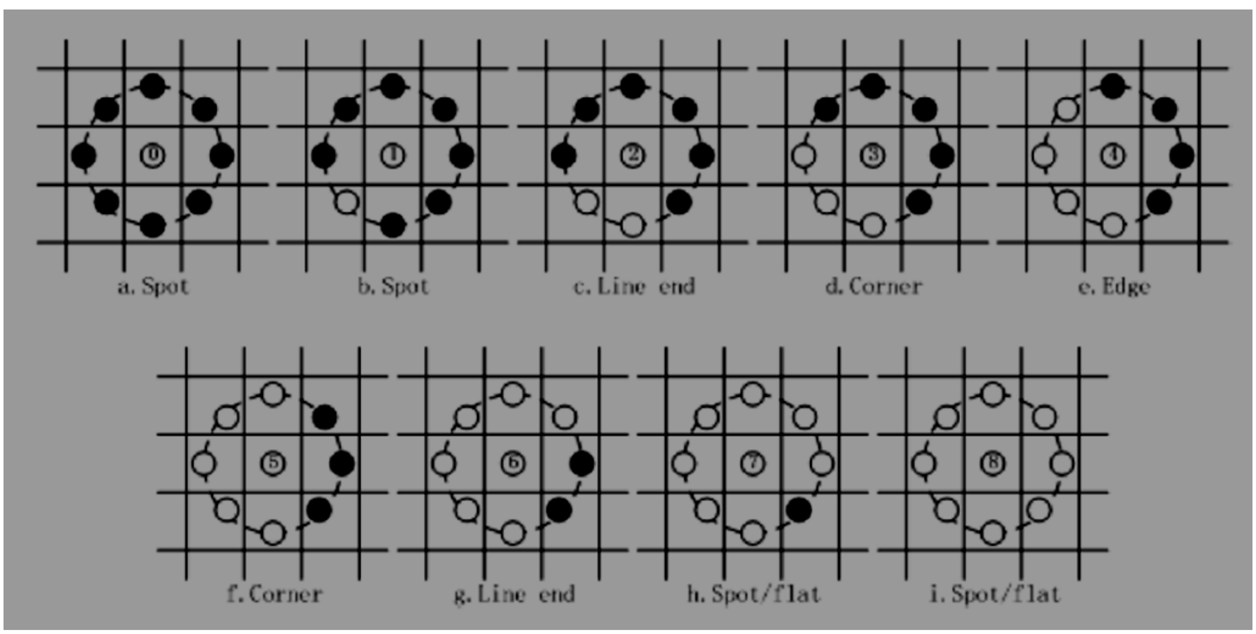


Figure 3.4: 0,1,7,8 are the Spots or Flat Area, 2, 6 are the Line End, 3,5 are the Corner, 4 is the Edge.

In target representation, the micro-textons such as edges, line ends and corners, by name of “major uniform patterns”, represent the main features of target, while spots and flat areas, called “minor uniform patterns”, are minor textures. In LBPriu2 8,1 the labels corresponding to minor uniform patterns are 0, 1, 7 and 8 respectively, and the label of non-uniform patterns is 9. The labels corresponding to main uniform patterns are 2–6, which have five patterns. background and extracts the main features. It reduces greatly the interferences induced by minor uniform patterns, which come from mainly the smooth background and noise in the target area.

Figure 3.5: uniform LBP pattern for nose

We use viola-jones [51] [52] cascaded classifiers which are consist of many stages, where each stage ensemble a weak classifier. The weak learners are simple classifiers called decision stumps. Each stage is trained using a technique called adaboosting. Boosting provides the ability to train a highly accurate classifier by taking a weighted average of the decisions made by the weak learners.

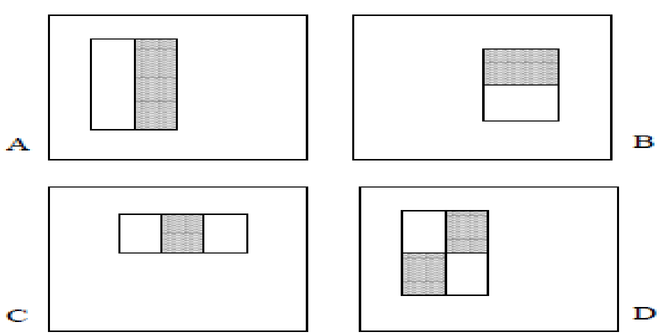


Figure 3.6: Two-rectangle features are shown in (A) and (B). Figure (C) shows a three-rectangle feature, and (D) a four-rectangle feature.

Rectangle features shown relative to the enclosing detection window. The sums of the pixels which lie within the white rectangles are subtracted from the sum of pixels in the grey rectangles. By the help of above integral features we find nose.

Chapter 4

Experiment and results

We have prepared our own Dataset, where we have taken a background which is similar to the skin color. The dataset is taken under variable lighting conditions. First the street lights were OFF and gradually it was made ON to make the situation complex. Pose variation, scaling and orientation are also taken. Partial occlusion is also involved.

The experiment is executed in MATLAB 2012 on an I3 Processor Laptop. The frames are in 640*480 resolutions and they are sampled at a rate of 20 samples per second. In order to verify and evaluate the performance of our face detection system, it is tested in different and challenging environments. The experiment is performed in different videos. At the first stage of our system, front face is detected using LBP. Then nose is detected using haar like features. Individual frame is taken and Image is then converted to HSV color space. Camshift is applied for tacking.

Dataset A is the dataset where background and foreground have similar colors with some orientation. Dataset B is the showing scaling. Dataset C is the showing multifaces with illumination changes. Dataset D is the showing occlusion; Dataset E is the showing orientation and Dataset E is the tracking multiple objects.

The table given below shows the comparative analysis of the various Datasets with our algorithm.

ws the outputs of different Datasets. The highest Detection Percentage is found out to be 98.9% of Dataset B, while the lowest is of Dataset C where the Detection Rate is 76.3%. The Detection percentage is less in Dataset C because there is very fast variation in no. of faces and background so the detection percentage falls. So, very fast changing background remained as a problem in our human detection system. It also shows a Detection Percentage of above 87.6% in case of Multiple Humans.

Chapter 5

SUMMARY

Advantages and Limitations

The system is very robust, fast and is applicable in any complex scene as shown above. It is testified to have a worthy efficiency in case of single face as well as for multiple faces. This system is also robust to various illumination changes, Occlusion Handling, orientation and scaling. However, if the tracked Human goes out of Frame for some time and again comes in the scene, it can be tracked. But sometime it takes time to re-detect the face. Another limitation is false detection. There is also scope of improvement in the Accuracy of the System.

Conclusion

In this paper, we present face detection and tracking system using Camshift, LBP and haar like structure. First we have detected a face and then we find nose for better tracking result. Minimum size for face required is [24, 24]. Face is detected with the help of LBP and nose is detected with cascaded object. The highest detection rate achieved by our system is 98.9% in Dataset B while the Lowest Detection Rate achieved by our system is 76.3% in Dataset C. The algorithm is also tested for Multiple Humans and the Detection Ratio calculated is 87.6%.

Various face Detection methods are mentioned in the table below along with the Dataset used, speed of implementation and the detection ratio of various methods along with our method.

Further Work

A process of Normalization can be incorporated to the system in order to provide a better images for further processing purpose. We can add the incorporate Multiple Feature to improve the accuracy of the system and reduced the false detection.

REFRENCE

[1] Gary R. Bradski,"Real time face and object tracking as a component of perceptual user interface", 1998,IEEE.

[2] A. Yilmaz, X. Li, and M. Shah, "object tracking survey," ACM Computing Surveys, Vol. 38, No. 4, Article 13, Publication date: December 2006.

[3] HUEY-MIN SUN, SHU-MIN CHIANG," Tracking Multitarget in Cluttered Environment", IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS VOL. 28, NO. 2 APRIL 1992

[4] In 2000 Yu Zhong, Anil k. Jain, Fellow IEEE, M.P.Dubuisson-Jolly, Member IEEE "Track an object using Deformable Templates", IEEE transactions on pattern analysis and machine intelligence, vol. 22, no. 5, 2000.

[5] Bing Han, William Roberts, Dapeng Wu, Jian Li ," Robust Feature-based Object Tracking", Air Force Research Laboratory under grant FA8650-06-1-1027.

[6] Ming-Yi Ju, Chen-Sen Ouyang, Hao-Shiu Chang," Mean shift tracking using fuzzy color histogram", In proceeding of: Machine Learning and Cybernetics (ICMLC), 2010 International Conference on, Volume: 6, IEEE.

[7] Juan Villalba Espinosa, José María González Linares, Julián Ramos Cózar, Nicolás Guil Mata," Kernel-Based Object Tracking Using a Simple Fuzzy Color Histogram", Advances in Computational Intelligence Lecture Notes in Computer Science Volume 6691, 2011, pp 513-519.

[8] Richard Y. D. Xu, John G. Allen, Jesse S. Jin," Robust Real-Time Tracking of Non-rigid Objects" Conferences in Research and Practice in Information Technology, Vol. 36, 2006.

[9] Bastian Leibe, Konrad Schindler, and Luc Van Gool,” Coupled Detection and Trajectory Estimation for Multi-Object Tracking”, in proceeding of: IEEE 11th International Conference on Computer Vision, ICCV 2007.

[10] Ning song Peng, jie yang, zhi liu,” Mean shift blob tracking with kernel histogram filtering and hypothesis testing”, Pattern Recognition Letters 26 (2005) 605–614.

[11] Yuhua Zheng and Yan Meng,” OBJECT DETECTION AND TRACKING USING BAYESCONSTRAINED PARTICLE SWARM OPTIMIZATION”, Computer Vision Research Progress, 2007.

[12] Jaideep Jeyakar, R. Venkatesh Babu, K. R. Ramakrishnan,” ROBUST OBJECT TRACKING USING LOCAL KERNELS AND BACKGROUND INFORMATION”, IEEE conferences, 2007.

[13] Du-Ming Tsai, Shia-Chih Lai: Independent Component Analysis-Based Background Subtraction for Indoor Surveillance. IEEE Transactions on Image Processing 18(1): 158-167 (2009).

[14] Michael D. Breitenstein, Fabian Reichlin, Bastian Leibe, Esther Koller-Meier, and Luc Van Gool, “Online Multi-Person Tracking-by-Detection from a Single, Uncalibrated Camera”, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, SUBMITTED JANUARY 2010, REVISED OCTOBER 2010.

[15] J. Sullivan, P. Nillius and Stefan Carlsson ,”Multi-target Tracking on a Large Scale: Experiences from Football Player Tracking”, Proceedings of the IEEE ICRA 2009 Workshop on People Detection and Tracking Kobe, Japan, May 2009.

[16] Leichter, I., Lindenbaum, M., Rivlin, E.: Mean shift tracking with multiple reference color histograms. Comput. Vis. Image Underst. (CVIU) 114(3), 400–408 (2010).

[17] Dorin Comaniciu, Visvanathan Ramesh, and Peter Meer,” Kernel-based object tracking”, IEEE Trans. Pattern Anal. Mach. Intell. **25**(5), 564–577 (2003).

[18] Azghani, M., Aghagolzadeh, A., Ghaemi, S., Kouzehgar, M.: Intelligent modified mean shift tracking using genetic algorithm. In: Proc. 5th International Symposium on Telecommunications, pp. 806–811 (2010)

[19] Amir Salarpour, and Arezoo Salarpour, and Mahmoud Fathi, and MirHossein Dezfoulian,” VEHICLE TRACKING USING KALMAN FILTER AND FEATURES”, Signal & Image Processing : An International Journal (SIPIJ) Vol.2, No.2, June 2011.

[20] Prajna , parimita dash,dipti patra,” EFFICIENT OBJECT TRACKING METHOD USING LBP BASED TEXTURE FEATURE AND OHTA COLOUR MOMENT” International Journal of Electronics and Communication Engineering (IJECE) ISSN: 2278-9901 Vol.1, Issue 2 Nov 2012 15-22 © IASET

[21] JIFENG NING, LEI ZHANG and DAVID ZHANG and CHENGKE WU :ROBUST OBJECT TRACKING USING JOINT COLOR-TEXTURE HISTOGRAM , International Journal of Pattern Recognition and Artificial Intelligence Vol. 23, No. 7 (2009) 1245–1263.

[22] Guo-wu YUAN, Yun GAO, Dan XU,” A Moving Objects Tracking Method Based on a Combination Of Local Binary Pattern Texture and Hue”, Procedia Engineering 15 (2011) 3964 – 3968, china.

[23] S.Selvarajah and S.R. Kodituwakku ,” Analysis and Comparison of Texture Features for Content Based Image Retrieval”, International Journal of Latest Trends in Computing (E-ISSN: 2045-5364) 108 Volume 2, Issue 1, March 2011.

[24] David Exner, Erich Bruns, Daniel Kurz, and Anselm Grundhöfer , Germany,” Fast and Robust CAMShift Tracking

[25] Zhenhua Guo, Lei Zhang, Member, IEEE, and David Zhang,” A Completed Modeling of Local Binary Pattern Operator for Texture Classification”IEEE transaction.

[26] B.Z. de Villiers, W.A. Clarke and P.E. Robinson,” Mean Shift Object Tracking with Occlusion Handling”.

[27] Valtteri Takala and Matti Pietik¨ainen,” Multi-Object Tracking Using Color, Texture and Motion”,Machine Vision Group.

[28] John G. Allen, Richard Y. D. Xu, Jesse S. Jin,” Object Tracking Using CamShift Algorithm and Multiple Quantized Feature Spaces”, School of Information Technologies University of Sydney.

[29] Fahad Fazal Elahi Guraya, Pierre-Yves Bayle and Faouzi Alaya Cheikh,” People Tracking via a Modified CAMSHIFT Algorithm”.

[30] Jifeng Ning, Lei Zhang, David Zhang and Chengke Wu,” Scale and Orientation Adaptive Mean Shift Tracking”.

[31] Yea-Shuan Huang, Zhi-Hong Ou, Hsiang-Wen Hsieh and Hung-Hsiu Yu,” A Rapid Texture-Based Moving Object Detection Method”, Proceedings of 2011 8th Asian Control Conference (ASCC) Kaohsiung, Taiwan, May 15-18, 2011.

[32] Michael Boyle,” The Effects of Capture Conditions on the CAMSHIFT Face Tracker”.

[33] A. Karahaliou et al., “ATexture Analysis Approach for Characterizing Micro calcifications on Mammograms”, 2006.

[34] G. N. Srinivasan, and G. Shobha, “Statistical Texture Analysis”, Proceedings of world academy of science, Vol. 36, 2008.

[35] Svetlana Lazebnik, Cordelia Schmid, Jean Ponce, “Affine-Invariant Local Descriptors and Neighborhood Statistics for Texture Recognition”.

[36] Shahera Hossain and Seiichi Serikawa,” Features for Texture Analysis”, SICE Annual Conference 2012 Akita University, Akita, Japan.

[37] Galloway M. M. (1975), Texture analysis using gray level run lengths, Computer Graphics and Image Processing, Volume 4(2), Pages 172-179

[38] Dengsheng Zhang, Aylwin Wong, Maria Indrawan, Guojun Lu, Content-based Image Retrieval Using Gabor Texture Features.

[39] R. Venkatesh Babu , Patrick Pe´rez , Patrick Bouthemy,” Robust tracking with motion estimation and local Kernel-based color modeling”, Image and Vision Computing 25 (2007) 1205–1216 .

[40] T. Ojala, M. Pietikäinen, D. Harwood,”A comparative study of texture measures with classification based on feature distribution”Pattern Recognition, 29 (1996), pp. 51–59.

[41] T. Ojala, M. Pietikäinen, T. Maenpaa,”Multiresolution gray-scale and rotation invariant texture classification with local binary patterns”,IEEE Trans. Pattern Anal. Mach. Intell., 24 (7) (2002), pp. 971–987.

[42] Xiao, L., Li, P.: Improvement on mean shift based tracking using second-order information. In: Proc. IEEE 19th International Conference on Pattern Recognition, pp. 1–4 (2008).

[43] Xiang, Z., Dai, Y.M., Chen, Z.W., Zhang, H.X.: An improved Mean Shift tracking algorithm based on color and texture feature. In: Proc. Wavelet Analysis and Pattern Recognition, pp. 38–43 (2010).

[44] Hongwei Ying, Xuena Qiu, Jiatao Song, Xiaobo Ren," Particle Filtering Object Tracking Based on Texture and Color", 2010 International Symposium on Intelligence Information Processing and Trusted Computing.

[45] Martin Armstrong Andrew Zisserman," Robust Object Tracking", in 1995.

[46] Martin Godec Peter M. Roth Horst Bischof ," Hough-based Tracking of Non-Rigid Objects", In Proc. Int. Conf. on Computer Vision, in 2011.

[47] J´ er ˆome Berclaz, Franc¸ois Fleuret, Engin T¨ uretken, and Pascal Fua ,'' Multiple Object Tracking using K-Shortest Paths Optimization", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2009.

[48] Mei Han Amit Sethiy Yihong Gong," A Detection-Based Multiple Object Tracking Method"2004.

[49] Vasilis Papadourakis , Antonis Argyros," Multiple objects tracking in the presence of long-term occlusions", Elsevier journal, Computer Vision and Image Understanding vol. 114, page no. 835–846, 2010.

[50] Mahbub murshed, Md. Hasanul kabir and oksam chae," moving object tracking an edge segment based approach", international journal of innovative computing, information and control, vol. 7, no. 7(A),2011.

[51] Hieu T. Nguyen and Arnold W.M. Smeulders, Member, IEEE," Fast Occluded Object Tracking by Robust Appearance Filter ", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 26, NO. 8, AUGUST 2004.

[52] Mei Han, Wei Xu Hai Tao Yihong Gong," An Algorithm for Multiple Object Trajectory Tracking", 0-7695-2158-4/04 2004 IEEE

[53] Yizheng Cai, Nando de Freitas, and James J. Little," Robust Visual Tracking for Multiple Targets", ECCV 2006, Part IV, LNCS 3954, pp. 107–118, 2006.

[54] Paul Viola Michael Jones,” Rapid Object Detection using a Boosted Cascade of Simple Features”, ACCEPTED CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2001.

[55] Paul Viola Michael Jones ,“Robust Real-Time Face Detection”, International Journal of Computer Vision 57(2), 137–154, 2004.

Appendix

Abbreviations

1) LBP – LOCAL BINARY PATTERN
2) CAMSHIFT – CONTINOUSELY ADAPTIVE MEAN SHIFT
3) HSV – HUE SATURATION VALUE
4) RGB - Red Green Blue
5) ICA – INDEPENDENT COMPONENT ANALYSIS
6) SDAC – SERVER DATA ACCESS COMPONENT
7) CCTV – CLOSED CIRCUIT TELEVISION
8) GLCM – GREY LEVEL CO-OCCURRENCE MATRIX
9) GLRLM – GREY LEVEL RUN LENGTH MATRIX