INTELLIGENT VIDEO SURVEILLANCE OF HUMAN MOTION: ANOMALY DETECTION

By

NORMALA BINTI TAJUDIN

FINAL PROJECT REPORT

Submitted to the Electrical & Electronics Engineering Programme
in Partial Fulfillment of the Requirements
for the Degree
Bachelor of Engineering (Hons)
(Electrical & Electronics Engineering)

Universiti Teknologi Petronas
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

INTELLIGENT VIDEO SURVEILLANCE OF HUMAN MOTION: ANOMALY DETECTION

by

Normala Binti Tajudin

A project dissertation submitted to the
Electrical & Electronics Engineering Programme
Universiti Teknologi PETRONAS
in partial fulfilment of the requirement for the
Bachelor of Engineering (Hons)
(Electrical & Electronics Engineering)

Approved:

Ms Zazilah May
Project Supervisor

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK

May 2011

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

Normala Binti Tajudin

ABSTRACT

Intelligent video surveillance is a system that can highlight extraction and video summarization that require recognition of the activities occurring in the video without any human supervision. Surveillance systems are extremely helpful to guard or protect you from any dangerous condition. In this project, we propose a system that can track and detect abnormal behavior in indoor environment. By concentrating on inside house environment, we want to detect any abnormal behavior between adult and toddler to avoid abusing to happen. In general, the frameworks of a video surveillance system include the following stages: background estimator, segmentation, detection, tracking, behavior understanding and description. We use training behavior profile to collect the description and generate statistically behavior to perform anomaly detection later. We begin with modeling the simplest actions like: stomping, slapping, kicking, pointed sharp or blunt object that do not require sophisticated modeling. A method to model actions with more complex dynamic are then discussed. The results of the system manage to track adult figure, toddler figure and harm object as third subject. With this system, it can bring attention of human personnel security. For future work, we recommend to continue design methods for higher level representation of complex activities to do the matching anomaly detection with real-time video surveillance. We also propose the system to embed with hardware solution for triggered the matching detection as output.

ACKNOWLEDGEMENT

Before saying anything I would like to thank God for His grace and Mercy that He gave me to get through this project. I would like to take this opportunity to thank all parties who has contributed their effort for the completion of this project.

With my sincere appreciation, I would like to express my utmost gratitude to my supervisor, Ms Zazilah, Lecturer of Electricals and Electronics Engineering Department. Her details and constructive comment along with continuous help, support, motivation and guidance from initial to the final level of the project, I am thankful for her limitless effort on me.

I would like to express my deepest gratitude to all staff of Electrical and Electronics Engineering Department at Universiti Teknologi PETRONAS, especially Final Year Project I and II coordinators and Computer Vision and Image Processing graduate assistants. I am thankful to them for their kind cooperation's and help.

The greatest appreciation goes to my family who did everything for me from the moment I was born to the present. I would not be a person who I have become today without their help and positive attitude towards me. Furthermore, I would like to thank my friends for their comments, help and positive criticism.

Thank you.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLE

LIST OF ABBREVIATIONS

2D	Two Dimensions
3D	Three Dimensions
AI	Artificial Intelligent
CCD	Charge Coupled Device
CCTV	Close Circuit Television
DSP	Digital Signal Processing
EFSA	Extended Finite State Automate
FSM	Finite State Machine
FYP	Final Year Project
IDS	Intrusion Detection System
PC	Personal Computer
PCA	Principle Component Analysis
SVM	Support Vector Machine

CHAPTER 1
INTRODUCTION

1.1 Background of Studies

In recent years, this has been considerable interest in visual surveillance of a wide range of indoor and outdoor sites by researcher and organization. This is manifested by the widespread and unabated deployment of CCTV cameras in public and private areas. In particular, the increasing connectivity of broadband wired and wireless IP networks, and the emergence of CCTV systems with smart sensors, enabling centralised or distributed remote monitoring, have further fuelled this trend.

Society needs for surveillance is going to be more and more intense. Parents nowadays are desperately able to do anything in order to protect their children from harmful condition. Setting up a spy camera like CCTV could well be the best way for parents to keep watch on the nanny and the children under their care [1]. Surveillance cameras can be far more useful tool if instead of passively recording footage, they can be used to detect events requiring attention as they happen and take action in real time. The system should be able to support real-time recording by detect anomalies and translate the output into communication medium to alert the parent as soon as possible.

The word '*nanny*' refer to the Oxford Dictionary [2] mean a person, typically a woman, employed to look after a child in its own home; a person or institution regarded as interfering and overprotective. The word meaning itself clearly states that her duty is to protect the employer's child safety.

1.2 Problem Statement

There were 16 cases of nanny torturing the children of employers reported since 2006 until 2009 [3]. Many busy working parents have to hire other in home workers to perform certain tasks for them out of sheer necessity. Common needs include; house cleaners, pet sitters, and repair workers. Each time they allow one of these workers into the home while they are not there, they have to trust that the workers have the employer best interest at heart. If the parents already have a hidden nanny surveillance system in place, they sure have a very practical way of keeping tabs over what the nanny are actually doing inside. If the parents can caught red handed their nanny, there would be difficult to charges the nanny without any evidence because she would claim the parent was abusing her too [4]. Even though alert parent set up hidden camera, it would be late to take action if their children had suffered from nanny abuse. Abused child would never be the same next day, or worst, their whole life.

In monitoring a visual scene that is cluttered and busy, the importance of detection and tracking of any number of moving objects of interest can never be overestimated. This is the central element of an object based intelligent video surveillance system, of which the two types of application are:

1) To allow real-time detection of unforeseen events that warrants the attention of security guards or law enforcement officers to take preventive actions,
2) To enable tagging and indexing of interesting (customer-defined) scene activities/statistics into the database for rapid forensic analysis.

In addition, object detection and tracking are the building blocks of higher-level vision-based or assisted event monitoring and management systems with a view to understanding the complex actions, interactions, and abnormal behaviours of objects in the scene. The range of applications includes detection of criminal behaviours in banks, marketing data analysis in shopping malls, and well-being monitoring at home.

1.3 Objectives and Scope of Studies

1.3.1 Objective

The objectives of the system are:

1. To record video with basic background (random and stationary background) for image acquisition with desire abnormal behaviour collection.
2. Be able to run motion segmentation to remove the background or separate the motion of the subject.
3. To collect detection block as individual representation and then reducing the noise like occlusion to happen.
4. Be able to running tracking box during the whole video surveillance play.

1.3.2 Scope of Studies

- MATLAB: Video & Image Processing; and Signal System
  - Image acquisition
  - Motion segmentation: Blob detection
  - Object classification
  - Tracking
  - Behavior recognition
  - High level Processing
  - Anomaly Detection Method*
  - Event matching
  - Transmitting online

CHAPTER 2
LITERATURE REVIEW

2.1 Image Acquisition

Video surveillance cameras are becoming more popular and, most importantly, more accessible to the average person. Huge store chains or gated communities are not the only places than can afford the benefits of surveillance cameras anymore. They are available for the home and office needs, no matter how simple or how elaborate a setup it requires.

Surveillance cameras are extremely helpful in protecting home or office environment. The parent can check up on the nanny or employers, make sure nothing suspect happens at home while they are away, monitor the children from another part of the house. Video cameras designed specifically for surveillance give an extra measure of safety and security.

2.1.1 Types of cameras available for video surveillance

There are a broad range of video surveillance cameras available nowadays. Understanding the different kinds will make it easier to determine which one suits to protect the best [5].

- **Fake security cameras** – These are not actual cameras. They are a very inexpensive alternative to purchasing an actual system. While these cameras can act as a deterrent, should something happen, there will not have a record of it.

- **Covert surveillance cameras** – These cameras look like regular items. A wall clock in a small store, a teddy bear in a baby's room, a potted plant by the front door—each one of these could very easily be a surveillance camera. Where we can record without anyone knowing it.
- **Wireless security cameras** – These offer more flexibility in set up. They are easy to install, can be moved easily, are often small, have no tell-tale wires, and are very discreet.
- **Wired surveillance cameras** – These cameras are appropriate for permanent setup. If the family has one location that they would like to constantly monitor and will not need it to be changed, like the family living room, a wired camera is an option. Some may have to be professionally installed.
- **Night vision security cameras** – These are ideal for any low-light areas, not just for the night time. The camera records in black and white in order to capture images best.
- **Home surveillance cameras** – When we install a setup for home, there will be a helpful system that often includes such bonuses as timers for the home lamps and motion sensors to go with the wireless camera.

2.2 Behavioural Analysis

The problem of using vision to track and understand the behaviour of human beings is a very important one. It has applications in the areas of human-computer interaction, user interface design, robot learning, and surveillance, among others. At its highest level, this problem addresses recognizing human behaviour and understanding intent and motive from observations alone. This is a difficult task, even for humans to perform, and misinterpretations are common. In the area of surveillance, automated systems to observe pedestrian traffic areas and detect dangerous action are becoming important.

Many such areas currently have surveillance cameras in place, however, all of the image understanding and risk detection is left to human security personnel. This type of observation task is not well suited to humans, as it requires careful concentration over long periods of time. Therefore, there is clear motivation to

develop automated intelligent vision-based monitoring systems that can aid a human user in the process of risk detection and analysis.

The capability of being able to analyze human movements and their activities from image sequences is crucial for visual surveillance. In general, the framework of a video surveillance system includes the following stages: modelling of environments, detection of motion, classification, tracking, and behaviour understanding and description. One of most difficult challenges in the domain of computer vision is semantic behaviour learning and understanding from observing activities in video surveillance. The research in this area concentrates mainly on the development of methods for analysis of visual data in order to extract and process information about the behaviour of physical objects (e.g., humans) in a scene.

From the security point of view, there were three different kinds of behaviours: normal, unusual and abnormal. Typically, in an area under surveillance in this project, the focus of attention falls into abnormal behavioural especially child abusing type. Child abuse is more than bruises and broken bones. While physical abuse might be the most visible sign, other types of abuse, such as emotional abuse or child neglect, also leave deep, long lasting scars. Some signs of child abuse are subtler than others. However, by learning common types of abuse and what to do next, it can make a huge difference in a child's life. The earlier abused children get help, the greater chance they have to heal from their abuse and not perpetuate the cycle.

2.2.1 Types of child abuse

There are several types of child abuse, but the core element that ties them together is the emotional effect on the child. Children need predictability, structure, clear boundaries, and the knowledge that their parents are looking out for their safety. Abused children cannot predict how their parents will act. Their world is an unpredictable, frightening place with no rules. Whether the abuse is a slap, a harsh comment, stony silence, or not knowing if there will be dinner on the table tonight, the end result is a child that feel unsafe, uncared for, and alone.

An abused or neglected child is a child who is harmed or threatened with physical or mental harm by the acts or lack of action by a person responsible for the child's care. Each state has its own laws concerning child abuse and neglect. There are several forms of abuse: physical abuse, emotional abuse, and sexual abuse. Child neglect is a form of abuse that occurs when a person responsible for the care of a child is able but fails to provide necessary food, clothing, shelter, or care. A brief discussion of each form of abuse follows [6].

2.2.1.1 Neglect

A child is neglected when the persons he depends on do not provide food, clothing, shelter, medical care, education, and supervision. When these basic needs are deliberately withheld, not because the parents or caregivers are poor, it is considered neglect. Often the parents or caregivers of neglected children are so overwhelmed by their own needs that they cannot recognize the needs of their children.

2.2.1.2 Sexual Abuse

When an adult or older child uses his or her authority to involve the child in sexual activity, it is child sexual abuse, and that person is a child molester. The molester might use tricks, bribes, pressure, threats, or force to persuade the child to join in sexual activity. Sexual abuse includes **any** activity performed for the sexual satisfaction of the molester. A common misconception about sexual abuse is that children are most likely to be molested by strangers when the fact is that a child molester is usually someone that the child knows and trusts. Child molesters are most often male, but females perform about one-fifth of the sexual abuse of boys under the age of 14.

2.2.1.3 Emotional Abuse

Emotional abuse is harder to recognize but is just as harmful to the child as other forms of abuse. Emotional abuse damages the child's self-esteem and, in extreme cases, can cause developmental problems and speech disorders. A child suffers from emotional abuse when constantly ridiculed, rejected, blamed, or compared unfavourably with brothers, sisters, or other children.

Unrealistic expectations in academic or athletic achievement are a common cause of emotional abuse by parents or other adults. When a child can't meet these expectations, he feels that he is never quite good enough. Emotional abuse is almost always present when other forms of abuse are identified.

2.2.1.4 Physical Abuse

Physical abuse is the deliberate injury of a child by a person responsible for the child's care. Physical abuse is often the result of unreasonable punishment, or punishment that is too harsh for the child. Sometimes, physical abuse is caused when caregivers react to stress. Drinking and drug abuse by caregivers are often contributing factors to physical abuse. Physical abuse injuries can include bruises, broken bones, burns, and abrasions. Children experience minor injuries as a normal part of childhood, usually in places such as the shins, knees, and elbows. When the injuries are found in the soft-tissue areas on the abdomen or back, or don't seem to be typical childhood injuries, it is possible that the child has been abused.

In physical abuse, unlike physical forms of discipline, the following elements are present:

- **Unpredictability.** The child never knows what is going to set the parent off. There are no clear boundaries or rules. The child is constantly walking on eggshells, never sure what behavior will trigger a physical assault.
- **Lashing out in anger.** Physically abusive parents act out of anger and the desire to assert control, not the motivation to lovingly teach the child. The angrier the parent, the more intense the abuse.
- **Using fear to control behavior.** Parents who are physically abusive may believe that their children need to fear them in order to behave, so they use physical abuse to "keep their child in line." However, what children are really learning is how to avoid being hit, not how to behave or grow as individuals.

Hence, there is important to define complex contexts of scenes or activities in details. The effectiveness of a behaviour profiling algorithm shall be measured by; how well anomalies can be detected and how accurately and robustly different classes of normal behaviour pattern can be recognizing.

2.3 Motion Segmentation

Segmentation of objects in image sequences is a crucial task for a wide variety of multimedia applications such as video object coding, manipulation and identification. The ideal goal of segmentation is to identify the semantically meaningful components of an image and to group the pixels belonging to such components. While it is very hard to segment static objects in images, it is easier to segment moving objects in video sequences. Once the moving objects are correctly detected and extracted, they can serve for a variety of purposes.

Subjective segmentation evaluation is necessary to study and to characterize the perception of different artefacts on the overall quality for different applications, but once this task has been accomplished successfully and an automatic procedure has been devised, systematic subjective evaluation can be avoided. In order to compare the performance of two objective metrics, in this paper, the subjective opinions are compared to their objective results.

Video tracking is the process of locating a moving object (or several ones) in time using a camera. An algorithm analyzes the video frames and outputs the location of objects as a function of time, optionally in real time. It is mainly used in video surveillance systems. The issues involved in video object tracking are different from those of video object segmentation evaluation since the ground truth on which these algorithms compare their performance is different. In fact, video surveillance systems concern algorithms for detecting, indexing and tracking moving objects and this evaluation requires specific considerations as follows. Most of segmentation methods use either temporal or spatial information in the image sequence. Several widely used approaches for motion segmentation include [7]:

2.3.1 Background subtraction

Background subtraction is very popular for applications with relatively static backgrounds as it attempts to detect moving regions in an image by taking the difference between the current image and the reference background image in a pixel by pixel fashion. However, it is extremely sensitive to changes of environment lighting and extraneous events. The numerous approaches to this problem differ in the type of background model and the procedure used to update the background model. The estimated background could be simply modelled using just the previous frame; however, this would not work too well.

The background model at each pixel location could be based on the pixel's recent history. Background subtraction methods store an estimate of the static scene, accumulated over a period of observation; this background model is used to find foreground (i.e., moving objects) regions that do not match the static scene. Recently, some statistical methods to extract change regions from the background are inspired by the basic background subtraction methods as described above. The statistical approaches use the characteristics of individual pixels or groups of pixels to construct more advanced background models, and the statistics of the backgrounds can be updated dynamically during processing. Each pixel in the current image can be classified into foreground or background by comparing the statistics of the current background model. This approach is becoming increasingly popular due to its robustness to noise, shadow, changing of lighting conditions, etc.

2.3.2 Temporal Differencing

In temporal differencing, video frames are separated by a constant time and compared to find regions that have changed. Unlike background subtraction, temporal differencing is based on local events with respect to time and does not use a model of the background to separate motion. Typically, two or three frames are used as separation time intervals, depending on the approach. A small time interval provides robustness to lighting conditions and complex backgrounds, since illumination changes and objects in the scene are more likely to be similar over short

periods of time. However, an image-stabilization algorithm is required when there is a significant movement of the camera.

Temporal differencing makes use of the pixel-wise difference between two to three consecutive frames in an image sequence to extract moving regions. Temporal differencing is very fast and adaptive to dynamic environments, but generally does a poor job of extracting all the relevant pixels, for example there may be holes left inside moving entities. This method also usually computationally inexpensive, but it regularly fail at properly extracting the shape of the object in motion and can cause small holes to appear. For these reasons, hybrid approaches often combine both background subtraction and temporal differencing methods to provide more robust segmentation strategies.

2.3.3 Optical Flow

Optical flow is the velocity field which warps one image into another (usually very similar) image, and is generally used to describe motion of point or feature between images. Optical flow methods are very common for assessing motion from a set of images. It also apply as a vector-based approach that estimates motion in video by matching points on objects over multiple frames. A moderately high frame rate is required for accurate measurements. It should be noted that a real-time implementation of optical flow will often require specialized hardware, due to the complexity of the algorithm. A benefit of using optical flow is that it is robust to multiple and simultaneous camera and object motions, making it ideal for crowd analysis and conditions that contain dense motion. However, most optical flow methods are computationally complex, sensitive to noise, and would require specialized hardware for real-time applications.

2.4 Object Tracking

After finding moving regions or objects in an image, the next step in the behaviour recognition process is object classification [8]. For example, a pedestrian crossing a street and a vehicle running a red light can be similar if there is no knowledge of the object causing the motion. Furthermore, object classification could

distinguish interesting motion from those caused by moving chair, mirror reflections, or other dynamic occurrences common happen in home recorded into transit videos. It is important to note here that there are multiple possible representations of objects before and after classification. Common geometric or topological properties used include height/width ratio, fill ratio, perimeter, area, compactness, convex hull, and histogram projection. Some of these properties are also used in post object classification to keep track of the object in sequential frames or separate cameras.

The object tracking module is responsible for the detection and tracking of moving objects from individual cameras; object locations are subsequently transformed into 3D world coordinates. The camera handoff and data fusion module (or algorithm) then determines single world measurements from the multiple observations. Object tracking can be described as a correspondence problem and involves finding which object in a video frame related to which object in next frame. Normally, the time interval between two successive frames is small, thus the inter-frame changes are limited, allowing the use of temporal constraints and/or object features to simplify the correspondence problem.

Tracking methods [9] can be roughly divided into four major categories, and algorithms from different categories can be integrated together. In the context of transit systems, tracking is defined as the problem of estimating the trajectory of a pedestrian in the image plane while he is in the transit station or vehicle. The increasing need for automated video analysis has motivated researchers to explore tracking techniques, particularly for surveillance applications.

2.5 Anomaly Detection

In anomaly detection, the difficulty is to detect activity, behaviour, objects, or substances that are atypical. Typical is defined with respect to historical data and is extremely scenario dependent. Algorithms for anomaly detection must adjust to the scenario and be robust to a vast range of possible assumptions. As a result, there is typically no model for an anomaly and the model for the location and time are derived from observations. Scenarios that need anomaly detection include perimeter, border, or gateway surveillance.

2.5.1 Anomaly Detection using Statistics

In statistical methods for anomaly detection, the system observes the activity of subjects and generates profiles to represent their behaviour. Typically, two profiles are maintained for each subject: the current profile and the stored profile. As the network events are processed, the system updates the current profile and periodically calculates an anomaly score by comparing the current profile with the stored profile using a function of abnormality of all measures within the profile. If the anomaly score is higher than a certain threshold, the system generates an alert.

Statistical anomaly detection has a number of advantages. Firstly, these systems do not require prior knowledge of security flaws and/or the attacks themselves. In addition, statistical approaches can provide accurate notification of malicious activities that typically occur over extended periods of time. However, statistical anomaly detection schemes also have drawbacks. Firstly, it can be difficult to determine thresholds that balance the likelihood of false positives with the likelihood of false negatives. In addition, statistical methods need accurate statistical distributions, but not all behaviours can be modelled using purely statistical methods.

Haystack [10] is one of the earliest examples of a statistical anomaly-based intrusion detection system. It used both user and group-based anomaly detection strategies, and modelled system parameters as independent, Gaussian random variables. Haystack defined a range of values that were considered normal for each feature. If during a session, a feature fell outside the normal range, the score for the subject was raised. It was designed to detect six types of intrusions. But, one drawback of Haystack was that it was designed to work offline.

2.5.2 Anomaly Detection using Classifier

Anomaly detection depends on the idea that normal characteristics behaviour can be distinguished from abnormal behaviour. A classifier can be used to predict the normal incoming event given the current event. If during the monitoring phase the

next event is not the one predicted by the classifier, it is considered as an anomaly. The classification process typically involves the following steps:

1. Identify class attributes and classes from training data.
2. Identify attributes for classification.
3. Learn a model using the training data.
4. Use the learned model to classify the unknown data samples.

A variety of classification techniques have been proposed in the literature. These include inductive rule generation techniques, fuzzy logic and genetic algorithms-based techniques.

Inductive rule generation algorithms typically involve the application of a set of association rules and frequent episode patterns to classify the audit data. The advantage of using rules is that they tend to be simple and intuitive, unstructured and less rigid. As the drawbacks they are difficult to maintain, and in some cases, are inadequate to represent many types of information. A number of inductive rule generation algorithms have been proposed in literature. Some of them first construct a decision tree and then extract a set of classification rules from the decision tree. Other algorithms directly induce rules from the data by employing a divide-and-conquer approach.

Fuzzy logic techniques have been in use in the area of network security since the late 1990's. Dickerson et al. [11] developed the Fuzzy Intrusion Recognition Engine (FIRE) using fuzzy sets and fuzzy rules. FIRE uses simple data mining techniques to process the network input data and generate fuzzy sets for every observed feature. The fuzzy sets are then used to define fuzzy rules to detect individual attacks. FIRE does not establish any sort of model representing the current state of the system, but instead relies on attack specific rules for detection.

Genetic algorithms, a search technique used to find approximate solutions to optimization and search problems, have also been extensively employed in the domain of intrusion detection to differentiate normal network traffic from anomalous connections. The major advantage of genetic algorithms is their flexibility and

robustness as a global search method. The earliest attempt to apply genetic algorithms to the problem of intrusion detection was done by Crosbie and Spafford [12] in 1995, when they applied multiple agent technology to detect network based anomalies.

2.5.3 Anomaly Detection using Machine learning

Machine learning aims to answer many of the same questions as statistics. However, unlike statistical approaches which tend to focus on understanding the process that generated the data, machine learning techniques focus on building a system that improves its performance based on previous results. In other words systems that are based on the machine learning paradigm have the ability to change their execution strategy on the basis of newly acquired information.

A Bayesian network is a graphical model that encodes probabilistic relationships among variables of interest. When used in conjunction with statistical techniques, Bayesian networks have several advantages for data analysis. Several researchers have adapted ideas from Bayesian statistics to create models for anomaly detection. Valdes et al. [13] developed an anomaly detection system that employed naive Bayesian networks to perform intrusion detection on traffic bursts.

Bayesian techniques also have been frequently used in classification and suppression of false alarms areas. Kruegel et al. [14] proposed a multi-senor fusion approach where the outputs of different IDS sensors were aggregated to produce a single alarm. This approach is based on the assumption that any anomaly detection technique cannot classify a set of events as an intrusion with sufficient confidence. Although using Bayesian networks for intrusion detection can be effective in certain applications, their limitations should be considered in the actual implementation. Since the accuracy of this method is dependent on certain assumptions that are typically based on the behavioural model of the target system, therefore, selecting an accurate model is the most important things towards solving the problem. Unfortunately selecting an accurate behavioural model is a difficult task as typical networks are complex.

Typical datasets for intrusion detection are very large and multidimensional. To tackle the problem of high dimensional datasets, researchers have developed a dimensionality reduction technique known as principal component analysis (PCA). PCA is a technique where n correlated random variables are transformed into d<n uncorrelated variables. The uncorrelated variables are linear combinations of the original variables and can be used to express the data in a reduced form. Shyu et al. [15] proposed an anomaly detection scheme, where PCA was used as an outlier detection scheme and was applied to reduce the dimensionality of the audit data and arrive at a classifier that is a function of the principal components.

2.5.4 Anomaly Detection using Finite State Machine

A finite state machine (FSM) is a model of behaviour composed of states, transitions and actions. In this model, state stores information about the past, a transition indicates a state change and is described by a condition that would need to be fulfilled to enable the transition. An action is a description of an activity that is to be performed at a given moment.

The finite state machine has been used to detect attacks on the DSR protocol in [16]. First, an algorithm for monitor selection for distributed monitoring all nodes in networks was proposed and then the correct behaviours of the nodes according to DSR were manually abstracted. Using this method has the advantage of detecting intrusions without the need of trained data or signatures, also unknown intrusions can be detected with few false alarms. As a result, a distributed network monitor architecture which traces data flow on each node by means of finite state machine was proposed. In reference [17], Sekar et al. present a specification-based model as well as a prototype with excellent detection performance. The model proposed by authors consists of developing protocol specifications by using Extended Finite State Automata (EFSA).

CHAPTER 3
METHODOLOGY

3.1 Procedure Identification

The overall project work follows the flow chart as below.

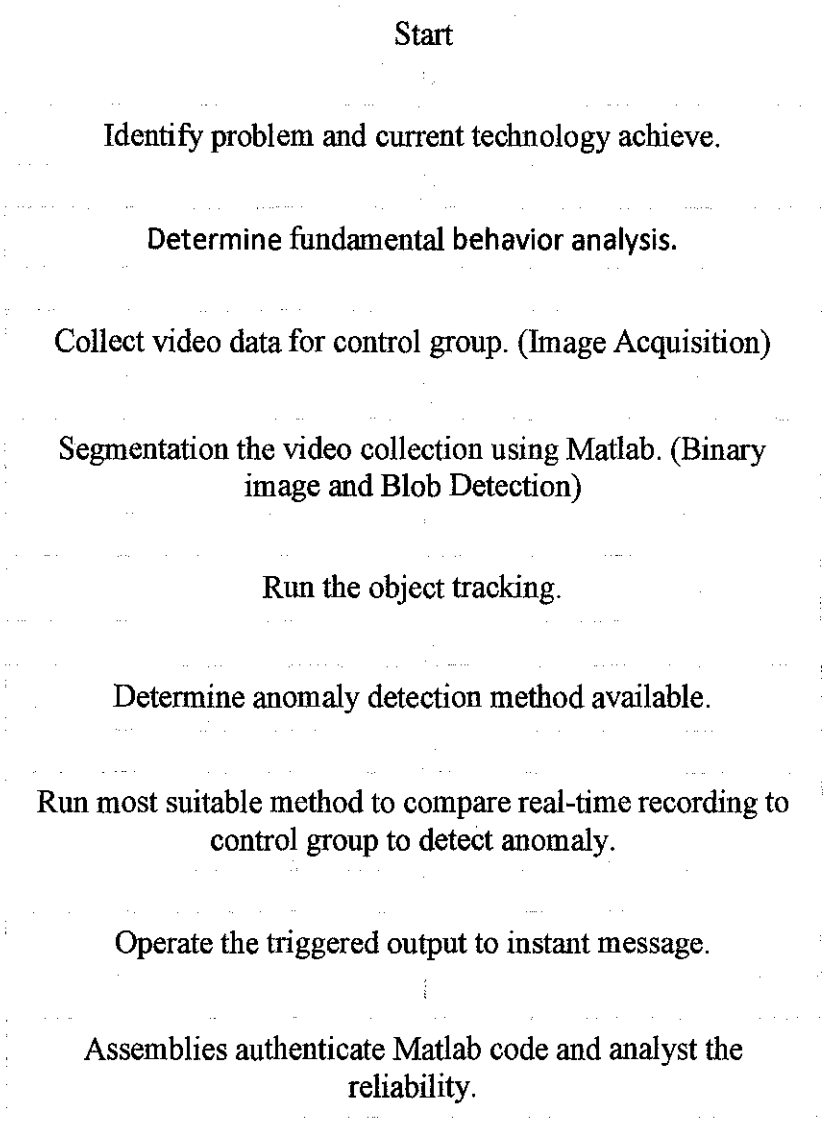


Figure 1: Flow Chart

3.2 Tools and Equipments Required

3.2.1 Hardware

Since the project covered all the software to develop the code, the system does not use any heavy electrical nor electronics hardware tools. In the beginning during collection of data for the control group, we use:

3.2.1.1 Human being for gesture modelling behaviour.

Figure 2: Solid human

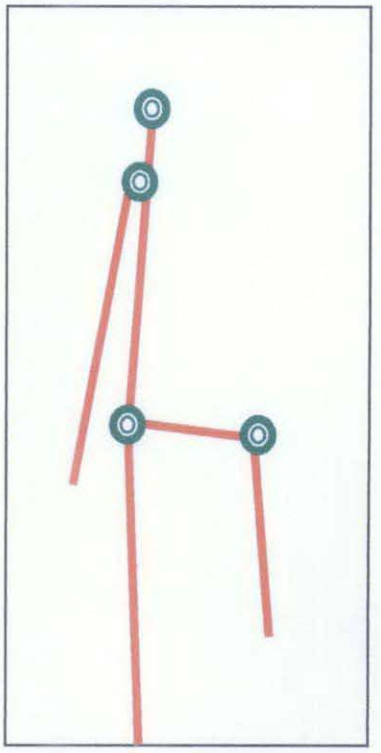


Figure 3: Stick representation

Above Figure 2 and Figure 3 show the example for stick representation for recording the control group video. We had classifies all the representative stick motion for every abusing motion to be in the control group, such as Stomping, Kicking, Slapping, Pointing Hard and Sharp Object. From the control group we had collective data, then when we masking the real-time recording. The control group will match if the abnormal targeted.

3.2.1.2 Teddy Bear for illustrate the child figure.

Figure 4: Child presentation

Figure shown respective model for child figure during control group video. Due to the collection needs, we practice improper motion like abusing, so it is not relevant to include any human being child. Since 90% baby bear figure resemble child figure, we manage to find one looks really like a child figure.

3.2.1.3 Samsung Video Camera HMX:

With Full HD and a1080i mode with 30 frames per second recording capability, with an integrated docking station, it is easy not to charge the battery, but use the HDMI input and component connectivity to easily view the recording movies directly on the PC for direct save control video.

Figure 5: Available Camera Recorder at 23.02.08

3.2.2 Software

Algorithm development is central to image and video processing because each situation is unique, and good solutions require multiple design iterations. By using Matlab software, it provides a comprehensive environment to gain insight into the image and video data, develop algorithms, and explore implementation tradeoffs.

With MathWorks tools for image and video processing:

- We can acquire images and video from imaging hardware
- Use graphical tools to visualize and manipulate images and video
- Develop the model using libraries of reference-standard algorithms
- Migrate the compiled designs to embedded hardware, in planning to the output triggered.

3.2.2.1 Video and Image Processing Blockset

The Video and Image Processing Blockset software is a tool used for the rapid design, prototyping, graphical simulation, and efficient code generation of video processing algorithms. Video and Image Processing Blockset blocks can process images or video data. These blocks can import streaming video into the Simulink environment and perform two-dimensional filtering, geometric and frequency transforms, block processing, motion estimation, edge detection and other signal processing algorithms.

Video and Image Processing Blockset blocks support floating-point, integer, and fixed-point data types. To use any data type other than double-precision and single-precision floating point installing Simulink® Fixed Point™ software is a must. Video processing systems require a stream processing architecture, in which video frames from a continuous stream are processed one (or more) at a time. This type of processing is critical in systems that have live video or where the video data is so large that loading the entire set into the workspace is inefficient. Video and Image Processing Blockset supports a stream processing architecture through System objects (for use in MATLAB) and blocks (for use in Simulink).

3.2.2.2 Design Real-Time Embedded Video Processing Systems

Once we have captured a control group video processing system design in MATLAB code or a Simulink block diagram, we can proceed to evaluate alternatives for impact on performance and fixed-point arithmetic. Next, for generate real-time embedded code from the block diagram and execute the code on a variety of supported target hardware for verification, debugging, and implementation, such that the program we develop can run on the real-time based and hardware. The module that includes optimized video processing libraries can improve the performance of the developed system even further.

Video applications present common but difficult challenges that require flexible analysis and processing functionality. Using MathWorks products, you can:

- Solve frequent pre- and post-processing problems, such as interfering noise, low contrast, out-of-focus optics, and artifacts introduced by interlacing
- Analyze video with methods such as edge detection, blob analysis, template matching, optical flow, and corner detection
- Develop solutions to common video processing challenges such as video stabilization, video mosaicking (MATLAB), target detection, and tracking

CHAPTER 4
RESULTS AND DISCUSSION

4.1 Image Acquisition

The analysis of human behaviour for surveillance, entertainment, intelligent domiciles and medical applications is of increasing interest. Video information is a valuable and easily acquired input for these domains. The research ranges from 2D tracking of human figures by the selected motion for the control group. In this paper an action denotes a short sequence of body configurations (e.g., *leg rising, arm still*). It is usually, but not exclusively, defined by one or a few body parts. An activity denotes a sequence of body configurations over a longer time span. Activities can be assembled from one or more actions, and actions can specify details of an activity (e.g., *walking with raised left arm*).

The discussion covers three areas: 1) motion analysis of the human body structure, 2) tracking human motion without using the body parts from a single view or multiple perspectives, and 3) recognizing human activities from image sequences.

The relationship among these three areas is depicted. Motion analysis of the human body usually involves low-level processing, such as body part segmentation, joint detection and identification, and the recovery of 3D structure from the 2D projections in an image sequence. Tracking moving individuals from a single view or multiple perspectives involves applying visual features to detect the presence of humans directly, i.e., without considering the geometric structure of the body parts. Motion information such as position and velocity, along with intensity values, is employed to establish matching between consecutive frames. Once feature correspondence between successive frames is solved, the next step is to understand the behaviour of these features throughout the image sequence.

After researching for correct motion for nanny abusing motion, here is the list of top human motion for abusing illustrate below:

Action	Video Clip
RANDOM BACKGROUND **Stomping**	
RANDOM BACKGROUND **Slapping**	
RANDOM BACKGROUND **Kicking**	

STATIONARY BACKGROUND Stomping	
STATIONARY BACKGROUND Slapping	
STATIONARY BACKGROUND Kicking	
STATIONARY BACKGROUND Blunt Object Hit	

STATIONARY BACKGROUND **Sharp Object Pointed**	

Table 1: Set of recorded image.

4.2 People Tracking using MATLAB Model

4.2.1 People Tracking Fundamental

This method detects and tracks people in a video sequence with a stationary or random background using the following process:

1) Use the first few frames of the video to estimate the background image. Background subtraction can detect moving regions in an image by taking the difference between the current image and the reference background image in a pixel by pixel fashion. For instance, we have 3 set of random background and 5 set of stationary background.

Random Background		Stationary Background	
Set #1	Stomping	**Set #4**	Stomping
Set #2	Slapping	**Set #5**	Slapping
Set #3	Kicking	**Set #6**	Kicking
		Set #7	Blunt Object Hit
		Set #8	Sharp Object Pointed

Table 2: Set of input video.

Below is the image for background subtraction, which resulting from selection of few second from the beginning of the video recorded. Since random background stating with person inside, it is considered the people inside as background also but it will not disturbs tracking results afterward.

Figure 6: Background subtraction for Random background.

Figure 7: Background subtraction for Stationary background.

2) Separate the pixels that represent the people from the pixels that represent the background. The motion regions are usually detected by subtracting the background from the current images.

Figure 8: Motion pixel differentiated after remove the background.

3) Group pixels that represent individual people together and calculate the appropriate bounding box for each person. Instead of tracking the whole set of pixels comprising an object, the algorithms track only the contour of the object.

Figure 9: Different pixel from each movement bounded together before bounded as a person in one full box.

4) Match the people in the current frame with those in the previous frame by comparing the bounding boxes between frames. Here, we used the idea of Kalman filter which is to model the background over time in conjunction with statistics based on a moving object model.

Figure 10: Input video played without tracked box.

Figure 11: Output video played with tracked box.

4.2.2 Demo Model

Now, we will discuss about how the process flows with the usage of MATLAB. The basic idea develop the tracking method is explained in the following flow chart.

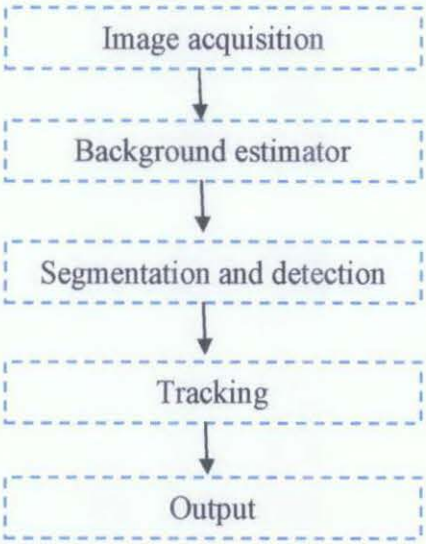


Figure 12: Tracking method flow chart.

Next is the following figure shows the People Tracking demo model:

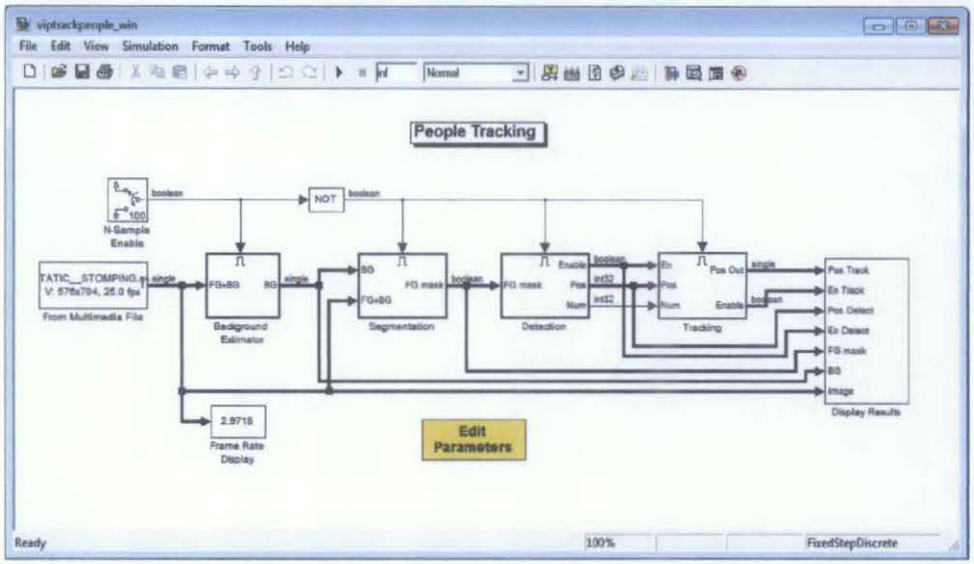


Figure 13: Demo Model block diagram.

Few variable needs to be selected before running the model. We are using video size 576 x 704, with 25.0fts. We need to be sure with the video file size because background subtraction will proceed with same size and sample time input.

4.2.2.1 Background estimator

Background subtraction compares an image with an estimate of the image as if it contained no objects of interest. It extracts foreground objects from regions where there is a significant difference between the observed and the estimated image.

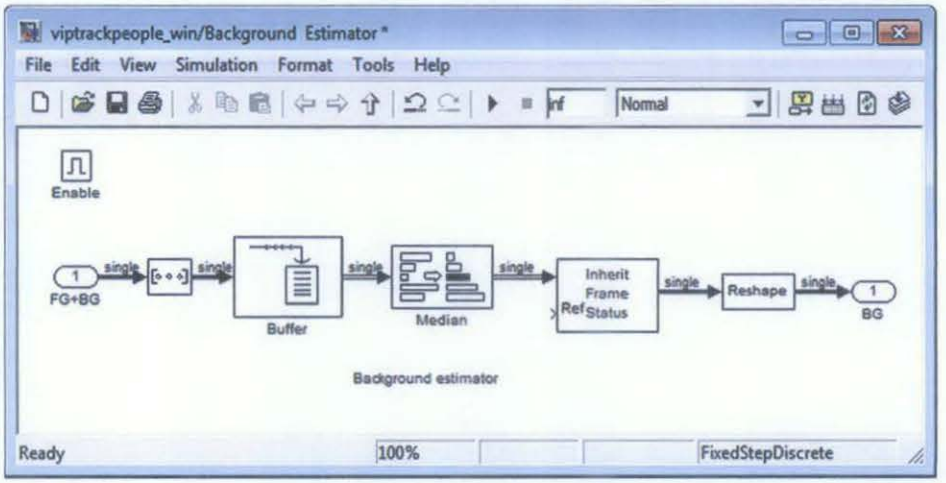


Figure 14: Background estimator block

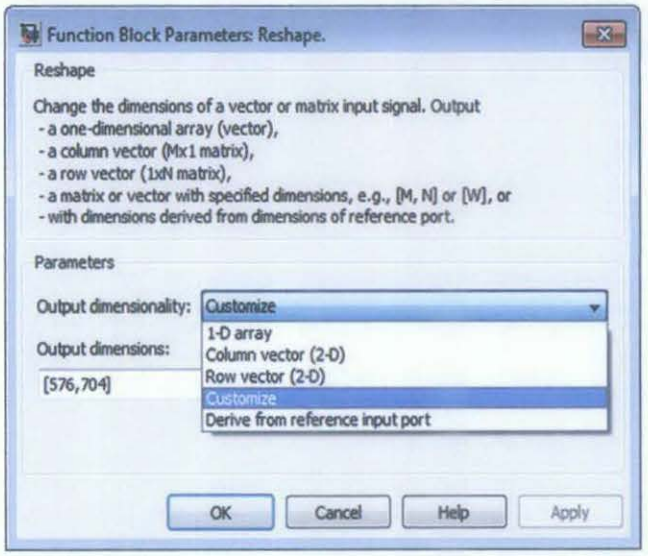


Figure 15: Background estimator dimension setup.

The output of the background subtraction method for each frame is a binary image, which is composed of foreground region. When an occlusion occurs, multiple objects may merge into the same area. This requires an object model that can address split-and-merge cases. Each pixel in the foreground indicates an object label according to which the product of colour and spatial probability is the highest.

4.2.2.2 Segmentation and detection

Most visual surveillance systems start with motion detection. Here, we start with segmentation then detection. In the behaviour detection process consists of foreground segmentation, blob detection, and tracking. Semantic descriptions of suspicious human behaviours are defined through groups of low-level blob based events. For example, abusing is defined as many blobs centroid moving together, merging and splitting, and overall fast changes in the blobs' characteristics.

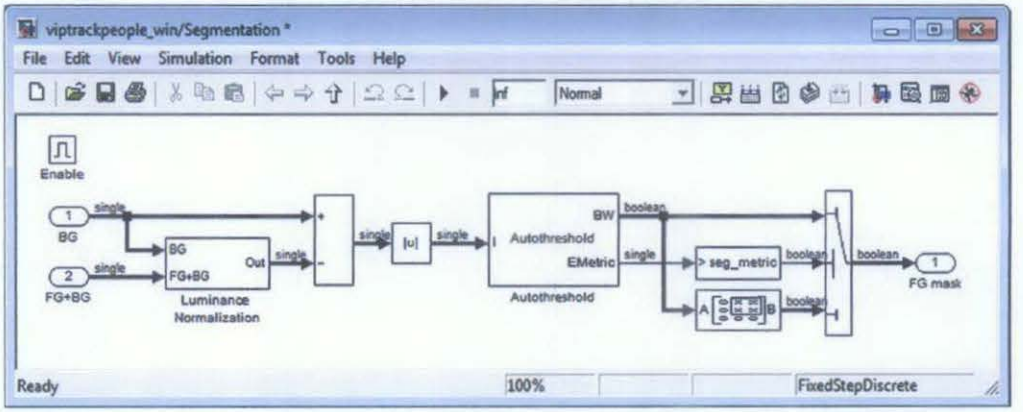


Figure 16: Segmentation block.

In the Segmentation subsystem, the Autothreshold block uses the difference in pixel values between the normalized input image and the background image to determine which pixels correspond to the moving objects in the scene.

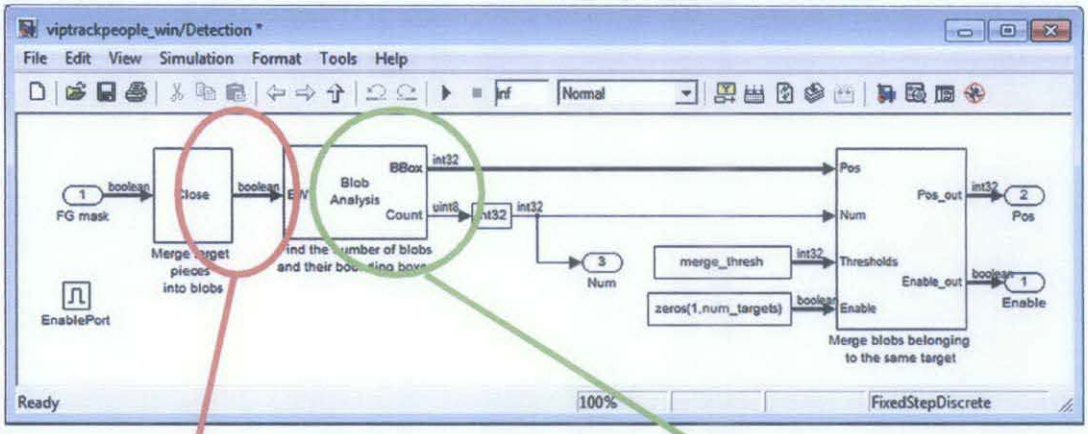


Figure 17: Detection block.

Figure 18: Close block function

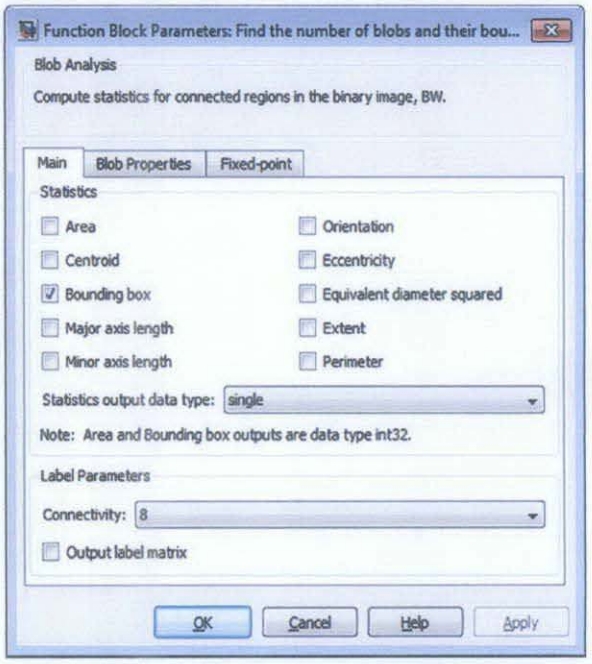


Figure 19: Blob analysis block.

In the Detection subsystem, the Close block merges object pixels that are close to each other to create blobs. For example, pixels that represent a portion of a person's body are grouped together. Next, the Blob Analysis block calculates the bounding boxes of these blobs. In the final step, the Detection subsystem merges the individual bounding boxes so that each person is enclosed by a single bounding box. Motion detection aims at segmenting regions corresponding to moving objects from the rest of an image. The motion and object detection process usually involves environment (background) modelling and motion segmentation. Subsequent processes such as object classification, tracking, and behaviour recognition are greatly dependent on it.

4.2.2.3 *Tracking*

Kalman filter was used to model the background over time in conjunction with statistics based on a moving object model. Commonly, a Kalman filter is used to provide better estimates of future trajectories of objects which move into an occluded zone.

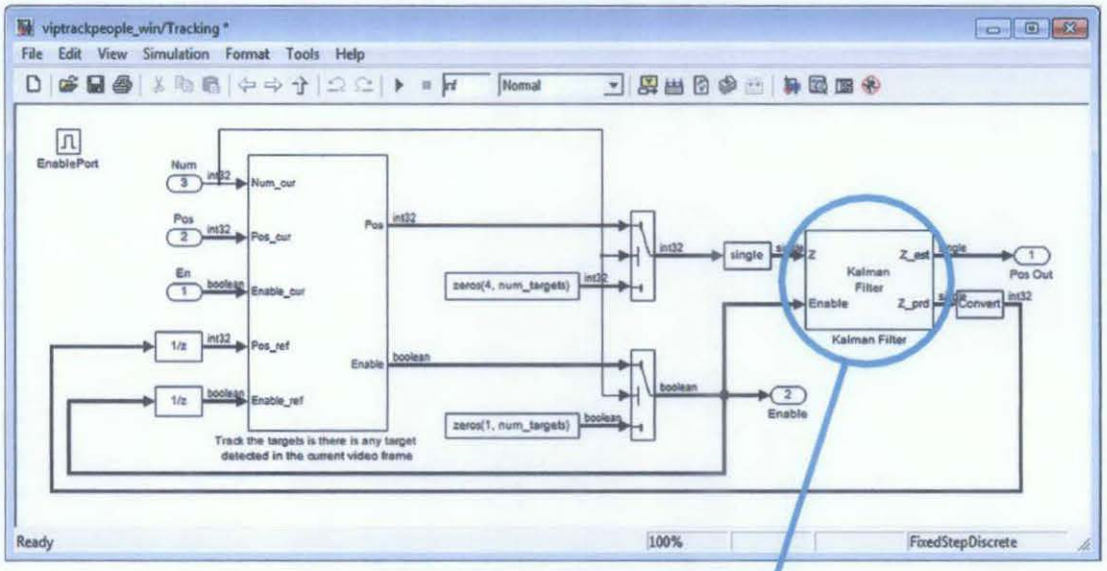


Figure 20: Tracking Block

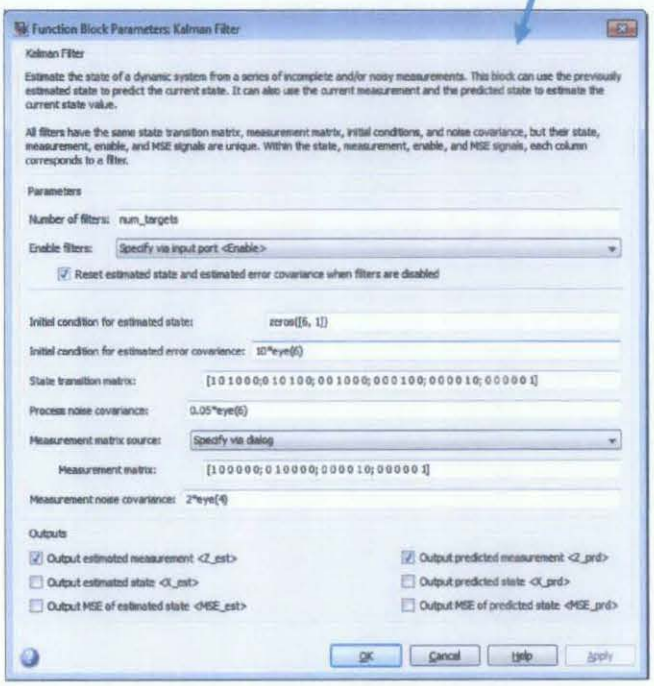
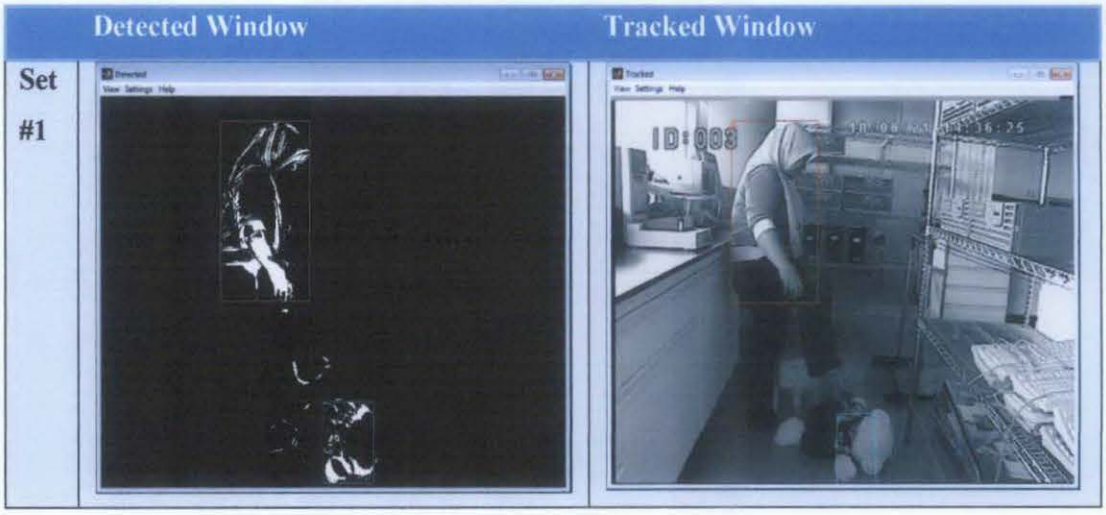


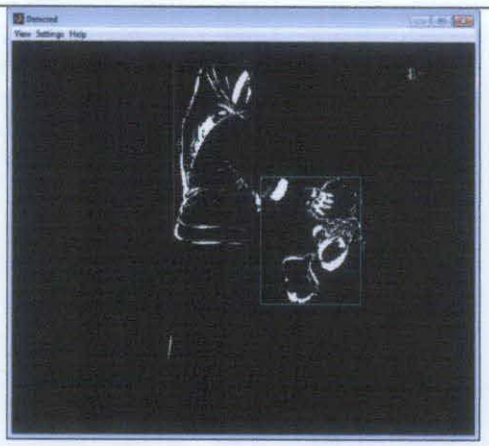
Figure 21: Kalman filter block

In the Tracking subsystem, the Kalman Filter block uses the locations of the bounding boxes detected in the previous frames to predict the locations of these bounding boxes in the current frame. To determine the locations of specific people from one frame to another, the demo compares the predicted locations of the bounding boxes with the detected locations. This enables the demo to assign a unique colour to each person. The demo also uses the Kalman Filter block to reduce the effect of noise in the detection of the bounding box locations.

4.2.3 Demo Model Result

In the Detected window, the people involved in the scene (the maid and the toddler) are surrounded by bounding boxes. The demo assigns each bounding box a colour based on the order that each person is detected. For instance, the maid is the person who come in first detected has a red bounding box and the toddler later detected has a green box. The colour of these boxes changes because the people in the scene are not tracked. In the Tracked window, each person has a unique bounding box colour and label for the duration of the video.



Set #2		
Set #3		
Set #4		
Set #5		

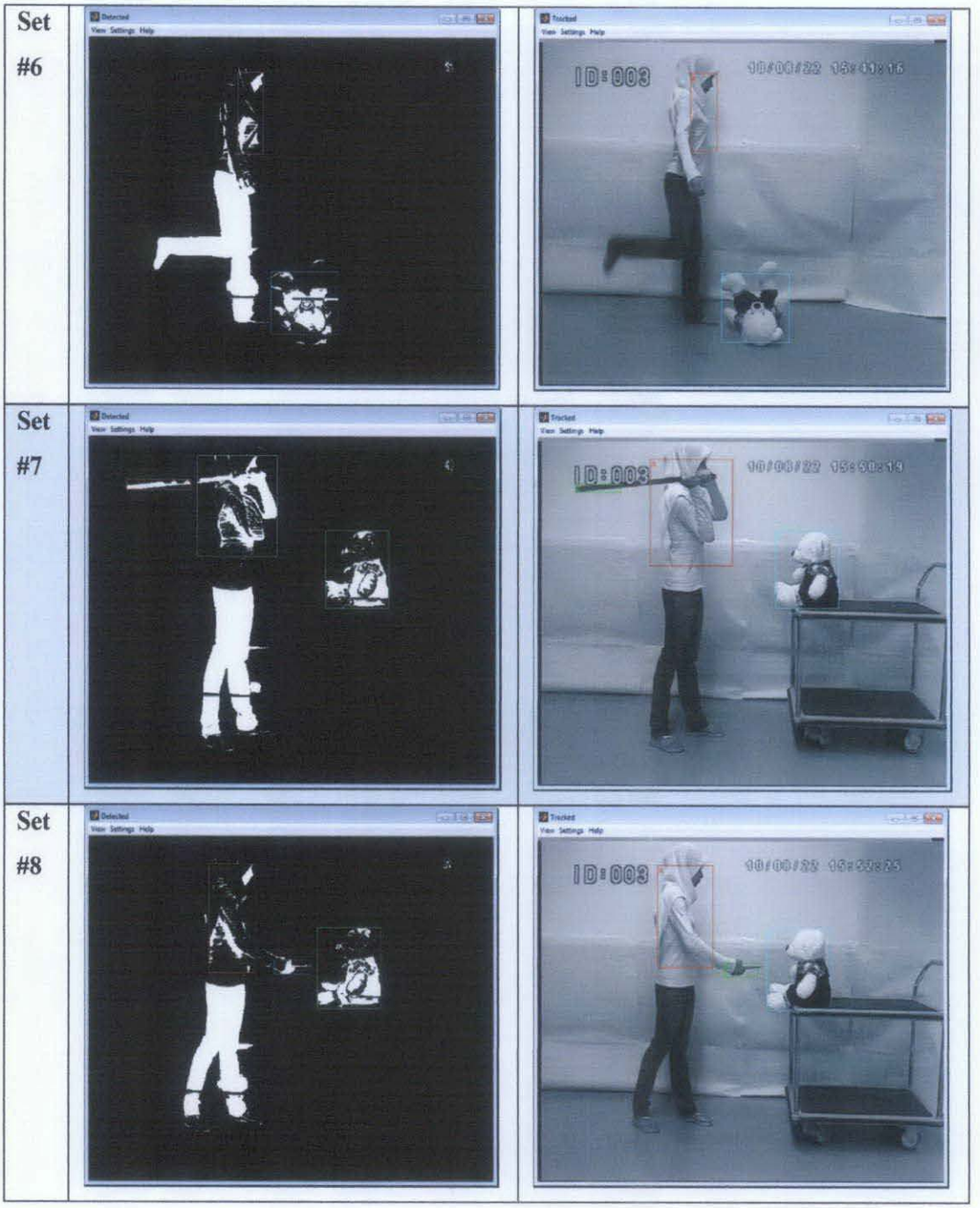


Table 3: Detected and tracked windows result of MATLAB model.

CHAPTER 5
CONCLUSIONS AND RECOMMENDATION

5.1 Conclusions

During the whole project, we successfully obtain the objection like mange to recording eight (8) set of video surveillance with desirable abnormal behavior like stomping, slapping, kicking, and pointed sharp or blunt object. We also does the abnormal behavior with different set of backgrounds. One with stationary background with white scene with no object involves. Next was a random background with different type of object involve like, table, chair, rack and all the furniture.

Then, from the surveillance input, we track and detect the motion of the subject that involve inside the recording. Even though the toddler does not move a lot like the adult we can detect both silhouettes with their own individual box. Next with the third object as harm object, the object does not become as one with the adult figure even though the object is connected. The result is successful since the occlusion can be removed to be happen. By using Kalman Filter, most of the noise is eliminated.

5.2 Recommendations

Since anomaly detection falls above image processing which complex event involve, further deep studies required to precede the artificial intelligent segment. The project also shall continue to develop output triggered by connected the compiled software into hardware programming. High frame rate is required for accurate measurements. Real-time implementation of video surveillance will often require specialized hardware, due to the complexity of the algorithm.

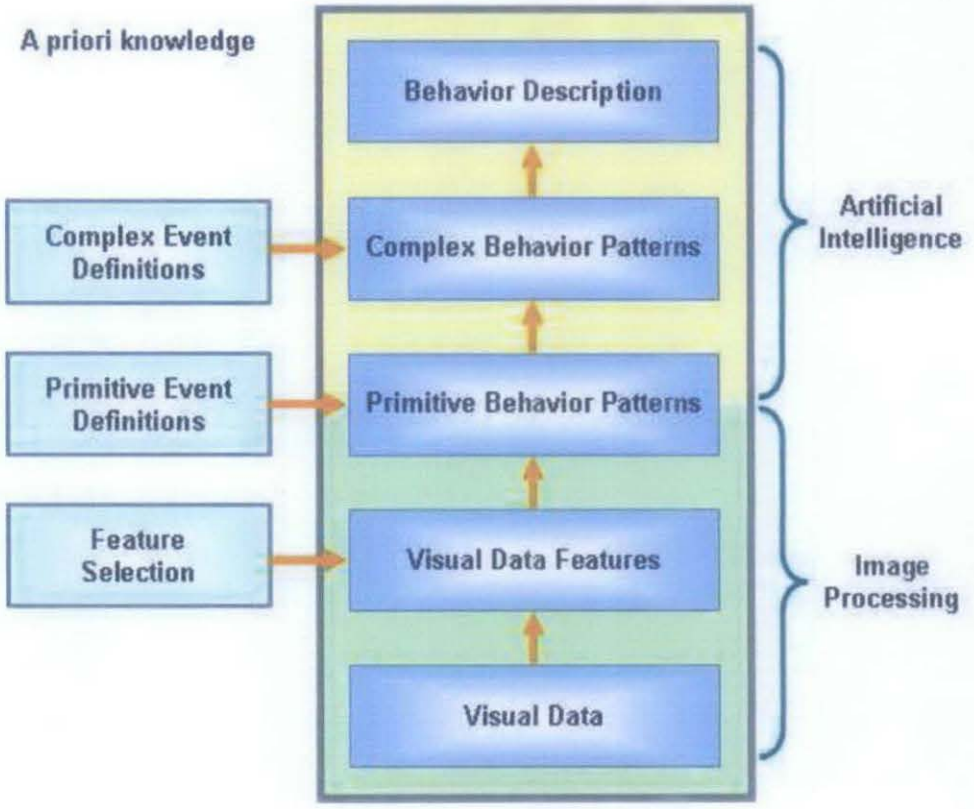


Figure 22: General architecture of video surveillance [19]

As shown in Figure 22, at the first level of a general video surveillance system, geometric features, like areas of motions, are extracted. Based on those extractions, objects are recognized and tracked. At the second level, events in which the detected objects participate are recognized. For performing this task, a selected representation of events is used that defines concepts and relations in the domain of human activity monitoring.

REFERENCES

[1] HelpGuide, Parenting & Family. Retrieve on: 10 August 2010 <http://www.helpguide.org/mental/child_abuse_physical_emotional_sexual_neglect.htm>

[2] Oxford Dictionary Online. Retrieve on: 16 August 2010 <http://dictionary.cambridge.org/dictionary/british/nanny_2>

[3] Foreign Domestic Helper. Retrieve on: 22 October 2010 <http://www.haz.com.my>

[4] TWOFORTYEIGHTAM: TheBside. Retrieve on: 10 August 2010 <http://www.248am.com/mark/kuwait/child-getting-abused-by-a-maid>

[5] Video Surveillance. Retrieve on: 10 August 2010 <http://www.video-surveillance-guide.com/nanny-surveillance.htm >

[6] News Strait Times, August 3, 2010: *"Nanny gets jail for theft, abandoning boss's baby"*

[7] Tao Xiang, and Shaogang Gong. *Video Behavior Profilinf for Anomaly Detection.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 5, May 2008.

[8] Xinyu Wu, Yongsheng Ou, Hunhuan Qian and Yangsheng Xu. *A Detection System for Human Abnormal Behaviour.* Intelligent Robots and Systems, 2005. IEEE/RSJ International Conference, 2-6 August 2005.

[9] J.K Aggarwal and Q Cai. *Human Motion Analysis: A Review.*Nonrigid and Articulated Motion Workshop, 1997. Proceeding IEEE 16 June 1997.

[10] S.E. Smaha. *Haystack: An Intrusion Detection System.* Proceeding of the IEEE Fourth Aerospace Computer Security Applications Conference, Orlando, FL. Pp. 222-232.

[11] J.E. Dickerson, J.A. Dickerson, *Fuzzy network profiling for intrusion detection.* Proceedings of the 19th International Conference of the North American Fuzzy Information Processing Society (NAFIPS), Atlanta, GA, 2000,pp. 301–306.

[12] M. Crosbie, G. Spafford, *Applying genetic programming to intrusion detection.* Working Notes for the AAAI Symposium on Genetic Programming, Cambridge, MA, 1995, pp. 1–8.

[13] A. Valdes, K. Skinner, *Adaptive model-based monitoring for cyber attack detection.* Recent Advances in Intrusion Detection Toulouse, France, 2000, pp. 80–92.

[14] C. Kruegel, D. Mutz, W. Robertson, F. Valeur, *Bayesian event classification for intrusion detection.* Proceedings of the 19th Annual Computer Security Applications Conference, Las Vegas, NV, 2003.

[15] M.-L. Shyu, S.-C. Chen, K. Sarinnapakorn, L. Chang, *A novel anomaly detection scheme based on principal component classifier.* Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop, Melbourne, FL, USA, 2003, pp. 172–179.

[16] P. Yi, Y. Jiang, Y. Zhong, and S. Zhang, *Distributed Intrusion Detection for Mobile Ad hoc Network.* Proceedings of the 2005 Symposium on Applications and the Internet Workshops (SAINTW'05),pp. 94-97.

[17] R. Sekar, A. Gupta, J. Frullo, T. Shanbhag, A. Tiwari, H. Yang, S. Zhou, *Specification-based anomaly detection: a new approach for detecting network intrusions*. Proceedings of the Ninth ACM Conference on Computer and Communications Security, Washington, DC, USA, November 18–22, 2002, pp. 265–274.

[18] Asia One News. Retrieve on: 5 August 2010 <http://www.asiaone.com>

[19] Teddy Ko, 2008, *A survey on behavior analysis in video surveillance for homeland security application.* 37th IEEE Applied Imagery Pattern Recognition Workshop (AIPR) 2008; Washington, DC; 15-17 October 2008.

[20] Joshua Candamo, Matthew Shreve, Dmitry B. Goldgof, Deborah B. Sapper, and Rangachar Kasturi, *Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms.* IEEE Transactions on Intelligent Transpotation Systems, Vol 11, No.1, March 2010.

[21] Anthony C Davies, Sergio A Velastin. *A Progress Review of Intelligent CCTV Surveillance Systems.* IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 5-7 September 2005, Sofia, Bulgaria.

[22] Tomi D Raty. *Survey on Contemporary Remote Surveillance Systems for Public Safety.* IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews, Vol.40, No.5, September 2010.

[23] News Strait Times, February 5, 2008: *"Indon nanny caught abusing baby on CCTV"*

[24] Malay Mail, August 31, 2007: *"Nanny kicked, stomped child"*

[25] Utusan Malaysia, July 10, 2007: "*16 kes amah dera anak majikan*"

[26] Lianyang Ma, Zhijing Liu. *Human Motion Recognition in Video.* Fifth International Conference on Fuzzy Systems and Knowledge Discovery.

[27] Weiyu Zhang, Qingbo Yang, Yushui Geng. *A Survey of Anomaly Detection Methods in Networks.* Computer Network and Multimedia Technology (CNMT) 2009, International Symposium on 18-20 January 2009.

APPENDIX A

Maid kicked, stomped child
MM - Aug 31, 2007


Maid kicked, stomped child

Car-theft gang busted

SHE was kicked so hard that she was sent sprawling to the floor. When the three-year-old got up, the maid grabbed her by the hand and started kicking her again repeatedly. Again, she was sent sprawling like a rag doll. Like a woman possessed, the maid repeatedly stomped on the girl's frail body. She stopped after a while, but continued stomping on the girl moments later.

This cruel act by an Indonesian maid was recorded by a young couple using a webcam installed in the living room of their condominium in Jalan Kuching. But it could have been a different story if the couple had not had the presence of mind to make the recording.

When confronted on Tuesday, the 18-year-old maid allegedly tried to jump off the couple's fourth floor condominium balcony, prompting a cleaner to call the police. When the cops arrived, she told them that she had been assaulted by the couple. Police then hauled the employers, their children and the maid to the Sentul police station for questioning. The couple spent hours explaining what had happened, and it was during questioning that they produced the footage to back up their claim that their daughter had been abused. The couple, who had been married for three years, also claimed that the maid had abused their one-year-old son as well. They lodged a report on the alleged abuse the next day.

The couple had told police that they had installed the webcam two weeks ago after their daughter frequently complained of stomach pain. When they asked the maid, she allegedly told them that it could be because the girl took too much yogurt drinks. Suspicious, the couple decided to install a webcam on the personal computer in the living room. In the first week, they caught the maid abusing their youngest. They confronted her and warned her not to repeat it.

On Tuesday, she did it again - this time, the couple's three-year-old girl was the target. But before they could lodge a complaint with the authorities, the maid turned the tables around. When police arrived after being alerted by the cleaner, the maid told police that she had been assaulted. The maid sustained injuries to the head and several other parts of her body. She was sent to Kuala Lumpur Hospital.

Police are investigating whether the injuries had been self-inflicted or caused by others. The maid, in her police report, had claimed that the couple was responsible for the injuries. It is learnt that the maid's 29-year-old female employer, who works with a shopping mall in the city, had returned home when she felt "uneasy" while working. She felt the need to check on her maid and children. Her husband of three years, a 26-year-old computer programmer, arrived later.

It was when they checked the recording that they discovered what the maid had done to their daughter. It was the third time unlucky for the couple. Their first two maids ran away after working with them for less than a year. The maid allegedly involved in the abuses had been working with them for more than a year.

The Malay Mail spoke to the couple who said their children had been sent to Kuala Lumpur Hospital for medical check-ups. The husband declined to give The Malay Mail access to the footage of the alleged incident. "We are afraid of the impact it might have on the children. We are sorry," he said. The wife said she did not know if they would get a maid after the episode. "There are those lucky enough to get good maids, and there are those who are not as fortunate," she said.

Kuala Lumpur CID chief senior Assistant Commissioner II Ku Chin Wah confirmed that police were investigating both reports. He said police have also retrieved the recordings from the couple. "We are also investigating the maid's claim that the couple had inflicted injuries on her with a weapon," he said. Ku said the couple was released early yesterday after police recorded their statements. The maid has been remanded.