THE REPUBLIC OF TURKEY

BAHÇEŞEHİR UNIVERSITY

A VIDEO SURVEILLANCE SYSTEM FOR DETECTION OF POTENTIALLY DANGEROUS EVENTS IN UNDERGROUND RAILWAY STATIONS

M. S. Thesis

Mehmet Ali DAĞLI

ISTANBUL, 2013

THE REPUBLIC OF TURKEY

BAHÇEŞEHİR UNIVERSITY
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
ELECTRICAL AND ELECTRONICS ENGINEERING

A VIDEO SURVEILLANCE SYSTEM FOR DETECTION OF POTENTIALLY DANGEROUS EVENTS IN UNDERGROUND RAILWAY STATIONS

M.S. Thesis

Mehmet Ali DAĞLI

Supervisor: Assoc. Prof. Çiğdem Eroğlu ERDEM

ISTANBUL, 2013

THE REPUBLIC OF TURKEY

BAHÇEŞEHİR UNIVERSITY
The Graduate School of Natural and Applied Sciences
Electrical and Electronics Engineering

Title of Thesis	:	A Video Surveillance System for Detection of Potentially Dangerous Events in Underground Railway Stations
Name/Last Name of the Student	:	Mehmet Ali DAĞLI
Date of Thesis Defense	:	25.04.2013

The thesis has been approved by the Graduate School of Natural and Applied Sciences.

Assoc. Prof. Tunç BOZBURA
Acting Director

This is to certify that we have read this thesis and that we find it fully adequate in scope, quality and content, as a thesis for the degree of Master of Science.

Examining Committee Members:

Assoc. Prof. Çiğdem EROĞLU ERDEM : ____________________

: ____________________

: ____________________

ACKNOWLEDGEMENTS

First of all I would like to thank my thesis advisor Assoc. Prof. Çiğdem Eroğlu Erdem very much for her endless, kind and optimistic support, guidance and encouragement during my studies and also about endearing image processing topics.

I am thankful to Asst. Prof. Devrim Ünay, Assoc. Prof. Hasan Fatih Uğurdağ and Asst. Prof. Övgü Öztürk who shared their valuable information and experience about developing my computer vision knowledge.

My special thanks to my colleagues from İstanbul Ulaşım A.Ş, Hüseyin Doğan and İbrahim Demirci for the video database support and guidances about the algorithms. Thanks to my other colleague Lütfü Akçil for the endless support and interest about the development of the project.

Finally, I would like to acknowledge my family for their infinite love and support during my life.

25.04.2013 Mehmet Ali DAĞLI

ABSTRACT

A VIDEO SURVEILLANCE SYSTEM FOR DETECTION OF POTENTIALLY DANGEROUS EVENTS IN UNDERGROUND RAILWAY STATIONS

Dağlı, Mehmet Ali

Electrical and Electronics Engineering
Supervisor: Assoc. Prof. Çiğdem Eroğlu Erdem

April 2013, 59 pages

Due to the excess number of security cameras in metro stations, security personnel cannot simultaneously monitor all of the cameras effectively. Therefore, they might miss some of the potentially dangerous events that require timely action. In this thesis, a video surveillance system, which automatically detects several potentially dangerous events, and generates warning messages for the security personnel is presented. The potentially dangerous events that are handled in this thesis consist of: safety line intrusion, proximity warning, fallen object detection, prohibited zone violation and open prohibited zone entry detection. Intrusion of the safety line along the tracks or an accidental fall on the tracks can cause lethal events if necessary precautions are not taken on time. Before the detection of the potentially dangerous events, it is important to detect the moving objects in the scene. For this purpose several state-of-the-art background estimation methods have been implemented and compared. Then moving objects are detected based on the estimated background. Another event that is informative is the state of the train, which can be either approaching, stopping, or leaving the station. For example, if the train has stopped, then detection of safety line intrusion is irrelevant, since passengers will be getting on and off the train.

The layout of the cameras at the pilot station has an active role in the development of the algorithms. Video streams from four cameras in the station, which have different field of views have been used for different purposes. For example, the camera used for the detection of the prohibited zone violation is different from the camera that is used for detection of the safety line intrusion. The developed system has been tested on video data obtained from a pilot station in İstanbul, where no such automatic video analysis system is currently being used. The experimental results are quite promising, and they have been verified by security personnel. The system can operate real-time and can be adjusted via a graphical user interface.

Keywords: video surveillance, security cameras, Train detection, railway security, Information Fusion, safety line intrusion, background estimation

ÖZET

METRO İSTASYONLARINDAKİ OLASI TEHLİKELİ DURUMLARIN KAMERA GÖZETİM SİSTEMİYLE ALGILANMASI

Dağlı, Mehmet Ali

Elektrik-Elektronik Mühendisliği
Tez Danışmanı: Doç. Dr. Çiğdem Eroğlu Erdem

Nisan 2013, 59 sayfa

Metro istasyonlarının güvenliğinin sağlanması için yerleştirilmiş kamera sayısının fazlalığından görevli personel tüm kameraları takip edememektedir. Bu nedenle acil müdahale gerektirebilecek olası tehlikeli birçok olayı kaçırabilmektedir. Bu tez çalışmasında olası tehlikeli durumların otomatik olarak algılanıp ilgili personele uyarı mesajı halinde döndürülmesini sağlayan ve bu olayları kötü sonuçlar doğurmadan önlemeyi sağlayacak bir sistem önerildi. Bu tez çalışmasında ele alınan olası tehlikeli durumlar şunlardır: güvenlik çizgisi ihlali, tren yaklaşım uyarısı, ray bölgesine düşen nesne, yasak bölge ihlalleri ve açık kalan yasak bölge girişi. Platform güvenlik çizgisi ihlali ve ray bölgesine yolcu düşmesi zamanında müdahale olmadığı takdirde ölümle sonuçlanabilecek olaylardır. Olası tehlikeli durumlar algılanmadan önce ortamdaki hareketli nesnelerin tespiti önem arz etmektedir. Bu nedenle gelişmiş arka plan kestirme yöntemleri uygulandı ve karşılaştırıldı. Sonrasında arka plan kestirimine göre hareketli nesneler tespit edildi. Olay tespiti için önemli olan bir diğer bilgide istasyona göre yaklaşma, durma ve ayrılma durumlarında olan trenin takibi. Örneğin tren duruyor iken yolcular trene girip çıktığından güvenlik çizgisi ihlali tespiti pasif hale getirilmektedir.

Pilot istasyondaki kamera yerleşimleri algoritmaların geliştirilmesinde etkin rol oynamaktadır. Görüş alanları farklı olan dört kamera farklı amaçlar için kullanılmıştır. Örneğin yasak bölge ihlali tespiti için kullanılan kamera ile güvenlik çizgisi ihlali tespiti için kullanılan kamera farklı konumlardadır. Geliştirilen sistem İstanbul metrosunda bir pilot istasyondaki, herhangi bir otomatik video analiz sistemi kullanılmayan kameralardan alınan veriler ile test edildi. Elde edilen umut verici deneysel sonuçlar güvenlik personeli tarafından da doğrulandı. Sistem gerçek zamanlı çalışabilmekte ve grafik arayüzü ile konfigüre edilebilmektedir.

Anahtar Kelimeler: video gözetim, güvenlik kameraları, demiryolu güvenliği, bilgi kaynaşımı, tren algılama, güvenlik çizgisi ihlali, arka plan tespiti

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Open Source Computer Vision Library	: OpenCV
Mixture of Gaussian	: MOG
Not Suitable	: N/S
Closed Circuit Television	: CCTV
Command Control Center	: CCR
Safety Line Intrusion	: SLI
Prohibited Zone Violation	: PZV
Fallen Object Detection	: FOD
System Modeling Language	: SYSML
Integrated Development Environment	: IDE
Graphical User Interface	: GUI
Object Oriented Programming	: OOP
C Plus Plus	: C++
Audio Video Interleave	: AVI
Probability Density Function	: PDF
Open Door Detection	: ODD
Region of Interest	: ROI

LIST OF SYMBOLS

Threshold	: τ
Pixel value at a given time	: p_t
Average value	: μ_t
Normal distribution	: η
Weight parameter	: w
Variation	: σ_t
Covariance	: $\sigma_t^2 I$
Result matrix	: R
Template image	: T
Source image	: I
Gradient of the Safety Line	: g_s
Foreground function	: f_t
Total velocity change of pixels	: E
Learning rate	: α

The most auspicious of people is the one beneficial to the people.

1. INTRODUCTION

Similar to many crowded places, public transportation stations such as metros are being monitored using many security cameras. For instance, London has 11000 CCTV (Closed Circuit Television) cameras for underground rail network only (Guardian 2011). However, most of these cameras currently play only a discouraging role and serve as an evidence material for proving and analyzing the details of a serious incident, after it has happened. This is because there are many cameras that have to be monitored at the same time by security personnel, which is a tedious task. Typically a CCR (Command Control Center) deals with video signals originating from 100 to 500 or more cameras (Velastin 2006). Only a small subset of them (5-30) is shown on an array of monitors called as the “video wall” for visual inspection (see Figure 1.1). Since the security personnel are unable to pay attention to all cameras at the same time, they may miss the opportunity to react to unwanted behaviors of passengers timely.

Figure 1.1: Command Control Center

In Istanbul Taksim - Haciosman metro line, which is being used by more than 250.000 people, there are about 800 CCTV cameras placed to observe train platforms, escalators, lifts, exits, toll gates etc. (ITC, 2012). This means there are on the average more than 50 cameras per station. At every station there is a responsible staff who observes all the station from the cameras in addition to other security personnel. Although these personnel has highly trained cognitive abilities to detect abnormal activities in stations, it is so difficult to track all of the cameras at same time. Also it is rather difficult to visually monitor for long periods of time.

Cameras which are located on the train platform of the station are the main focus of this thesis, since most of the possibly dangerous events occur there. There are two platforms in every station of the Taksim-Haciosman line. One is for north, and the other is for the south directions. In total, 10 cameras are located on these platforms. This means that one of the platforms, which belong to the trains of north direction, is observed by four steady cameras and one moveable camera. Two of these cameras can observe the arrival of the train, and two of them can observe the departure. The last one can move in all directions and observe both.

On the platforms, dangerous events may happen, which have the potential to result in death or injuries. For example, there is electrical current on the metro line, which may cause death when the train is not in platform. Also with the train arrival there is crashing risk, when a person accidentally falls on the tracks. In order to prevent such events, passengers should stay away from the end of the platform when there is no train in the platform. This should be monitored by the security personnel. Dropping objects to the metro line is hazardous for the operation of the train, which should also be detected before the incident happen. Another important issue is to detect the violation of the forbidden zones at the end of platform, which are dangerous for the passengers. Passengers should be detected if they enter forbidden zones as shown in Figure 1.2.

Figure 1.2: Forbidden zone

All of these events should be detected before the dangerous events happen. However, for a single staff it is very difficult to monitor all cameras effectively. Therefore, automatic detection of these potentially dangerous events and giving warning message to the security staff is very important.

1.1 MOTIVATION

Around the world many people die or have injuries in railway stations. Therefore, detection of potentially dangerous events before bad incidents actually happen is a crucial task. Since there are many surveillance cameras at stations automatic detection of such events may be possible, and people can be warned beforehand and most of these accidents may be prevented.

Dropped objects on the tracks are dangerous for the operation of the trains. Therefore, it is important that the security personnel becomes aware of such objects without losing time, and preferably automatically. Such incidents cause interruptions in the train schedule and eventually result in financial loss. If it is detected automatically there is a chance to clear the track without interrupting the trips.

Vandalism and graffiti drawn on the walls of the stations are also unwanted events at underground railway stations. Therefore, real-time monitoring of forbidden zones and warning the security personnel is an important issue, which may prevent some of these events.

In stations there are proximity sensors, which detect the arrival of the train to the platform. This detection triggers some functions such as giving warnings on the video walls. With an intelligent video surveillance system detection of the train arrival/departure can also be done with cameras. This will provide energy and cost efficiency.

1.2 LITERATURE REVIEW

Since CCTV cameras are being used intensively on transportation areas the number of applications using these security systems is increasing every year. Due to difficulties in tracking all events in stations from the control center alone; there are many research efforts towards automatic detection of important events.

Understanding behaviors of the passengers is one of the most needed applications that will be helpful to the security personnel. From the CCTV recordings, passengers can be tracked and their positions/postures/actions can be recognized (Fujimura, Kamijo, Yoshimitsu and Naito 2011) (see Figure 1.3).

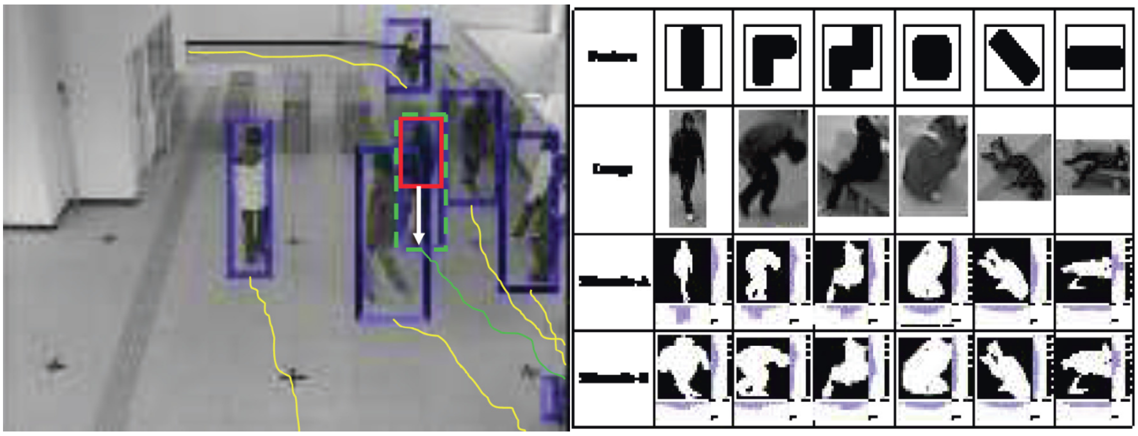

Figure 1.3: Tracking pedestrian locations on the left and posture recognition on the right

Source: Fujimura, Kamijo, Yoshimitsu and Naito (2011)

Another application for railway stations is obstacle detection on the tracks. For the automatically operated railway systems it is an important issue to detect obstacles before a train arrives. The steps of this algorithm for obstacle detection are shown in Figure 1.4. When the program is initialized, it means that independent from the environment such as train, obstacle start of the program is available. In area definition step the platform rail area, the track area, the platform edge areas are determined. With the data organization processed data categorized such as object, message, image etc. In the image input step one gray-scale image is provided from the camera. In image preprocessing part several methods such as shifting, resizing, mirroring are applied to the image. Then in segment generation certain image features are extracted. After segmentation features are classified and defined as object. Finally if there is any object in an unwanted area a message generated (Oertel, Dimter and Szoska 2002).

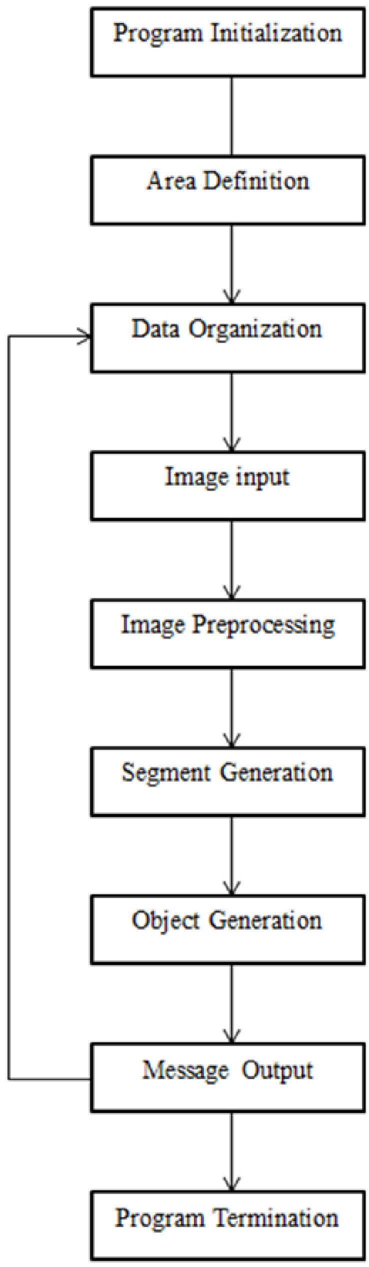

Figure 1.4: Algorithm for obstacle detection on the tracks

Source: Oertel, Dimter and Szoska (2002)

Here are other reference studies at railway stations which aim to detect events automatically. These studies tried to get more and more data from these cameras without human labor. Researches mainly focus on background and foreground estimation in underground stations, determining the motion vectors, object tracking and information fusion from multiple cameras.

1.2.1 Background Estimation

Removing the stationary features of a scene is the first and essential step of video analytics. If the background can be estimated successfully, then object tracking can be done more accurately. There are several methods in the literature to estimate the background of a scene.

A background image may be estimated based on averaging pixel's intensities of a sequence over time if objects in the scene are always moving and small. Detection of movement by calculating inter frame difference can be used to remove objects from background if their intensities are not homogeneous, since movement on homogeneous areas cannot be easily detected using inter frame difference (Velastin, Boghossian and Lazzarato 1999).

Another approach is related to motion. Motion features are used to control a statistical estimation of the background image, whereby only stationary blocks are considered to be background updating candidates. In the context here, stationary blocks are those determined as not being foreground and having a filter block-matching motion vector of zero and having an inter frame difference of zero (Velastin, Boghossian and Vicencio-Silva 2006).

Wei and his group presented a dynamic background estimation for the railway track region. They use the stationary property of the track region (Han, Lin, Ming and Wei 2004). Similarly Nassu used the track stationary for background estimation and rail extraction (Nassu, Ukai 2012).

1.2.2 Foreground Estimation

Once information on the background becomes available, it can be used to determine which image blocks represent foreground data i.e. those whose illumination is different from the corresponding area of background (Velastin, Boghossian and Vicencio-Silva 2006).

1.2.3 Determining Motion Vectors

The changes in illumination that occur from one image to another can be represented by the so called motion vectors. At any given position in the image, the change is measured by a displacement and a direction (see Figure 1.5). Given the large amounts of data involved (for example in images of 512 x 512 pixels), this process is computationally intensive, but needs to be robust in a range of small and large displacements. The net result is a (motion) vector field in image space that correlates to changes in the scenes arising from either object movements, illumination changes or camera movements. This is useful information for a computer-based video analysis system (Velastin, Boghossian and Vicencio-Silva 2006).

Figure 1.5: Sample image with raw motion vectors on the left and image with foreground filtered motion vectors on the right

Source: Velastin, Boghossian and Vicencio-Silva (2006)

1.2.4 Object Detection and Information Fusion

Object detection process can be initialized after the foreground regions are extracted. The most important object for the stations is the train. Most of the events are evaluated and decisions are made according to the arrival and departure of the train.

In the work of Oh, Park and Lee (2007) cameras and the detection processors comprised the information acquisition unit. The information acquisition unit detects and perceives dangerous factor, such as fallen passenger, a disastrous fire and so on, in the monitored area. The detection processor conducts a series of processes i.e. train detection, object detection, object recognition and object tracking. The fusion unit makes more intelligent and meaningful information by using inputted the monitored results from every single camera sensor for analyzing the situation (see Figure 1.6). According to the results from situation analysis, it generates different alarm messages for local station and CCR employees and train driver (Oh, Park and Lee 2007).

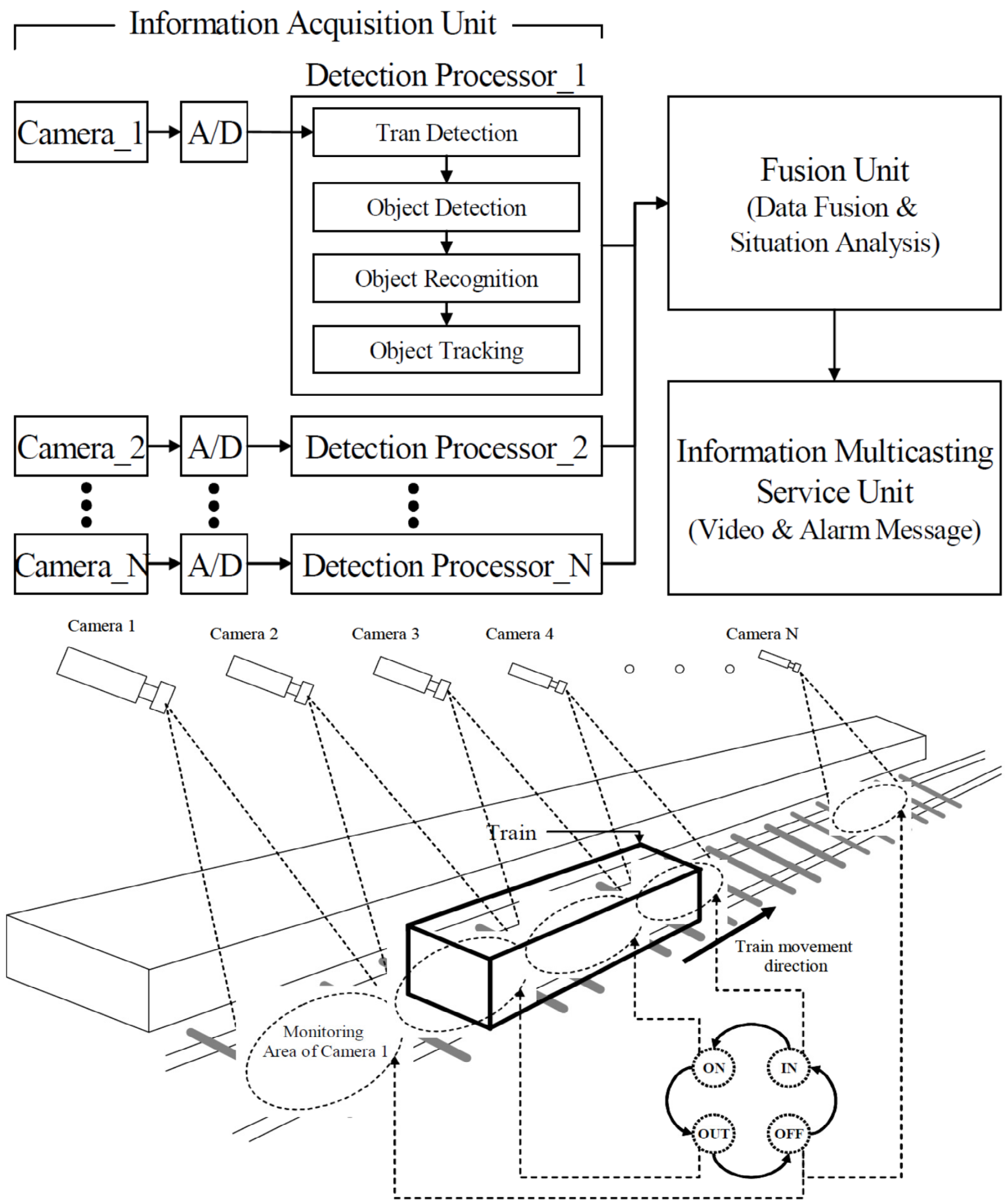

Figure 1.6: System configuration above and Train states for each camera below

Source: Sehchan Oh, Sunghyuk Park and Changmu Lee (2007)

The detection process is divided mainly into two steps i.e. train detection and object/human detection and recognition. The train detection determines the train state to prevent a train from being mistaken for a fallen passenger. The entire detection process for each camera sensor is described in Fig 1.6. The object/human detection process is performed in OFF modes, i.e. a state train does not exist in the monitoring area (Oh, Park and Lee 2007).

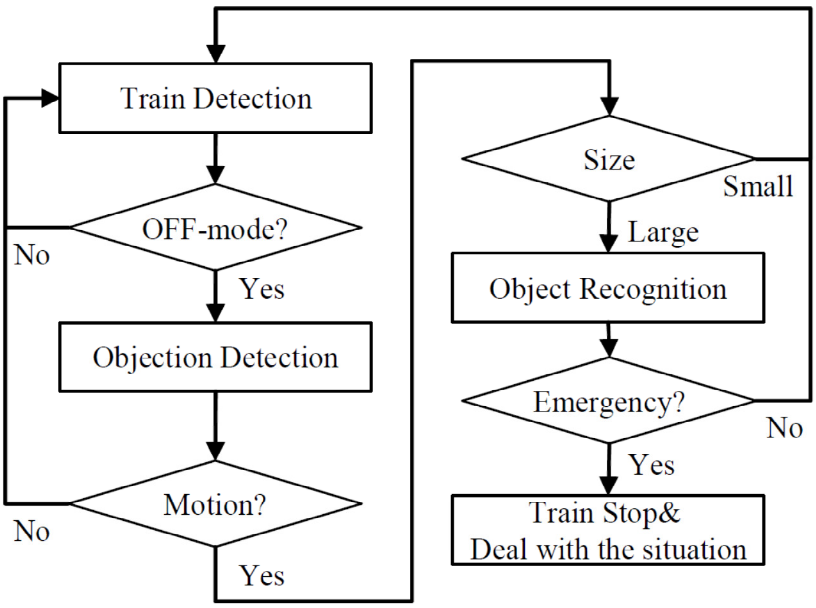

Figure 1.7: Flowchart of Detection Process
Source: Sehchan Oh, Sunghyuk Park and Changmu Lee (2007)

To make a decision about a dangerous factor such as a fallen object in monitoring area, it is important to find the accurate train states in the area for every single camera. The proposed system uses a vision sensor, camera sensor for finding train states in the monitored area with combining the detection results of laser sensors (see Figure 1.7). The four states of the train using camera sensor can be decided as shown in Figure 1.8 (Oh, Park and Lee 2007).

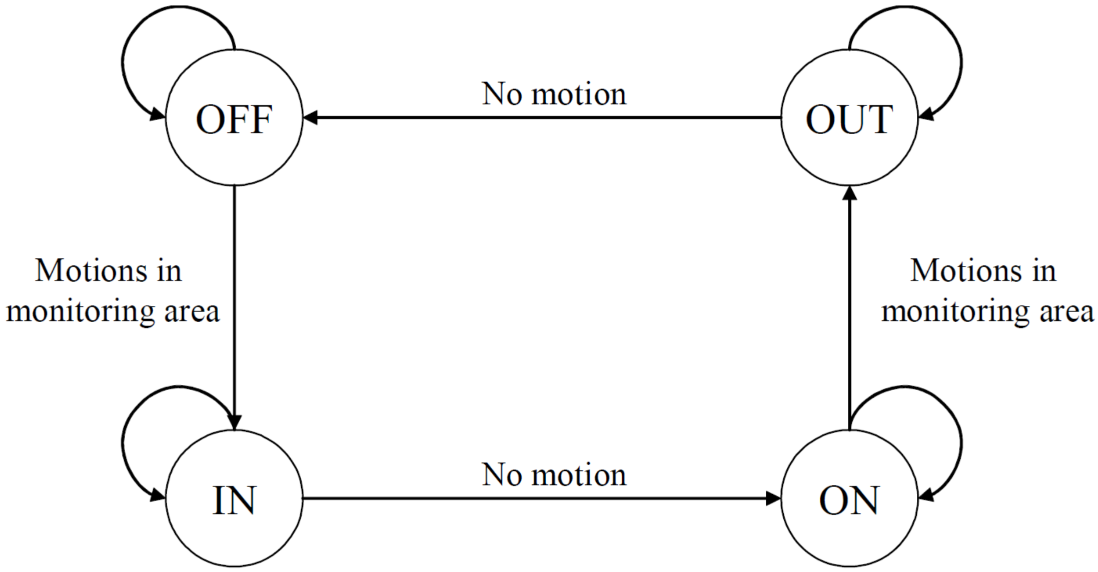

Figure 1.8: State diagram for distinguishing each state of train

Source: Sehchan Oh, Sunghyuk Park and Changmu Lee 2007

Since it is illumination invariant for the indoor environments like underground stations it is easier to detect motion. For outdoor environments illumination calibration, masking is the vital step for the studies (Vazquez, Mazo, Lazaro, Luna, Urena, Garcia, Cabello and Hierrezuelo 2004).

1.2.5 Detection of Dangerous Events

One of the most vital image processing applications on railway station is detection of dangerous events. There are many situations detected by operators which may be potentially dangerous.

Since several studies focused on the tunnel area where the trains are moving (Qiu, Liang, Li, Wu 2009); most of the studies are about the places where passengers are moving. Many of the dangerous events occurred in passenger platforms, turnstiles area, escalator etc. (Carincotte, Naturel, Hick, Odobez, Yao, Bastide and Corbucci 2008). Dangerous events can vary according to the parts of the station.

One of the most active railway operators, RATP (Autonomous Operator of Parisian Transports) defined some of these situations as: Proximity warning, dropping objects on the tracks, launching objects across the platforms, person trapped by the door of a moving train, walking off the rails, fall on the tracks, crossing the rails (Seyve 2005).

Safety line intrusion is the most frequent event in railway stations. This line should not be passed due to the risk of falling to the tracks. Typically this intrusion is detected by the virtual margin on the image which designated manually. Wen, Chen and Lin (2012) presented a method which employs the linear equation of straight line formed from margins using point-slope computing as given in Equation 1.1. Because if there is a point that is located upon the designated segment, it will then be planted into the linear equation and the outcome is 0. If there is a point that is located at the lower designated segment, the outcome obtained will be a negative number, yet if it is located at the upper segment it will be positive number. Therefore, the outcome will turn from negative to positive before and after a point is moved, and we can learn that a motion object has exceeded the margins. In Figure 1.9, the indication chart of motion object that exceeds the margins.

$$mx - y + (y_1 - mx_1) = 0 \qquad \textbf{(1.1)}$$

where $m = (y_1 - y) / (x_1 - x)$ is the slope, $A(x_1, y_1)$, $P(x, y)$ are the initial and terminal coordinates of margins designation, respectively (Wen, Chen and Lin 2012).

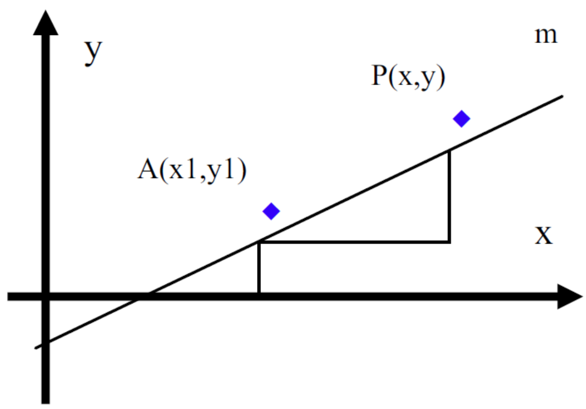

Figure 1.9: Indication chart of motion object in excess of margin
Source: Wen, Chen and Lin 2012

Another important event which has a life-sustaining effect is *Fallen Passenger Detection*. Sasaki and Hiura (2003) defined the platform as 3 parts; Areas on the right side of Fig. 1.10 can be divided into the area on the platform (1), the area below the platform (2) and the area outside the track (3). The fallen passenger detection area of this system corresponds to area (2) having a height ranging from 50 cm to 1 m.

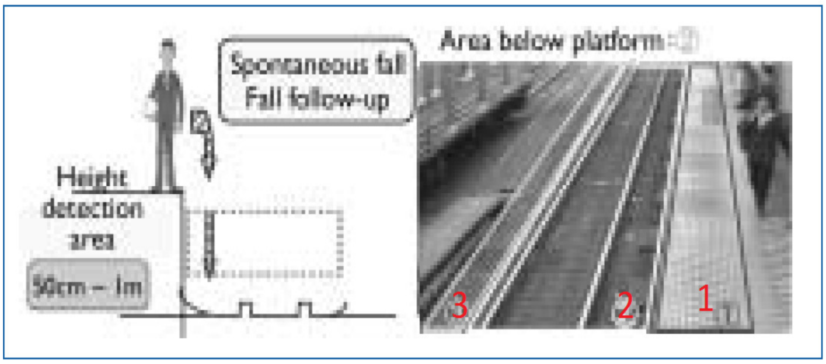

Figure 1.10: Detection area

Source: Sasaki and Hiura (2003)

The system detects a passenger who has fallen into this area. Ordinary fall is a spontaneous fall from area (1) to area (2). Thus, information on spontaneous fall from area (1) to area (2) as well as information about fall follow-up are used for this evaluation to improve the accuracy in distinction between an object other than a human. It is also possible that a passenger could enter area (2) from area (3), although it would not be likely. Detection is carried out in area (2) only when a human being with a height of 70 cm or more has been determined to be in area (3). Fig. 1.11 is a flowchart representing the case where a passenger has fallen into area (2).

(1) -> (2): Extracts features of the image on each of the right and left.

(2) -> (3): Extracts parallax from each of the feature-extracted images.

(3) -> (4): Checks parallax-extracted image against detection area.

(4) -> (5): If the status in area (2) continues for two or more consecutive frames and has a height of from 50 cm to 1 m and a size of 5 or more blocks (a size where a 30 square centimeter box is visible from a distance of 40 meters), the size is evaluated according to the distance. If the size exceeds a certain level, it is extracted as representing a fallen person.(Y.Sasaki, N.Hiura 2003)

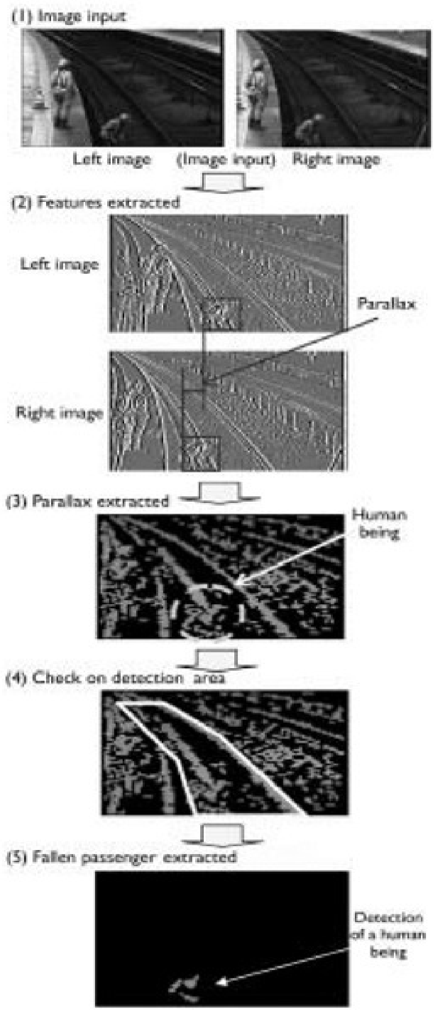

Figure 1.11: Flow of fallen passenger detection
Source: Sasaki and Hiura (2003)

Overcrowding states are dangerous for the railway stations too. It may cause negative results when the train arrives. In order to prevent these results it should detected beforehand. Also the operators need to divert the passengers to the less crowded sections of the station. Velastin (2006) presents an approach to use the perspective calibration and correction method to compute a weighted sum of the number of foreground blocks. This sum correlates to the number of pedestrians in the scene. In fact, what is computed is the (normalized) ratio between the perspective-weighted number of foreground blocks and the maximum perspective-weighted number of blocks in the Region of Interest that has been chosen for detection. This gives a normalized indication of density (Velastin, Boghossian and Vicencio-Silva 2006).

1.3 CONTRIBUTIONS OF THE THESIS

In this thesis, a system for automatic detection of potentially dangerous events using the videos captured by surveillance cameras has been realized. According to the corporate records of Istanbul Transportation, it is the first computer vision study about metro stations in Istanbul. The system has been customized according to the camera locations for the Taksim-Hacıosman metro line. Currently, although there are many surveillance cameras in Şişli-Mecidiyeköy station, no automatic processing exists, and the security personnel is overwhelmed by the abundance of video data that need to be monitored. Hence automatic detection of potentially dangerous events may be life- saving for accidental fall events of a passenger (see Figure 1.12) and may prevent also serious economic losses.

Figure 1.12: A passenger felt to the tracks

The following events can be detected by the realized system:

- **Safety line intrusion and Fallen object detection:** Passengers who get close to the end of passenger platform are automatically detected, and a warning signal is given to the security personnel Also the attention of the security personnel is directed to the specific camera that is capturing the event from the best angle so that the personnel can then take the necessary precautions on time. This system can also be used for making automatic announcements to the passengers warning them to stay behind the safety line.

- **Proximity warning (Train detection)**: We detect the arrival/departure of the train, since this is required to assess the relevance/urgency of a security line intrusion. For example an intrusion while the train is approaching needs to be urgently warned, whereas an intrusion when the train has stopped is inevitable. Also with the arrival of the train advertisements are finalized and warning messages displayed on the information screens. (see Figure 1.13)

  There are proximity sensors that detect the arrival of the train. With this study arrival of the train can be detected by the security cameras. This will provide energy and cost efficiency.

Figure 1.13: Information panels of the station

- **Prohibited Zone Violation and Open prohibited zone violation:** People intrude forbidden zones for drawing graffiti or vandalism can be detected. Due to the high number of cameras, security personnel can sometimes miss such incidences. They are detected by security cameras, which will help prevent material damages and economic loss. Also with the surveillance of the forbidden zones, if any of the entrances of those zones are open the system warns the security personnel about that.

The list of the former computer vision studies in railway operation and contributions of the thesis are shown in Table 1.1. This study has also contributions about combining different purposes in a single solution.

	Safety Line Intrusion	Fallen Object Detection	Prohibited Zone Violation	Train Detection	Overcrowd Detection	Open Prohibited Zone
Sasaki, 2003		✓		✓		
Velastin, 2006					✓	
Wen, 2012	✓			✓		
Seyve 2005		✓		✓		
Oh, 2007				✓		
Dagli, 2013	✓	✓	✓	✓		✓

Table 1. 1: Former Computer Vision Studies in Railway Environment

1.4 OUTLINE OF THE THESIS

The outline of the thesis is as follows. In Chapter 1, a literature review of the current state of the art methods for detection of potentially dangerous events in metro stations is provided. Also detailed examples about estimation of the foreground scene of the stations and tracking objects are given.

In Chapter 2, three different methods for the background estimation are explained in detail and afterwards foreground extraction from the frames is described.

In Chapter 3, camera layout of the pilot station and the features of each camera location are mentioned. Then one of the most important events "Train Detection" is described.

In Chapter 4, tracking the moving objects process which has importance for the event detection is explained. Then the algorithms for the events "Safety Line Intrusion (SLI)", "Fallen Object Detection (FOD)", "Prohibited Zone Violation (PZV)" and "Open Door Detection (ODD)" are described in detail.

In Chapter 5, experimental results for the comparison of the background-foreground estimation is given. Then experiments about the train detection, SLI, FOD, PZV, ODD events are also presented with statistical results.

In Chapter 6, a summary and conclusions of the thesis are presented. Then further steps of the thesis for future works are mentioned.

2. BACKGROUND ESTIMATION IN THE STATION

In any given image containing pedestrians, it is important to remove permanent environment features (the background) prior to any analysis such as pedestrian counting. In some situations, it is physically possible to capture an image frame with no people, in which case that frame is used as a reference image for further processing. When this is not possible, e.g. a site that always has some pedestrian traffic as shown in Figure 2.1; different techniques are needed to capture such background image automatically from images containing pedestrians.

Figure 2.1: Background crowded with passengers

2.1 BACKGROUND ESTIMATION IN QUIET HOURS

Most of the underground railway lines are operated in a time interval. Istanbul Metro lines are operational between 06:00 and 23:59. During the non-operating period maintenance, controls and tests are done and generally there are no passengers on the station platform. So the platform background is almost stationary. A video captured in these non-operational time intervals can be used as background reference image. An image, which can be used as background is shown in Figure 2.2.

Figure 2.2: Background with no passenger

The cameras which are used for the experiments on the platform are stationary. Also their field of view is not changing because they cannot zoom, pan or tilt. If there were no changes in the background; the screenshot of any day could be used as background for all days. But the background is also changing a bit due to maintenance, construction or operational processes as shown in Figure 2.3(a).

Figure 2.3 (a): There are construction items on the image (b): Another day's background image

When different test videos of different days are used it has been observed that that although the cameras are stationary and they cannot tilt, zoom there are some changes at their field of view. It is understood that because of some calibration works there are changes on the field of view of the cameras and also the reference background image. Due to day-to-day changes in the background; a background image has to be captured for that day's calculations.

2.2 DYNAMIC BACKGROUND ESTIMATION

If it is guaranteed that there is no foreground object in an image then this image can be used as a background image. In underground metro stations, there are less illumination changes as compared to other railway stations like tram stations that operate under daylight. However, there are many new objects that are added slowly to the scene, so it is difficult to capture an image without any foreground objects.

A better strategy is to dynamically build the background model by regularly updating it. This can be accomplished by computing what is called a running average (also called moving average). It is a way to compute the average value of a temporal signal which takes into account the latest values received. If p_t is the pixel value at a given time t, and μ_{t-1} is the current average value, then this average is updated using the following formula:

$$\mu_t = (1 - \alpha)\, \mu_{t-1} + \alpha\, p_t \qquad \textbf{(2.1)}$$

The parameter α is called the learning rate and it defines the influence of the current value over the currently estimated average. The larger this value is, the faster the running average will adapt to changes in the observed values. To build a background model, one has just to compute a running average for every pixel of the incoming frames. The decision to declare a foreground pixel is then simply based on the difference between the current image and the background model (Laganiere 2011).

Foreground objects can be found quickly with this method. If the learning rate is high, the background is renewed quicker. If learning rate is low then there are some problems with the foreground.

Flow diagram for the dynamic background estimation is shown in Figure 2.4. After program initialization a frame captured from the selected test videos. For fast processing it is converted to grayscale.

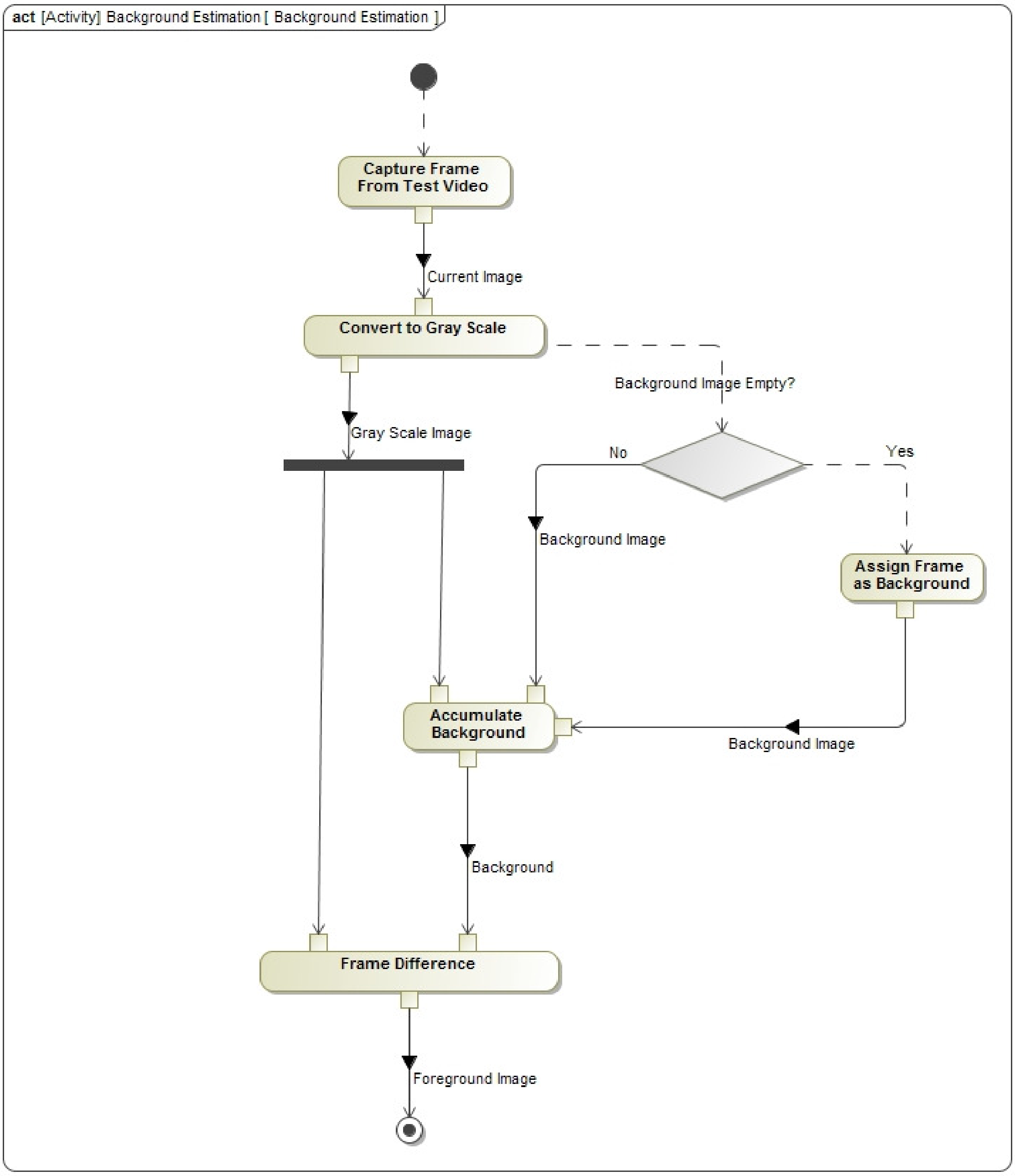

Figure 2.4: Flow diagram for the dynamic background estimation.

If there is no background image assigned before, first frame assigned as background image. Then with OpenCV (Open Source Computer Vision Library) *absdiff* method foreground objects are extracted from background and current image's absolute difference. Also foreground pixels are thresholded by the defined value and set to 0 with the OpenCV *threshold* method. That means if the difference between the any pixel of background and the current image greater than the threshold value, it is the foreground pixel and shown at black color. This thresholded foreground image provides a mask for accumulating the background for the next cycles. Background image is renewed by OpenCV *accumulateWeighted* method with the arguments *learning rate*, *foreground image* and the last *captured frame*. With this renewed background image new foreground objects extracted in the next cycle.

In the previous method background is estimated using a fixed threshold value. In order to cope with the illumination changes or little camera movements adaptive methods can be used. One of the adaptive background estimation methods is the Mixture of Gaussian method. Grimson presented a statistical background modeling method (Stauffer and Grimson 1999). In this method they tried to match a Gaussian distribution for the histogram of each pixel. Then they achieved the PDF (Probability Density Function) of the background model. The probability of a pixel value in time t is as follows:

$$p(p_t) = \sum_{k=1}^{n} w_k \eta(p_t; \theta_k) \quad \textbf{(2.2)}$$

Where $\mathrm{p_t}$ is the current pixel value, w_k is the weight parameter of the k^{th} Gaussian component and $\eta(p_t; \theta_k)$ is the normal distribution of k^{th} component. It is defined with the mean value of the pixel μ_t, the covariance $\sigma_t^2 I$ and the variation σ_t.

$$\eta(p\,; \theta_k) = \eta(p\,; \mu_t\,, \sigma_t^2 I) \quad \textbf{(2.3)}$$

Then the distributions are ordered and the minimum value from w_k / σ_k is used as model. According to the order for each pixel, intensity value is checked whether it fits to a Gaussian distribution as defined in (2.4).

$$|\, p_t - \mu_t| \leq 2.5 * \sigma_t \quad \textbf{(2.4)}$$

If there is a matching on the current pixel then it is defined as background. Then the distribution parameters are updated as follows

$$\mu_t = (1 - \alpha)\mu_{t-1} + \alpha * \mathrm{p_t} \quad \textbf{(2.5)}$$

$$\sigma_t^2 = (1 - \alpha)\, \sigma_{t-1}^2 + \alpha(\mathrm{p_t} - \mu_t)^2 \quad \textbf{(2.6)}$$

$$w_t = (1 - \alpha) w_{t-1} + \alpha * f_t \quad \textbf{(2.7)}$$

f_t shows whether a pixel is foreground or not.

$$f(w_t|\mathrm{p_t}) = \begin{cases} 1, & if\ w_t\ is\ matched\ with\ Gaussain\ component \\ 0, & otherwise \end{cases} \quad \textbf{(2.8)}$$

As similar to the previous dynamic background estimation method, when the α learning rate is high there is higher effect of the current pixel on the background pixel.

When a given Gaussian model is not hit sufficiently often, it is excluded as being part of the background model. Reciprocally, when a pixel value is found to be outside the currently maintained background models (that is it is a foreground pixel), a new Gaussian model is created. If in the future, if this new model becomes frequently hit, then it becomes associated with the background.

For applying this algorithm there is an OpenCV implementation called BackgroundSubtractorMOG and is defined as a subclass of the more general BackgroundSubtractor class. For the experiment in the next section this class is used with default parameters.

2.3 FOREGROUND EXTRACTION

When the reference background image is ready the foreground objects can be extracted easily. By frame differencing the foreground objects can be acquired as shown in (2.9). $p(x, y, t)$ is the pixel density value of the location (x, y) in the current frame and $p(x, y, t-1)$ is the value from the previous frame. Since the cameras and stations are stationary steady background images from quiet hours are used. So the $p(x, y, t-1)$ is steady for all of the related day's calculations. τ is the threshold value and with the experiments it is chosen 40 for this primitive method.

$$foreground(x) = \begin{cases} 0, & p(x,y,t) - p(x,y,t-1) \leq \tau \\ 1, & p(x,y,t) - p(x,y,t-1) > \tau \end{cases} \quad \textbf{(2.9)}$$

For the train detection the method above used. With dynamic background estimation foreground objects are obtained as defined as (2.8)

3. CAMERA LAYOUT OF THE STATION AND TRAIN DETECTION

In this section the camera layout at the selected station is described, since the video captured by each camera will be used for a different purpose depending on its field of view.

3.1 LAYOUT OF THE CAMERAS MOUNTED AT THE STATION

At Istanbul Metro Line every station has at least two platforms. There are four stationary cameras on each platform. Two of them cover the north direction and the other two cover the south. Each platform is used for only one direction except for emergency situations. In this thesis, the east platform of the pilot station "Şişli-Mecidiyeköy" has been chosen meaning that the platform hosts trains coming from north and has a trip name as "Hacıosman to Taksim".

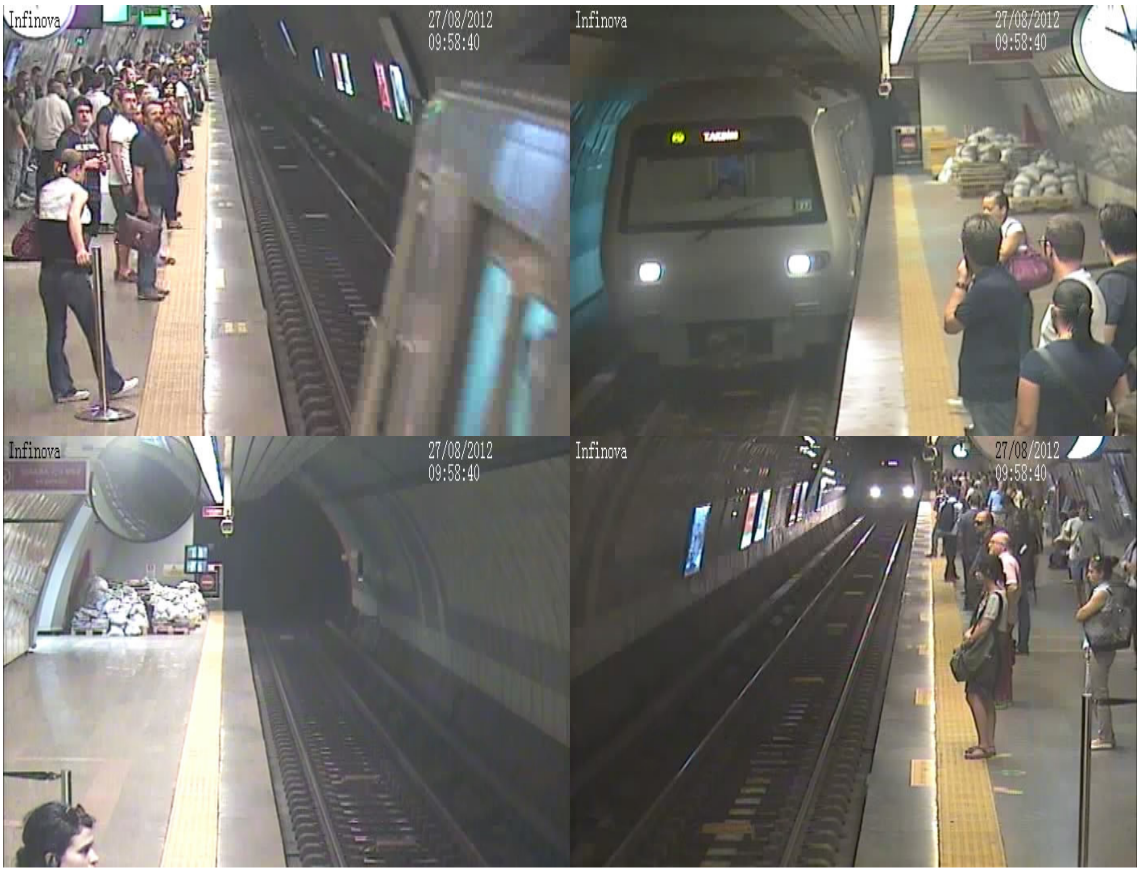

Figure 3. 1: Images from all four cameras captured at the same time instant.

Two of the cameras observe the north direction so they are observing the arrival of the train. Other two observe the south direction and they are observing the departure of the train. (see Table 3.1)

	1	2	3	4
Arrival - Departure	Arrival	Departure	Arrival	Departure
Direction	North	South	North	South

Table 3. 1: List of camera views indicating their direction and state of movement

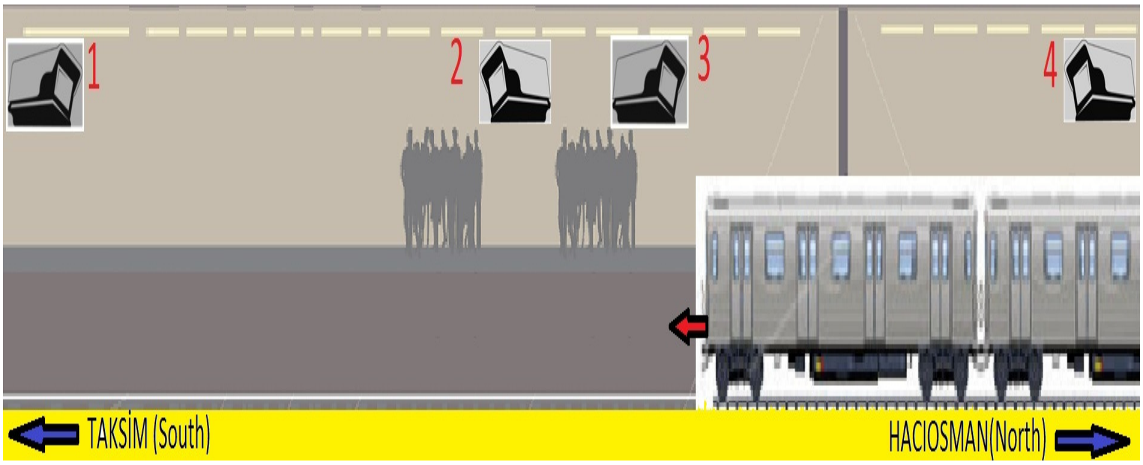

Figure 3. 2: The layout of the cameras mounted at the selected metro station.

Each camera has a location which has certain advantages:

a) Camera 1 is the most suitable one for monitoring the platform for yellow line violation and fallen object detection since the platform area is in its field of view. Since this camera sees the north side of the tunnel, it can also observe the arrival of the train (see Figure 3.3).

Figure 3. 3: Arrival of the train can be detected with images captured by camera 1

Figure 3. 4: Safety line intrusion can be detected with images captured by camera 1

b) Camera 2 observes the south side tunnel exit and has the most suitable field of view for detection of the train departure. Also it can see the entrance of the restricted area at south side. (see Figure 3.5)

Figure 3. 5: An image captured by camera 2

c) Camera 3 observes the tunnel entry at the north side and has the most suitable view for detection of the train arrival. Because it is closer to the entry of tunnel, its field of view for detection of train is better than camera 1. Also it has a sight of the restricted area entrance at the north side. (see Figure 3.6)

Figure 3. 6: An image captured by camera 3

d) Camera 4 has a good location for monitoring the platform for yellow safety line violation and fallen object detection at the opposite direction of Camera 1. Since this camera can see the south side of the tunnel it can also monitor the departure of the train but not as well as camera 2, because it is far from the south side of the tunnel exit. (see Figure 3.7)

Figure 3. 7: Departure of the train can be detected with images captured by camera 4

Figure 3. 8: Safety line intrusion can be detected with images captured by camera 4

	Safety Line Intrusion	**Fallen Object Detection**	**Prohibited Zone Violation**	**Train Detection**
CAMERA 1	Suitable	Suitable	N/S	Suitable
CAMERA 2	N/S	N/S	Suitable	Suitable
CAMERA 3	N/S	N/S	Suitable	Suitable
CAMERA 4	Suitable	Suitable	N/S	Suitable

Table 3.2: Most suitable tasks for different cameras

3.2 TRAIN DETECTION

Detection of the state of the train is the first step of making an inference about potentially dangerous events. The train can be in one of the four states according to each camera:

1) **NONE:** There is no train in the monitored area.
2) **APPROACH:** The train is approaching the station.
3) **STOP:** The train has stopped at the station.
4) **LEAVE**: The train is leaving the station and/or the monitored area.

Operators want to watch the screen especially when the train arrives. They should take more care on the platform when the train is approaching and has stopped. Also yellow line intrusion warnings should be done according to the presence of the train. When there is no train, passengers should not pass the yellow line. When the train arrives they have to pass over the yellow line to get into the train.

The steps of detection of the train shown in Figure 3.9. With the program initialization a frame from the quiet hours taken as background image. Quiet hour means non-operating hours. There are almost no people at these hours and there is least change at the scene. So that frame is used as background image. When the background image is ready, capturing process from the test video starts. After converting them to gray scale a frame difference step is realized. Then the foreground image achieved. There is a defined region of interest (ROI) which represents the tunnel entry as shown in Figure 3.10. If the coordinates of the ROI which is rectangle shaped consist of the corner points $(x_i,\ y_i)$, $(x_j,\ y_j)$, $(x_k,\ y_k)$, $(x_l,\ y_l)$ and $p_{(i)}$ is the intensity value at $(x_i,\ y_i)$ then the sum of the pixel values in that region is

$$p_t = \sum_{y=0}^{k-i} \sum_{z=0}^{j-i} p_{(i+z+y)} \tag{3.1}$$

and the average pixel density is

$$\mu_t = \frac{p_t}{(j-i)(l-k)} \tag{3.2}$$

In the foreground image tunnel entry area is observed.

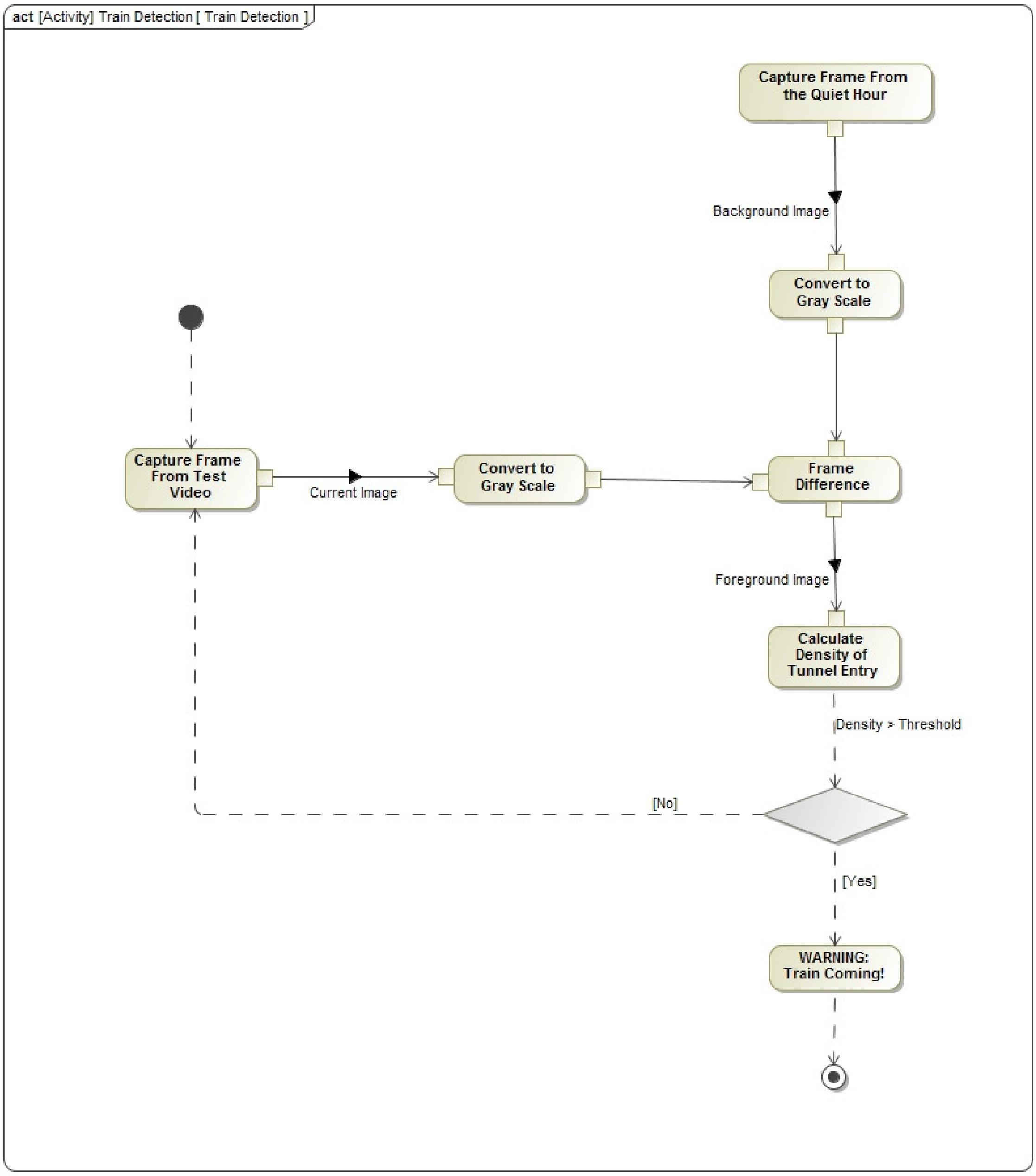

Figure 3. 9: Flow diagram for train detection

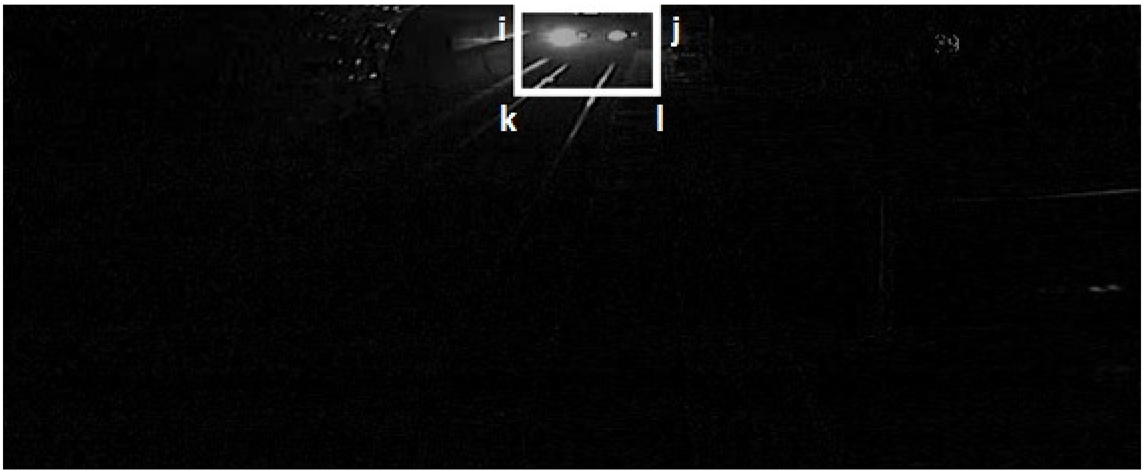

Figure 3. 10: ROI where the train enters the station

$$detection(x) = \begin{cases} 0, & \mu_t \leq \tau \\ 1, & \mu_t > \tau \end{cases} \quad \textbf{(3.3)}$$

As shown in (3.3) if μ_t which is the average pixel density of this ROI is greater than the threshold value this means the train is approaching. There is a big change on the density because of the light of the train. The threshold value determined with several tests of train arrival. The optimal value is chosen. Also after the detection there is a waiting interval for 5 seconds when detection process is stopped. This interval is for the entering of the train into the station. This means there will be no second detection until 5 seconds passed, in order to prevent multiple detection.

Since the camera locations are constant during the day; the train enters into the scene from the same point. This means, in a video the region for train entrance is fixed. However, if there is a camera maintenance their angles may change slightly, which affects the region of interest for train entrance. When the angle of the camera changes, the algorithm may fail since the tunnel may not be in the field of view of the camera causing false detections of false alarms.

4. OBJECT TRACKING AND DETECTION OF POTENTIALLY DANGEROUS EVENTS

In metro stations there are many potentially dangerous events occur that may result in death or injuries. In order to detect these events before they happen; some of the features should be detected and tracked. The most important objects to track are the passengers. If a passengers falls on to the tracks then the high voltage in the railway may kill him. Another potentially dangerous event is the violation of the restricted areas (such as the tunnel entrance), which may threaten the passenger's life and the security of rail system.

4.1 PASSENGER TRACKING USING MOTION VECTORS

Tracking the passengers on the platform is useful for detecting potentially dangerous events such as violation of restricted areas and falling on the tracks accidentally or on purpose.

In order to track the passengers on the platform the optical flow estimation methods implemented in OpenCV were used. This methods use the optical flow equation and Lucas-Kanade approach. The optical flow equation depends on the assumption that the intensity of the pixel $I(x, y)$ is constant along its motion trajectory (4.1).

$$\frac{dI(x,y)}{dt} = 0 \qquad \textbf{(4.1)}$$

And the motion vector of the related pixel which is defined as $v = [v_x, v_y]$ is very small (4.2).

$$\frac{dI(x(t),y(t),t)}{dt} = 0 \qquad \textbf{(4.2)}$$

To solve the problem there is need for one more assumption about the velocity of the pixel. Furthermore Lucas-Kanade feature tracking algorithm is used. According to Lucas-Kanade the displacement of all points in the neighborhood of the feature point is the same. So the sum of the displacement of the neighbor pixels square which is defined as E is minimum.

$$\frac{dE}{dv_x} = 0 \tag{4.3}$$

From (4.3) for the x coordinate one more equation can be obtained as follows

$$\frac{dI}{dx} = I(x + 1, y) - I(x, y) \tag{4.4}$$

With the both coordinates of the pixel if the displacement $v = [v_x, v_y]$ defined as (u, v) then the optical flow equation will be as follows

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v = -\frac{\partial I}{\partial t} \tag{4.5}$$

In OpenCV implementation there is a chance to adjust the resolution about neighborhood. Also in this OpenCV *calcOpticalFlowPyrLK* method there are flags which tell whether the tracking is successful or not.

An instance about the method is shown in Figure 4.1. The obtained feature points are tracked during the man is walking through the safety line.

Figure 4.1: Optical flow of feature points of the passenger is marked

4.2 SAFETY LINE INTRUSION

Tracks where the trains are moving and the platform where the passengers standing are separated by a level difference. The platform is at a higher level than the tracks. Passengers should not get into or even approach to the track level due to the dangers of the high voltage and the arrival of the train. In order to prevent possibly dangerous events such as accidentally falling on to the tracks, passengers should stay away from the end of the platform.

In order to mark the secure region on the platform, a yellow security line has been mounted close to the edge of the platform. When a passenger passes beyond this yellow line a warning should be triggered. Yellow security line and the intrusion samples are shown in Figure 4.2.

Figure 4.2: Sample images for SLI (security line intrusion)

For this purpose a virtual security line is defined in the video. The location of this line is determined manually. It is not on the real yellow line, it is at the left side of the end of platform. Because if the real yellow line chosen, then there will be many intrusion events most of which are false alarms. For example, if a passenger leans to the track region when he is behind of the yellow line there will be intrusion. Furthermore virtual safety line defined at the end of the platform as shown in Figure 4.3. If it is required to be located on a different place then it can be changed.

Figure 4. 3: Passenger lean to the track region dangerously

Then the foreground objects are extracted from the frame with the Mixture of Gaussian method in OpenCV BackgroundSubtractor class. The foreground image is binary inverted in order to show the foreground objects as black pixels. Since passengers should pass through that safety line in order to get into the train, when the train arrives the control of the intrusion stopped. For detecting the train arrival left side of the safety line is observed pixel by pixel before the intrusion control initialization. When the train arrives, train is the foreground object in the track region. When there is no foreground object in the track region then this means the train has departed. Again the foreground objects which are at the right side of the safety line are observed whether they intrude the safety line or not.

Intrusion is detected by checking all of the points of the safety line. The coordinates of the two initial points of the safety line $(x_1,\ y_1)$ and $(x_2,\ y_2)$ are chosen. According to these points the gradient of the safety line which is defined as g_s is calculated as (4.2).

$$g_s = \frac{(x_2 - x_1)}{(y_2 - y_1)} \tag{4.2}$$

In every frame from the initial point $(x_1,\ y_1)$ the pixels are checked whether there is any foreground pixel on it. Then the x_1 is increased with respect to g_s and y_1 is decreased one pixel and the same control mechanism continues until the $(x_2,\ y_2)$ end point is reached. If any of the black foreground pixel intrudes to the virtual safety line, then this is defined as Safety Line Intrusion. The flow chart of the SLI is shown in Figure 4.4.

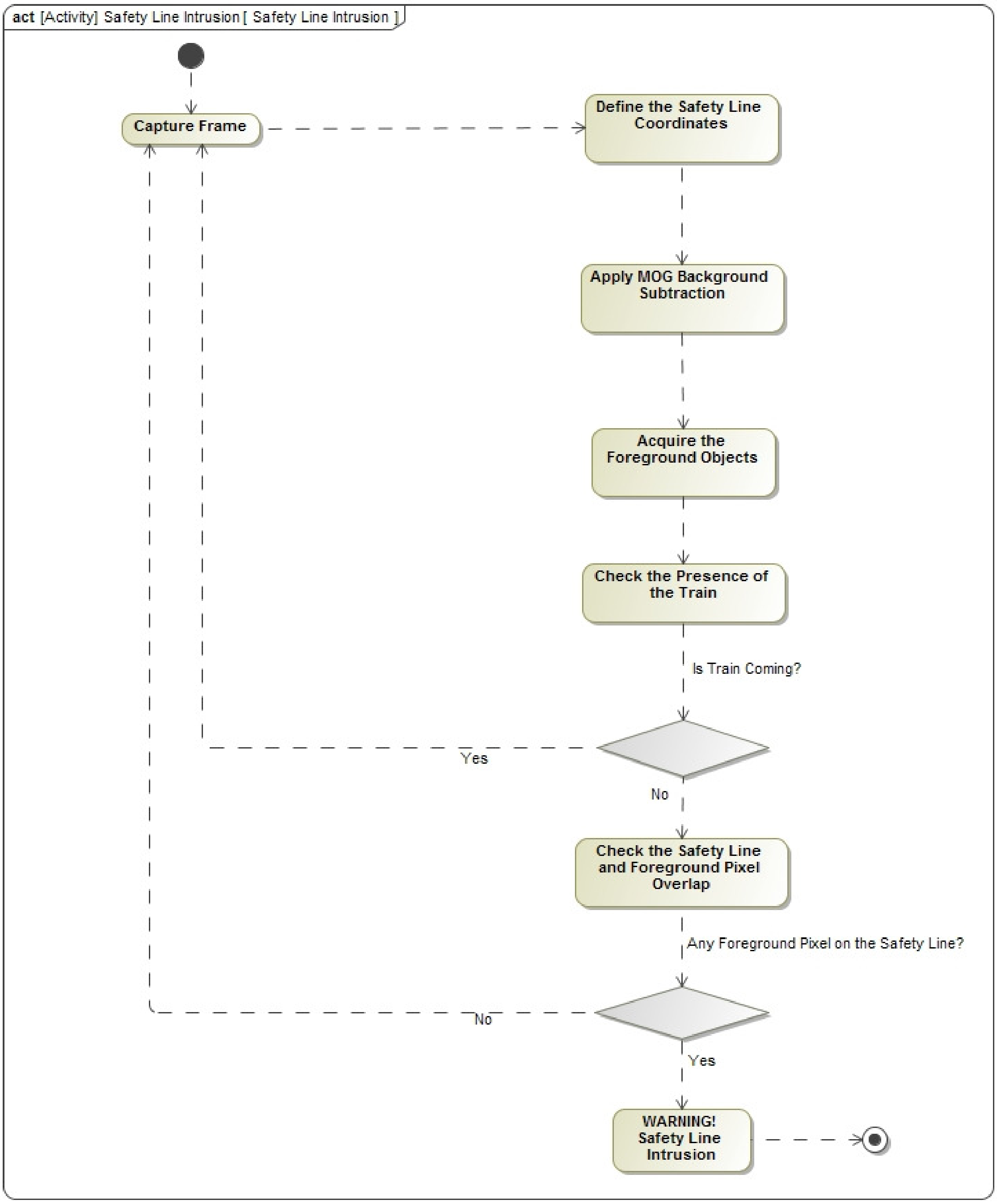

Figure 4. 4: Flow chart for the detection of Safety Line Intrusion with tracking pixel density

When a passenger or part of his body violates the security line, a warning is generated by the system. This warning can be implemented as a visual effect on a screen for the security personnel or can be formed as an audible warning for the staff and the passengers.

4.3 FALLEN OBJECT DETECTION

In a daily operation there are many yellow line intrusion events. This should be detected for the possible dangerous events. But rarely an object falls to the rail tracks. This can be a passenger, a bag or another thing. If a passenger falls to the tracks this is a crucial risk due to high voltage and the train arrival. Driver sees track region of the station when he enters the platform area. For the driver mostly it is hard to stop the train in a second. In order to prevent this event fallen object should be detected and warning should be generated to the personnel for the halting action. Also if a bag or another object falls to the dangerous track region which is shown in Figure 4.5; this may cause damage for the rail system, this should be detected too.

Figure 4. 5: The dangerous track region shown with yellow color

For that needs track region is controlled continuously and system generates warning even if an object falls. Sample fallen object event is shown in Figure 4.7.

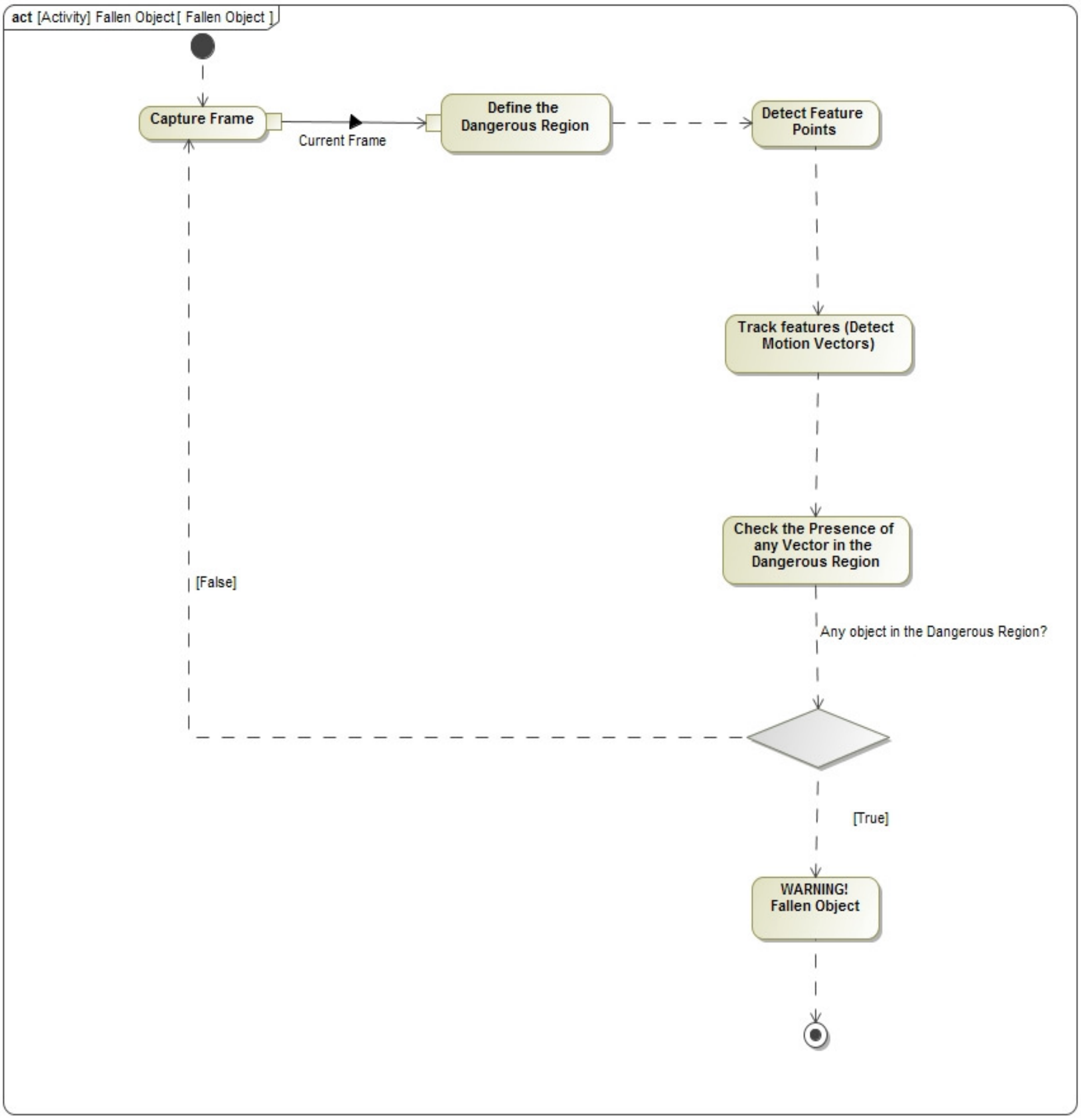

Figure 4.6: Flow diagram for the FOD

The flow chart of the "Fallen Object Detection" is shown in Figure 4.6. On the captured frame from the test video, dangerous region for the passengers is defined. This region is located at the left of the platform. If it is required to be located on a different place then it can be changed.

Figure 4. 7: Personnel entering the track region

The program is initialized to detect the feature points. After detection, feature points are tracked and if there is any change on their location their displacement defined as motion vectors. If there is any intersection between the dangerous region and the motion vectors then it is defined as fallen object detection. In some of the areas which are far from the camera motion vectors could not be tracked. In order to cope with this problem the intensity level of the dangerous region is also tracked. If the change in the intensity level is greater than the threshold value then it is defined as a fallen object.

4.4 PROHIBITED ZONE VIOLATION

In every metro station there are system rooms in tunnel areas. Most of the important components of the system are controlled from these places. The entrances of these rooms are at the end of the platforms. There are two entrances in each platform. (See Figure 4.8)

Figure 4. 8: The two entrances for the tunnel area

Security personnel should be warned if any people enter that area due to the risks about the passenger and the system. Tunnel area is dangerous for the passengers. Also if any damage occurs at the system it may causes failure on the trips. Other possible event is vandalism. People might draw graffiti to the wall of the tunnel. This gives visual and material damage to the metro line.

PZV is detected as shown in Figure 4.9. With the program initialization prohibited zone entry is defined first. The coordinates of the entry given manually. According to the defined region of interest, system detects violations. After tracking the motion vectors, any violation into the entrance area is examined. If any of the feature points get into that area then software gives warning as “Prohibited Zone Violation”.

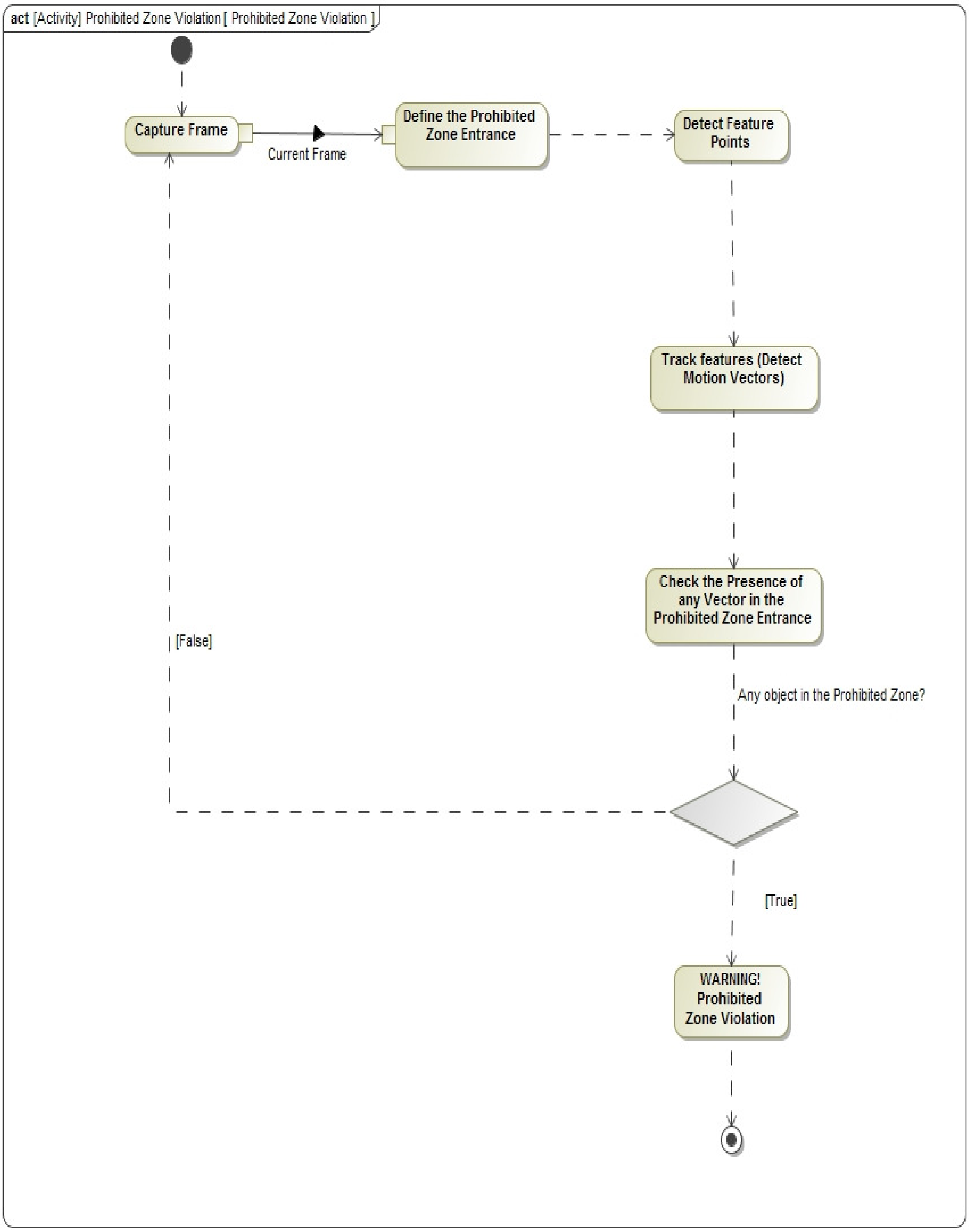

Figure 4.9: Flow chart for the detection of the Prohibited Zone Violation

4.5 OPEN PROHIBITED ZONE ENTRY DOOR DETECTION

Maintenance staff uses the door with a red “no entrance” sign to enter the system rooms and the tunnel region. Sometimes that staff forgets the door open (see Figure 4.10). When the operation time initialized this is a danger factor for the tunnel and the restricted area. For that reason red colored entrance door should be observed continuously. If the door is open a warning for the personnel should be given.

Figure 4. 10: Personnel might forget the restricted area entry open

The algorithm is working as shown in Figure 4.11. With the program initialization a template for the restricted area entry is selected manually and saved. After that this defined template is being searched in every capture continuously. If the template image cannot be found in the image then it means the door is open for the system.

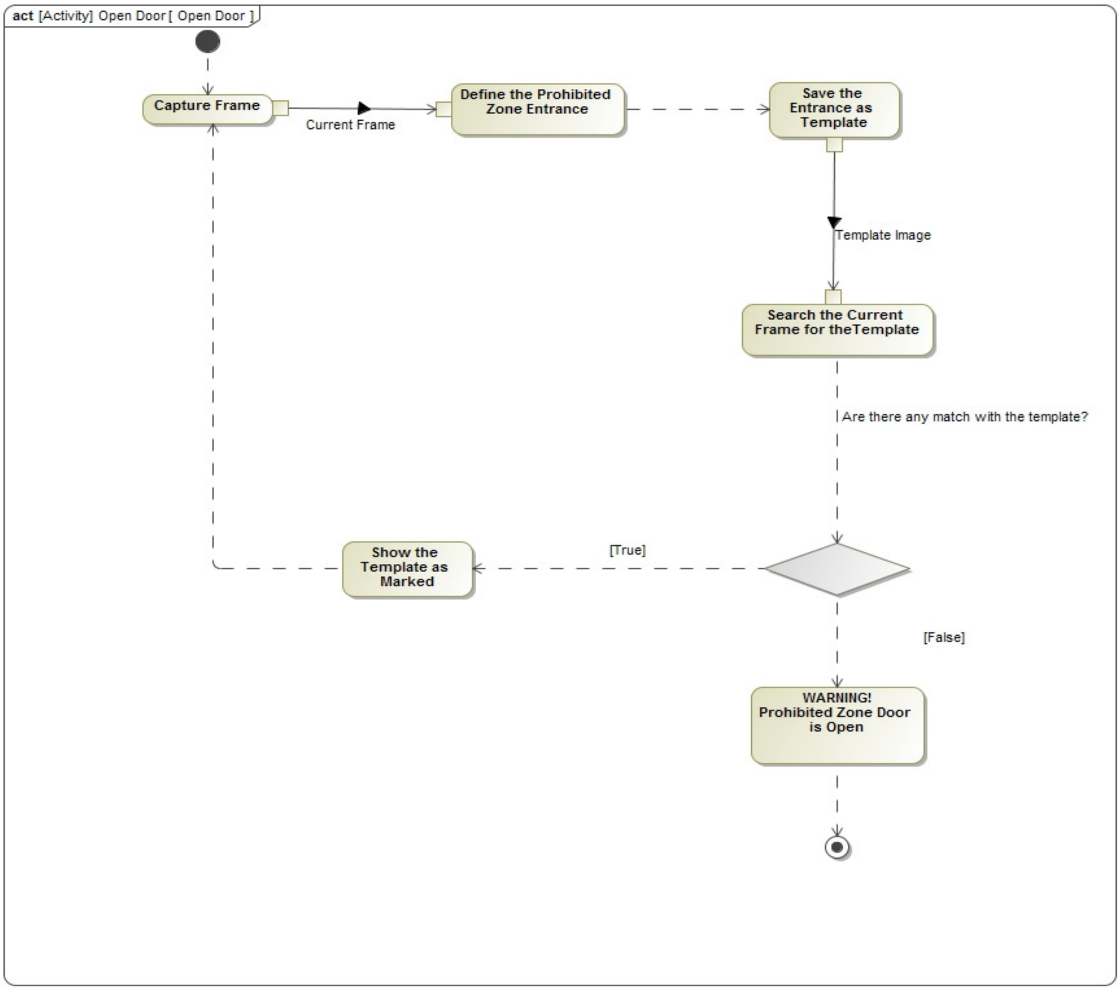

Figure 4.11: Flow chart for the detection of the Open Restricted Area Entry

For the detection of the open restricted area entry OpenCV Template Matching algorithm is used. In this algorithm a template image is searched in the source image. The template image is sliding pixel by pixel through all pixels of the source image. For each location a metric value is stored in a result matrix. There are different matching methods in OpenCV (OpenCV 2.4.0 Doc 2013). In this thesis method with a metric named CV_TM_SQDIFF is used. In this method the result matrix is defined as

$$R(x, y) = \sum_{x',y'} \left(\left(T(x', y') - I(x + x', y + y') \right)^2 \right) \tag{4.3}$$

where R is the result matrix, T is the template image and I is the source image. After finding the result matrix it is normalized, then the minimum and maximum values in the matrix found. According to the CV_TM_SQDIFF method the minimum value is the best matched location with the template image.

5. EXPERIMENTAL RESULTS

In order to test the algorithms videos with a total length of 200 hours are obtained from the Istanbul Metro Control Center. The videos are from the Şişli-Mecidiyeköy station. The resolution of the test videos are 704x288 and the frame rate is 12 frames per second. All of the tests are realized successfully in real-time and most of the algorithms can run at rates higher than 12 frames per second.

All of the algorithms have been designed with SYSML (System Modeling Language). The software development is realized with C++ (C++ 2013) OOP (Object Oriented Programming) Language with QT (QT Creator 2013) Creator IDE (Integrated Development Environment). QT framework (QT 2013) is used for the development of GUI (Graphical User Interface) and the OpenCV library is used for the computer vision implementations.

5.1 BACKGROUND AND FOREGROUND ESTIMATION

In this thesis for detection of the train arrival this simplest background estimation method which is above is used because it was successful enough for this process as shown in Figure 5.1.

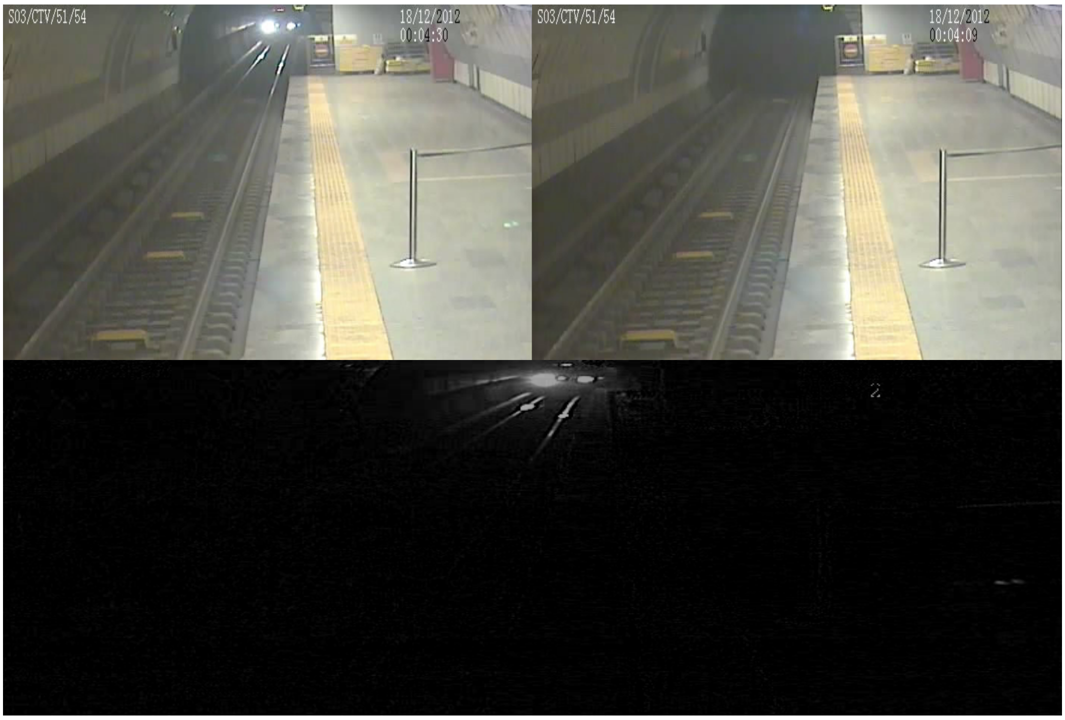

Figure 5. 1: Foreground image obtained by frame differencing only

For different kinds of environments and requirements two different background estimation methods are tested. Since there can be construction works there are changes on the station environment. Therefore background image is renewed every day before the operation starts.

But for the other stations this simple method is not available. There are some illumination differences between different hours and days. Also there can be some personnel or object at the background in that hour. As a result foreground extraction might not be successful enough.

For those reasons dynamic background estimation is used. When algorithm of this method is initialized, first frame is assigned as background image as shown in Figure 5.2.

Figure 5. 2: Background frame

Since this is the first frame there are no foreground object detected as shown in Figure 5.3.

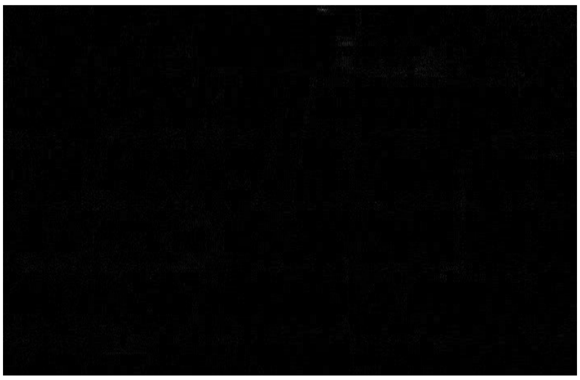

Figure 5. 3: Foreground objects

After several objects enter the scene as shown in Figure 5.4 background should adapt itself with the new environment.

Figure 5. 4: Captured color frame with several objects are included

Background frame is accumulated proportionally with the learning rate and the current frame as shown in Figure 5.5.

Figure 5. 5: Accumulated background image

Foreground objects are gathered by frame differencing between the images shown in Figure 5.4 and Figure 5.5. After differencing, acquired foreground objects are shown in Figure 5.6.

Figure 5. 6: Foreground objects

When there are more crowded environment the results are like shown in Figure 5.7.

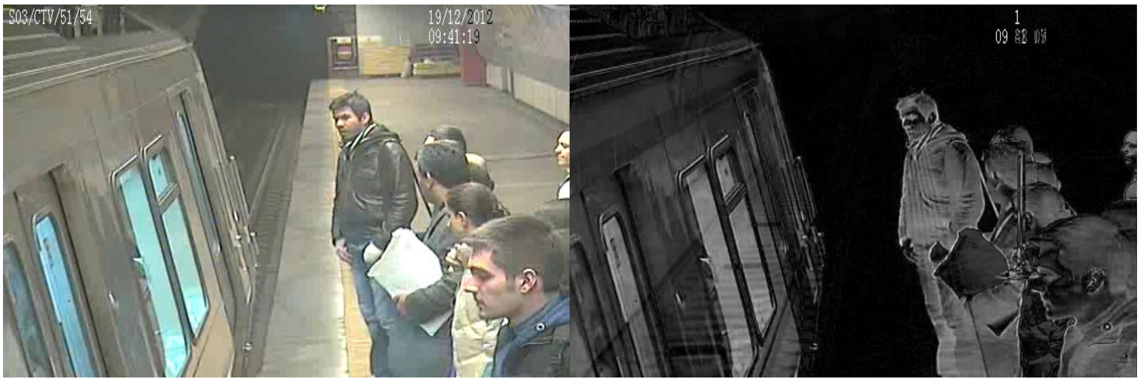

Figure 5. 7: More crowded scene at the left and the foreground objects are at the right

The learning rate for these examples above was 0.01. When the learning rate is increased the background frame adapted quicker as shown in Figure 5.8.

Figure 5. 8: Background image with the learning rate is 0.11 at the left and 0.51 at the right

Then the Mixture of Gaussian method is tested for background estimation. With the same threshold values and 0.01 learning rate applied on the same frames, results are shown in Figure 5.9. It is seen there is faster adaptation to the last frame. The train starts to disappear from the foreground scene. Since it stays stationary for a long time.

Figure 5. 9: A crowded scene at the left and the foreground objects are at the right

The learning rate for these examples above was 0.01. Like previous dynamic background estimation, when the learning rate is increased the background frame adapted quicker.

Figure 5. 10: Foreground image with the learning rate is 0.11 at the left and 0.51 at the right

In Figure 5.10 it is observed that with the learning rate 0.11 renewing of the foreground objects is faster as shown in the left image, but it is more apparent with the value 0.51 as shown in the right picture.

Figure 5. 11: The comparison between the dynamic background estimation method and the MOG method

With the same threshold and the learning rates of 0.11 and 0.51 a comparison between the dynamic background estimation method and the MOG method is shown in Figure

5.11. With MOG background changes and foreground objects disappear quicker than the other method.

At very crowded environments for detecting motion MOG is the best method. But for the stationary regions of the station like the tunnel and the restricted area, dynamic estimation methods or static background estimation on quiet hours which were described in Chapter 2.1 are sufficient.

5.2 TRAIN DETECTION

In order to test the performance of train detection algorithm, a dataset containing 200 hour long videos has been used. On weekdays, there are 225 trips in one direction. That means, 225 trains are passing through the pilot station.

For example on 18th December 2012 it was observed that 229 trains passed. The ground-truth data for the arrival of the train was generated by hand-annotation. All of the arrivals are counted for one day. Also it is proved from the CCR statistics. It is reported from CCR that there 225 trips and also there are 4 empty train passage before starting of the operation.

For this day, using the background estimated in quiet hours, 214 train arrivals were detected by our algorithm. 204 of these detections were true positives and 10 were false positives (see Figure 5.12).

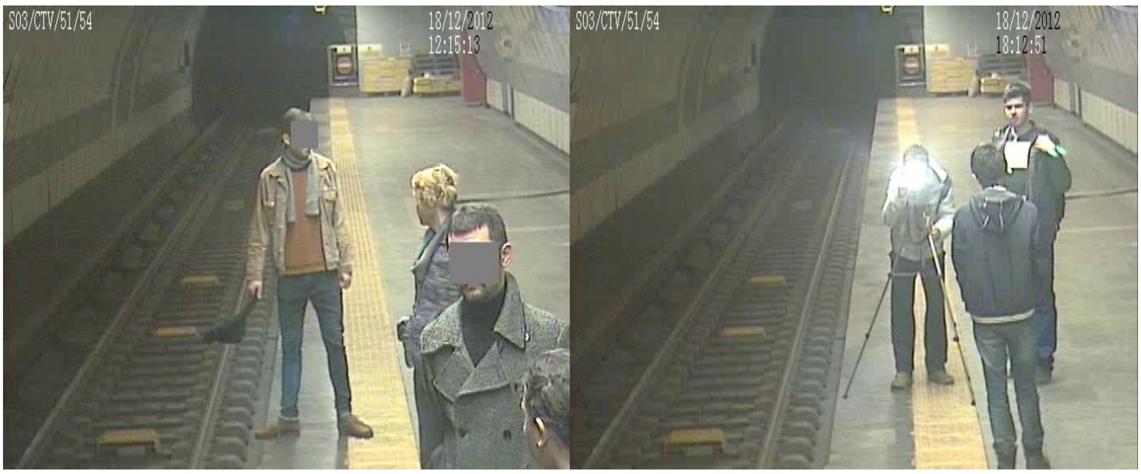

Figure 5. 12: When different day's background image used there are false positives

The configuration about tunnel location was belonging to another day's. It is observed that there are some changes about the angle of the camera. Then when it is implemented with the day which experiment done 228 train detected. When a train enters to the station it is saved as a snapshot image to a definite folder.

Using	Train Detection	# of Detected Trains	F1 Score
Another day's background	214	204	87.8
Current day's background	239	228	97.6

Table 5. 1: Number of detected trains and the Success rate

With the one week's videos an experiment done and the F1 Score for the experiment was 97.6. (see Table 5.1) A sample test results from a rush hour are seen in Figure 5.13.

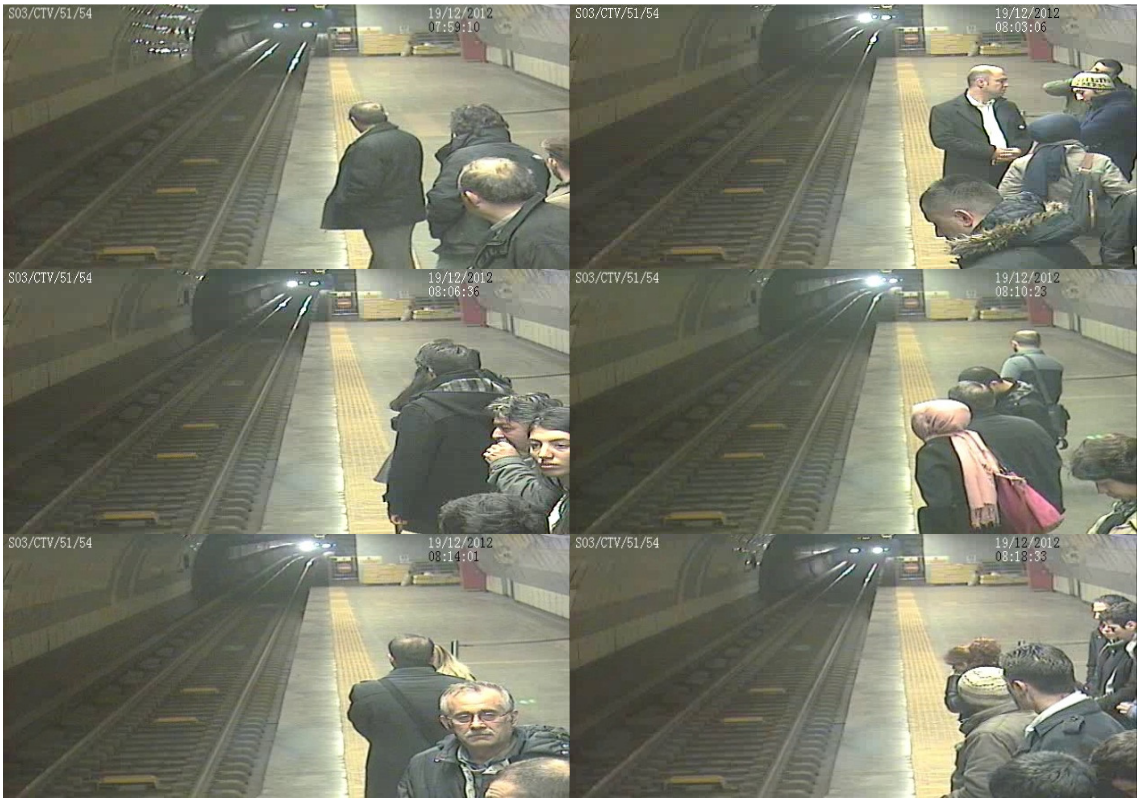

Figure 5. 13: All of the 12 train arrivals in a rush hour have been detected

In rush hours such as 08:00-09:00, there are many frequent train trips (every 4 minutes). The number of train trips between 08:00 – 08:40 is generally around 12. The presented train detection algorithm which uses the quiet hour background estimation method detects all of the train arrivals. (see Figure 5.13)

5.3 SAFETY LINE INTRUSION

Safety line intrusion tests are applied with the videos from the first camera. For the day 18.12.2012 the ground-truth data for the SLI was generated by hand-annotation. There were 243 SLI events. Using the algorithm explained in chapter 4.2 all of the SLI events and 10 false positives have been detected. False positives were because of false results about the train detection. With respect to the presence of train SLI control activated or deactivated. When the train stands in the station then it is assigned as background object and it is assumed that there is no train in the station. Furthermore SLI is activated and when a passenger intrude the safety line then a false positive result saved as true positive as shown in Figure 5.14.

Figure 5. 14: False positives for the SLI detection

Another reason for false alarms is sudden illumination changes as shown in Figure 5.15. After the operation ends most of the lamps are shut down and an important change occurs in the illumination level. Furthermore this causes a change on the background too.

Figure 5. 15: Sudden illumination change at station

A list of SLI detection statistics are given in Table 5.1. There are less SLI event between 00:00-06:00 but the SR (success rate) is low. There are false positives because of the sudden illumination changes. During the operation hours there are less extraordinary states false train arrival detection or illumination change furthermore the SR is at high levels. For the CCR officer it should be at high levels in order not to waste his time with false alarms.

	00:00-06:00	06:00-12:00	12:00-18:00	18:00-24:00	Total
Ground-Truth Data	2	37	91	113	243
True Positive	2	37	91	113	243
False Positive	3	3	2	2	10
Total Detection	5	40	93	115	253
F1 Score	80.0	96.1	98.9	99.1	98.0

Table 5. 2: List of SLI detection statistics by different time intervals

Some of the true positives are shown in Figure 5.16. Intrusion events which are far from the camera are detected successfully.

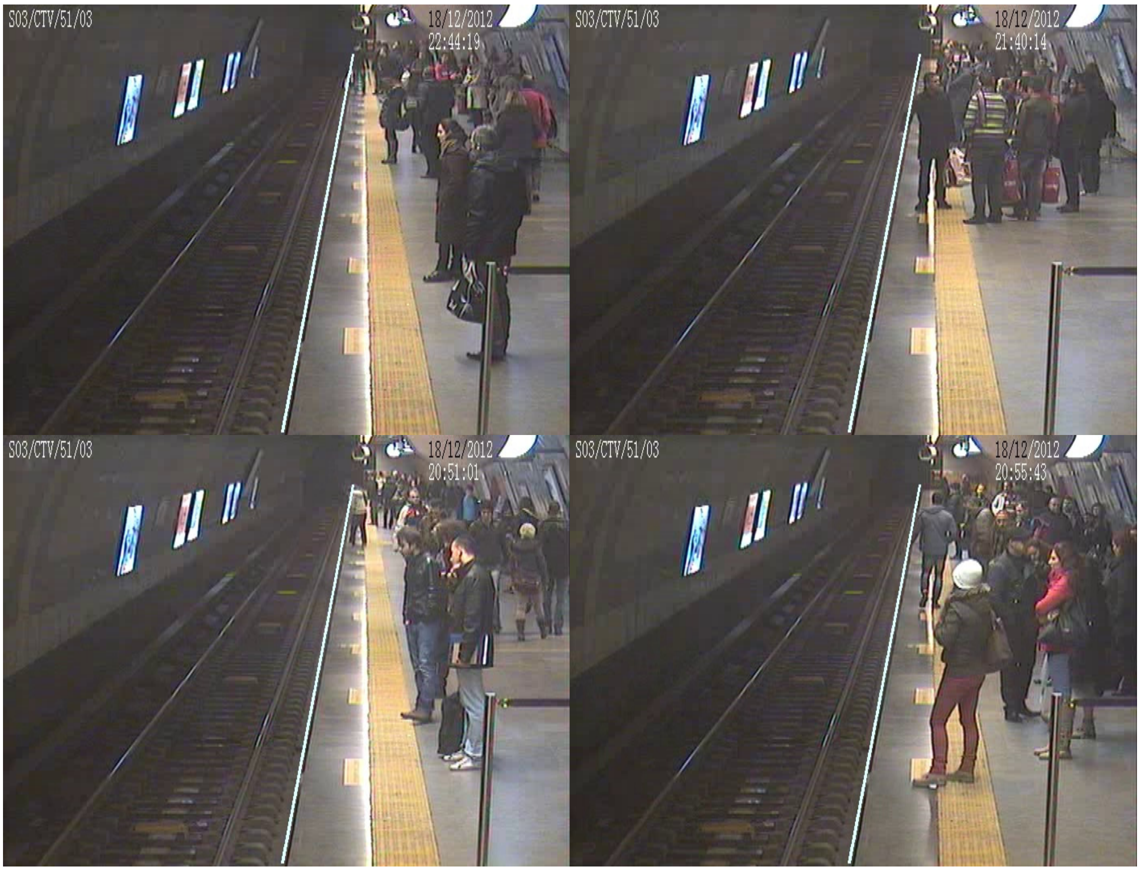

Figure 5. 16: True positive results for the SLI

The software generates warning message as there is a “Safety Line Intrusion” as shown in Figure 5.17.

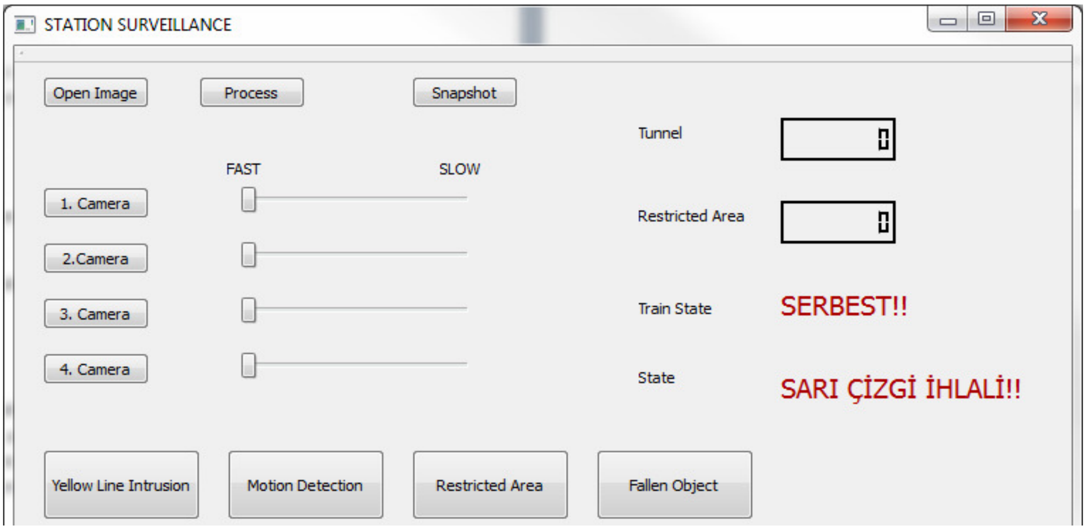

Figure 5. 17: Warning message for the Safety Line Intrusion

5.4 FALLEN OBJECT DETECTION

Fallen Object Detection tests are applied with the videos from the first camera. For the day 18.12.2012 the ground-truth data for the FOD was generated by hand-annotation. There were 3 FOD events.

Figure 5. 18: True positive results for the FOD

With tracking the motion vectors only 2 of the real events are detected (see Figure 5.18). Also there were many false positives. One of them is shown in Figure 5.19. The reasons were the similarity of the intensity levels of the pixels and noise in the scene. In order to cope with this effect likewise SLI algorithm MOG background subtraction applied first. If any foreground object passes beyond the safety line then tracking motion vectors process activated. With this method false positives are discarded.

Figure 5. 19: A false positive result for the FOD

Nevertheless all of the FOD events could not be detected (see Figure 5.20). Undetected event was very far from the first camera and feature cannot be inspected. Therefore fourth camera have to be used for detection and the information should be fused with first camera.

Figure 5. 20: Undetected FOD event

The software generates warning message as there is a "Fallen Object Detection" as shown in Figure 5.21.

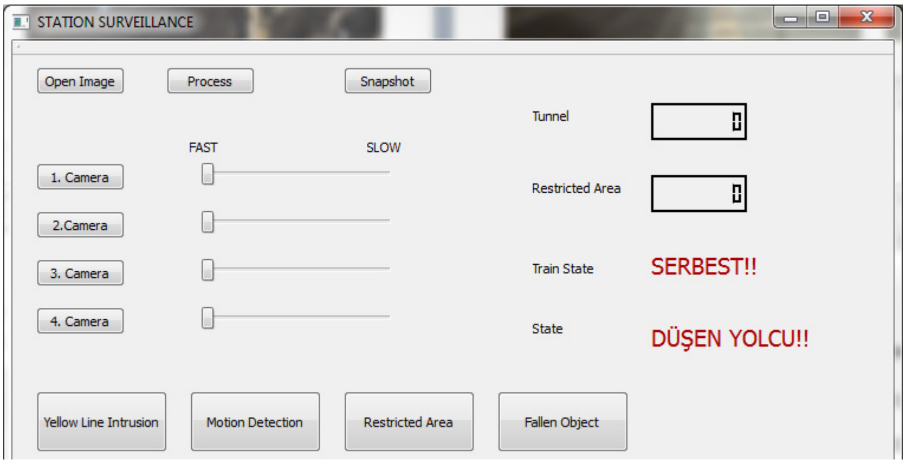

Figure 5. 21: Warning message for the Fallen Object Detection

5.5 PROHIBITED ZONE VIOLATION

Prohibited Zone Violation tests are applied with the videos from the third camera. PZV events are less than SLI or FOD events. Furthermore videos of 18.12.12 – 25.12.12 are used for the tests. For that week the ground-truth data for the PZV was generated by hand-annotation. There were 5 PZV events.

Figure 5. 22: A detected PZV event

With tracking the motion vectors 5 of the real events are detected. One of them is shown in Figure 5.22. But there were many false positives. The reason was the illumination changes and noise in the scene. In order to cope with this effect likewise SLI algorithm MOG background subtraction applied first. If any foreground object enters to the prohibited zone entry it is defined as PZV event. But because of the noise still there were false positives as shown in Figure 5.23.

Figure 5. 23: False positive because of noises in the video

Since there are noise effects on the image there are very little black pixels on the prohibited zone entry so it causes false positives. Finally the mean of the prohibited zone entry intensity is calculated and if it is less than 100 it is defined as PZV event. Thus with this step all of the false positives are discarded. Some of the true positives are shown in Figure 5.24.

Figure 5. 24: True positives for the PZV

The software generates warning when there is a PZV. (see Figure 5.25)

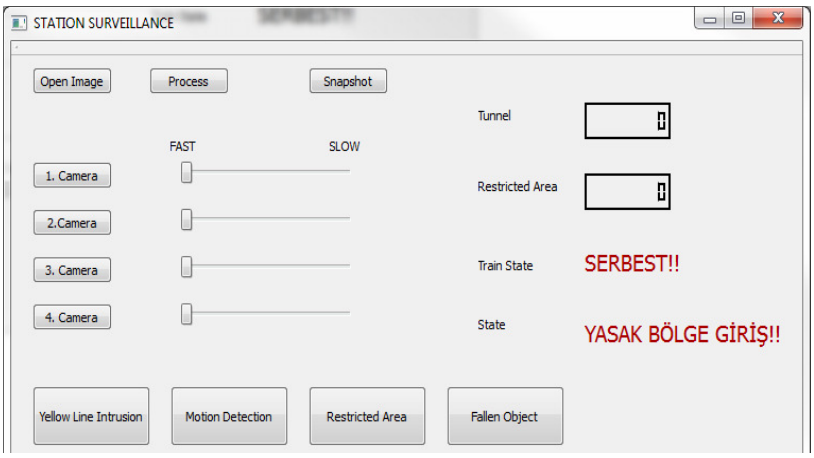

Figure 5. 25: Warning Message for the Prohibited Zone Violation

5.6 OPEN PROHIBITED ZONE ENTRY DOOR DETECTION

Likewise PZV, Open Door Detection (ODD) tests are applied with the videos from the third camera. Timeline for the dataset was 18.12.12 – 25.12.12. For that week the ground-truth data for the ODD was generated by hand-annotation. There were 5 ODD events.

With template matching method prohibited zone entry inspected continuously. When there is no match then it is defined as ODD as shown in Figure 5.26.

Figure 5. 26: No "no entrance" sign can be seen since the door has been left open.

The software generates warning when there is an ODD. (see Figure 5.27)

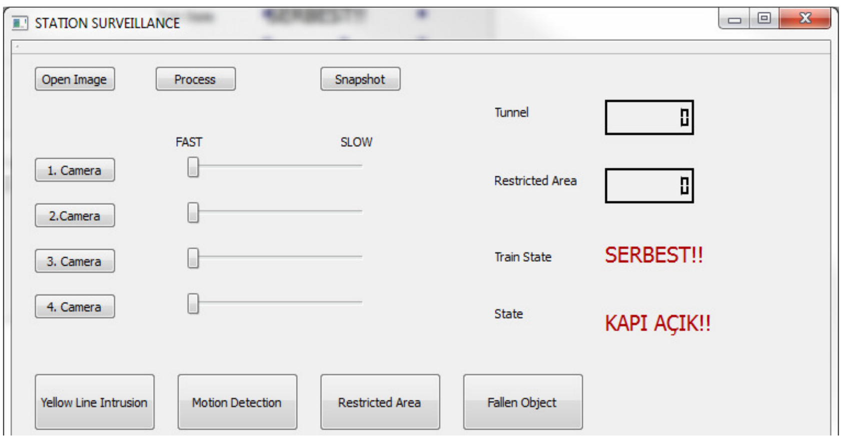

Figure 5. 27: Warning Message for the Open Prohibited Zone Entry Door

6. CONCLUSIONS AND FUTURE WORK

In this thesis a system for automatic detection of several potentially dangerous events which are seen in metro stations has been realized. These potential events are intended to be detected before bad results occurred. The potentially dangerous events considered are: Safety Line Intrusion, Proximity Warning, Fallen Object Detection, Prohibited Zone Violation and Open Prohibited Zone Entry Detection.

First for the estimation of the foreground objects, we focused on the estimation of the background. Then, the foreground objects are detected based on the detected background using two different methods which are manual selection from quiet hour and dynamic modeling with Gaussian mixture model. The system has been tested on 200 hours of video data captured at the Şişli-Mecidiyeköy metro station. The system can successfully detect all of the safety line intrusions and the software gives warning message to the security personnel. There were only two instances for the fallen object detection. Both of them have been detected successfully. Similarly, all of the prohibited zone intrusions and open restricted area entry events have been correctly detected by the system. Unlike previous studies, different events happened at the stations are handled and combined in this study.

A publication about the study named "*Automatic Detection of Potentially Dangerous Events at Metro Stations Using Security Cameras*" is accepted and presented at the *IEEE 20th Signal Processing and Applications Conference (SIU).*

As future work the system might be integrated with the cameras to operate in real-time. This will provide a real-time warning to the security personnel to enable them interfere with the situation without delay. The test environment of the system developed in this thesis was the Şişli-Mecidiyeköy metro station's second platform which serves for the trains of Taksim direction trips. In the future, the system will be tested for all the other 13 stations. The information fusion from all 4 cameras will be enhanced in the future.

REFERENCES

Books

Bradski, G, 2008, “*Learning OpenCV Computer Vision with the OpenCV library*”, Book, O’Reilly Publishing.

Laganiere, R, 2011, “*OpenCV 2 Computer Vision Application Programming Cookbook*”, Book, Packt Publishing.

Tekalp, M, 1995, “*Digital Video Processing*”, Book, Prentice Hall Signal Processing Series

Periodicals

C. Stauffer and W. E. L. Grimson, “Adaptive background mixture models for real-time tracking”, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '99)*, volume 2, pp. 246-252, June 1999.

Carincotte, C.; Naturel, X.; Hick, M.; Odobez, J.-M.; Yao, J.; Bastide, A.; Corbucci, B., "Understanding metro station usage using Closed Circuit TeleVision cameras analysis," *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference* on, vol., no., pp.420, 427, 12-15 Oct. 2008

Chih-Hsiang Wen; Shi-Huang Chen; Po-Chuan Lin; , "A Study of Station Platform Security Using Digital Image Processing Technology," *Computing, Measurement, Control and Sensor Network (CMCSN), 2012 International Conference* on , vol., no., pp.210-213, 7-9 July 2012

Dongwei Qiu; Qinghuai Liang; Qiuping Li; Runze Wu, "Safety Monitoring Information System of Urban Subway," *Information Engineering, 2009. ICIE '09. WASE International Conference* on, vol.2, no., pp.426, 428, 10-11 July 2009

Fujimura, K.; Kamijo, S.; Yoshimitsu, Y.; Naito, T.; , "Behavior understandings based on state transition of postures and locations at railway station," *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference* on , vol., no., pp.91-96, 9-12 Oct. 2011

Li Han; Liu Ji Lin; Yang Wen Ming; He Zhi Wei, "A novel algorithm for background updating of rail scenery images using SDL method," *Signal Processing, 2004. Proceedings. ICSP '04. 2004 7th International Conference* on, vol.2, no., pp.1111, 1114 vol.2, 31 Aug.-4 Sept. 2004

M. A. Dagli, C. E. Erdem, “Automatic Detection of Potentially Dangerous Events at Metro Stations Using Security Cameras”, *IEEE 21th Signal Processing and Applications Conference (SIU)*, Girne, Northern Cyprus, April 2013

Nassu, B.T.; Ukai, M., "A Vision-Based Approach for Rail Extraction and its Application in a Camera Pan–Tilt Control System," *Intelligent Transportation Systems, IEEE Transactions* on , vol.13, no.4, pp.1763,1771, Dec. 2012

Oertel, W.; Dimter, T.; Szoska, D.; "A video-based approach for stationary platform supervision," *Intelligent Transportation Systems, 2002. Proceedings. The IEEE 5th International Conference* on, vol., no., pp. 892- 897, 2002

S.A. Velastin, B.A. Boghossian, and A. Lazzarato, "Detection of Potentially Dangerous Situations involving Crowds using Image Processing", *in Proc. IIA/SOCO*, 1999

Sehchan Oh; Sunghyuk Park; Changmu Lee, "A platform surveillance monitoring system using image processing for passenger safety in railway station," *Control, Automation and Systems, 2007. ICCAS '07.* International Conference on, vol., no., pp.394,398, 17-20 Oct. 2007

Sergio A. Velastin, Boghos A. Boghossian, Maria Alicia Vicencio-Silva, “A motion-based image processing system for detecting potentially dangerous situations in underground railway stations”, *Transportation Research Part C: Emerging Technologies*, Volume 14, Issue 2, April 2006, Pages 96-113, ISSN 0968-090X, 10.1016/j.trc.2006.05.006.

Seyve, C.; "Metro railway security algorithms with real world experience adapted to the RATP dataset," *Advanced Video and Signal Based Surveillance, 2005*. IEEE Conference on, vol., no., pp. 177- 182, 15-16 Sept. 2005

Vazquez, J.; Mazo, M.; Lazaro, J.L.; Luna, C.A.; Urena, J.; Garcia, J.J.; Cabello, J.; Hierrezuelo, L., "Detection of moving objects in railway using vision," *Intelligent Vehicles Symposium, 2004 IEEE* , vol., no., pp.872,875, 14-17 June 2004

W.R. Christopher, A. Azarbayejani, T. Darrell and A. Pentland. “Pfinder: Real-Time Tracking of the Human Body,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 19, pp. 780-785, 1997

Y.Sasaki, N.Hiura. “Development of Image Processing Type Fallen Passenger Detecting System,” *JR-EAST Technical Review Special Edition Paper*, No. 2, pp.66-72, 2003.

Other Publications

C++. 2013, Available from: http://www.cplusplus.com/doc/tutorial/ (cited Jan 11, 2013).

ITC, 2012. Istanbul Transportation Company Annual Report, 2012

Guardian, 2011 http://www.guardian.co.uk/uk/2011/mar/02/cctv-cameras-watching-surveillance (cited Dec 7, 2012)

OpenCV. 2011, Available from: http://sourceforge.net/projects/opencvlibrary/ (cited Jan 11, 2013).

OpenCV 2.4.0 Doc 2013, http://docs.opencv.org/doc/tutorials/imgproc/histograms/

template_matching/template_matching.html (cited Jan 11, 2013)

QT Creator 2013, http://releases.qt-project.org/qtcreator/2.7.0/ (cited Jan 11, 2013).

QT 2013, http://qt-project.org/downloads (cited Jan 11, 2013).

CURRICULUM VITAE

FULL NAME	: Mehmet Ali DAĞLI
ADDRESS	: Bahçeşehir University Faculty of Engineering
BIRTH PLACE / YEAR	: Kilis / 1987
LANGUAGE	: Turkish (native), English
HIGH SCHOOL	: Meram Anatolian High School
B. Sc	: Electronics Engineering, Istanbul Technical University
M. Sc	: Electrical and Electronics Engineering, Bahçeşehir University
NAME OF INSTITUTE	: Natural and Applied Sciences
NAME OF PROGRAM	: Electrical and Electronics Engineering,
WORK EXPERIENCE	: From September 2009 – Currently Istanbul Transportation Co, Electronics Engineer