Republic of Iraq

Ministry of High Education and

Scientific Research

University of Babylon

College of Information Technologies

Department of Software

Video Objects Detection and Shadow Removal Based on Statistical and Spatial Features

A Thesis

Submitted to the Council of the College of Information Technology for Postgraduate Studies of the University of Babylon in Partial Fulfillment of the Requirements for the Degree of Master in Information Technology – Software

By

Mohamed Qasim Mohamed joda

Supervised by

Prof. Dr. Tawfiq Abdul-Khaleq Abbas Al-Asadi

2022 A.D. **1444 A.H.**

بسم الله الرحمن الرحيم

((قالوا سبحانك لا علم لنا إلا ما علمتنا ۖ إنك أنت العليم الحكيم))

صدق الله العظيم

(سورة البقرة : الآية ٣٢)

Dedication

This study is wholeheartedly dedicated to my beloved mother, a strong and gentle woman who taught me to trust in Allah, believe in hard work and that nothing is impossible

To my Father
For supporting and encouraging me to believe in myself and for having faith in me

To my Brothers Ali and Hassan
Thank you for being there for me every time I needed you, the efforts you made, and patience

Mohamed
2022

Acknowledgment

First of all, I thank "**The Greatest Allah"** for granting me strength, willing and patience to accomplish this work.

My grateful deep appreciation and thanks to my supervisor **Prof. Dr. Tawfiq A. Al-Asadi** for his continuous guidance and support, for his patience, encouragement, meaningful and valuable instruction throughout the whole year during working on this thesis.

Finally, I wish to thank my dear family for support and love from the beginning of this journey till this day.

Mohamed Q. Mohamed

2022

Abstract

Object detection in video surveillance systems has great importance in monitoring sensitive security areas such as borders, banks, highways, public places, etc., the advance of large-capacity storage, computing power and availability led to the development in this field. Making the system requires robust and reliable algorithms for the detection of moving objects, shadow removal. In this thesis, moving object detection and shadow removing system is proposed which consists of two stages.

The first stage is moving object detection, this stage includes two steps, the first step is the construction of the background model by selecting N frames to construct the statistical model and spatial models. In the statistical model, the mean and standard deviation for pixels will be calculated. The spatial model consists of a set of histogram models, which are built by using Centre Symmetric Local Binary Pattern (CS-LBP). The second one is foreground detection If the selected pixel value in the current frame is not in the specified range of threshold, that means it is hard to specify if it belongs to the background or foreground. So it will be classified by using the spatial model.

The second stage is shadow detection and removing, shadow removing method is done in two steps, the first one is the selection of candidate shadow pixels by simple shadow detector based on spectral features, the second is the classification of the candidate pixels as either foreground or shadow based on texture correlation in the current frame with the texture in the background models, in addition, the colour feature is used as a complement to texture feature.

Experiments on video challenging sequences illustrate the efficiency of the object detection method compared to the benchmark

methods (Gaussian Mixture Model (GMM), Kernel Density Estimation (KDE) , sliding window and self-regulated learning-based background updating methods for changing detection in videos (SWCD) by increasing precision metrics to (3.6%). In terms of the shadow removing method, the results indicate the efficiency of the proposed method compared to the benchmark methods by increasing precision metrics to (7.6%).

Table of Contents

List of Abbreviations

Symbol	Meaning
BMP	Bitmap
CGI	Computer Generated Imagery
CS-LBP	Center Symmetric Local binary pattern
FN	False Negative
FP	False Positive
fps	Frame per second
GMM	Gaussian Mixture model
HCI	Human Computer Interaction
HSI	Hue, Saturation, Intensity
HSV	Hue, Saturation, Value
KDE	Kernel Density Estimation
LBP	Local Binary Pattern
Pr	Precision
Re	Recall
RGB	Red, Green, Blue
ROI	Region of Interest
SILTP	Scale Invariant Local Ternary Pattern
SWCD	Sliding Window and Self-Regulated Learning-Based Background Updating Methods for Changing Detection in Videos
TN	True Negative
TP	True Positive
YUV	one luma (Y′) and two chrominance (UV) components

List of Figures

List of Tables

CHAPTER ONE

1.1 Introduction

Lately, expecting terrorist acts and giving security to residents at home and abroad has turned into a high need for nations throughout the planet. To accomplish this goal, a lot of data needs to process, decipher and investigated. Several computer technologies can offer incredible help in taking care of these difficulties. Mechanizing the recognition of critical events through cameras needs to utilize the procedures of computer vision which are ready to distinguish objects from the scene, describe them and their conduct, and identify the important events [1].

Video surveillance systems are one of the most important topics of computer vision that tries to detect objects through a series of frames. It also attempts to describe and understand the behaviour of objects by replacing the ancient old way of observing cameras with human operators. The first common step in video surveillance is to identify moving objects in the video sequence and to classify the pixels of these objects into two main classes either background or foreground pixels. Detection of the object is an important and challenging task in several computer vision applications such as vehicle navigation, surveillance and autonomous robot navigation. The detection of objects includes locating objects in the video sequence frames. Many computer vision applications need a moving object detection mechanism either in each frame or when the moving object first appears in the video [2].

Moving object detection is one of the fundamental interests in the field of the video surveillance system. In moving object detection the primary objective is to analyze a video sequence to detect the objects that are in motion with respect to a background scene [3].

Moving object detection is done using three different kinds of approaches: optical flow, temporal differencing and background subtraction. Background subtraction approaches consist of two important steps: the proper generation and updating of a reference background image, and the suitable subtraction between the current image and the background model [4].

Moving objects detection has been used for a wide range of applications like [5]:

1- Detection of moving objects from the video is essential for different surveillance and security applications. This is especially required to trace an abandoned object (e.g. a briefcase) that could be a security threat, locating a vehicle in a stray area in a parking lot, an intruder in a security-sensitive area etc.

2- In Human Computer Interaction (HCI) systems it is used to detect and track different body components (e.g. a moving hand for gesture recognition) for building systems that can interact with a computer in a more natural human-like manner.

3- In gait analysis, this is used for extracting out the moving human silhouette from a video stream.

4- There are various other usages of moving object detection; e.g. human activity recognition and real-time object classification from video.

5- Moving object detection has also found its application in medical image processing, virtual and augmented reality, robotics etc.

The prevalent challenge in Moving object detection is that the shadow will be part of the detected object [6]. Now, recognizing the shadow from the object represents a fundamental task because the shadow causes problems in numerous stages in video surveillance, like analysis of object shape and segmentation. For these reasons, shadow detection and removal is an extremely huge task to increase the execution of video surveillance[7].

Sometimes, the shadow is larger than the actual object. In this case, the rate of misclassification of objects will increase, which leads to reduce the performance of the video surveillance system. There are many states correlated to the performance of the detection of objects that can be affected by the following elements: several objects are overlapped together due to their moving cast shadow in addition to object appearance model reliability will be decreased because of the presence of pixels of shadow which causes the degradation of the performance of video surveillance systems as a whole [8].

1.2 Research Problems

1- The detection and shadow removal stay one of the most common problems of video surveillance .

2- The difficulty level of design and implementation an accurate, robust and high-performance system is remaining a great challenge yet, because depends on how the object is defined to be detected.

3- If only one feature is used such as colour to represent the object, this leads to failure for object detection which is the most important step of the video surveillance system, because is greatly affected by the change of illumination,shadow and many other factors .

4- To tackle this problem more than one feature must be selected, but It should be noted that used different features for each step of the system, this leads to increase the time complexity and storage requirements of the system,so should look for features that can be used in more than one step in the system.

1.3 Aim of the thesis

1- Design and implementation of robustness and reliable object detection and shadow removal system based on local texture feature and chromacity features.

2- The thesis aims to build a system that can face several challenges related to object detection stage which are: illumination (sudden, gradual) change, dynamic background .

3- The thesis aims to build a system that can face several challenges related to shadow removal stage which is : shadow pixels are detectable as foreground pixels since they typically differ significantly from the background, and shadows have the same motion as the objects casting them .

1.4 Contributions

The main contribution of this thesis is:

1- Using the same features (spatial, statistical) in two stages (object detection, shadow removing) leads to increasing the accuracy of the proposed system.

2- Proposed a hybrid object detection method, based on (statistical, temporal, spatial) features in another word, spatial, statistical and temporal features are used in object detection.

3- Proposed a hybrid shadow removing method, based on the simple detector (chromacity features) and followed by finding a correlation between the current frame and background models in terms of center-symmetric local binary pattern (SC-LBP) local texture feature and Quantized Hue spectral feature.

1.5 Literature review

To establish a deep understanding of the problem of this work, a wide study of the up to date researches and technologies should be considered. This is done by establishing a literature review.

1.5.1 Object Detection

First, over recent years, the importance of object detection in surveillance cameras are raised rapidly. Many applications in security, accidents, aids, behaviour analysis, safety, and much more. However, the video signal from the camera, has many challenges that make it a very hard task to analyse the video signal and detect humans. These challenges include mainly the noise and disturbance associated with the video signal. To overcome these types of challenges, many different approaches developed in the last years to tackle these types of problems.

Zhang et al., in (2016) [9] present a new model consisting of two layers based on codebook technique combined with local binary pattern descriptor to solve the dynamic background and illumination variations problem. The block-based on codebook is used to construct the local binary pattern histogram and average values of RGB colour channels which are used to construct the first layer. The stability of the local binary pattern (LBP) features combined with the monotonic illumination variations allows the layer to create an output that is forgiveness about the illumination changes, by creating a block-wise detection. To remove the false-positive error, the pixel-based codebook is used to utilize the accuracy from the result of the number one layer.

A.M. Rahma and N. B. Abd., in (2017) [10] Proposed a method to detect the moving object by using invariant moments. define the region of interest used background subtraction and then applied invariant moments for the region of interest and according to the database information, the moving car will be detected and recognized.

Samera Shams., in (2017) [11] the researchers proposed a model that can automatically detect the distribution of foreground objects that can be found inside the connected regions and this model is wholly depending on the object-based consideration hypothesis and extracts foreground objects straight from the image via feature grouping, in the begging apply this model to the task of foreground object detection and determine integrated foreground by determining the object details rarity. Automatically, the image locations that deviate via the remainder of an image needs to be foreground. To reflect the content of the image, the visual colour and word histograms have been utilized for it. In the end, the foreground has been demonstrated depending on these features outperforms standard information maximization foreground and regular spectral recurring foreground for that task of the detection foreground object.

Muna Gazi., in (2018) [12] The researcher proposed a foreground objects detection approach based on the chrominance and texture features with canny enhance filter. The input is the background image and current image and the output are the detecting foreground objects. The proposed approach consists of three steps: first, the features extracting which are chrominance and texture features from a current and background image. Then, the similarity matching is computed for each feature. Finally, canny filters are used to enhance the results.

Chih-Yang Lin et al., in (2020) [13] proposed a method to model the background without using a threshold value which is extracted from using the approach of bit-planes. This method is proposed because the detection of the mobile objects require an advanced threshold value to detect the illumination changes in the background model. This new proposed approach will employ and optimize the colour property by improving the spatial and temporal.

S. Mohanty and S. Rup., in (2020) [14] proposed an enhanced approach that solves the problems of illumination changes, non-static background, low contrast, and noisy sequences. To detect the temporal motion and spatial textural features from videos frames, the system uses a spatial-temporal local binary pattern (STLBP) technique. To minimize the false error percentage, and make the algorithm detect the object in the foreground more precise, the values of learning rate and threshold are adjustable. This replaced the classical way of setting a constant value for the learning rate.

Table (1.1) illustrate the characteristic of methods that mention in literature review of bject detectin as well as the characteristic of proposed method.

Table (1.1) Comparison between proposed method and methods in literature review of object detection

Method	**Robust to**			**Using of**		
	Illumination variations	Dynamic background	Noise	Spectral information	Spatial information	Temporal information
Zhang et al.[9]	✓	✓		✓	✓	
A.M. Rahma and N. B. Abd.[10]		✓			✓	✓
Samera Shams.[11]		✓			✓	
Muna Gazi.[11]		✓	✓	✓	✓	
Chih-Yang Lin et al.[12]	✓	✓			✓	✓
S. Mohanty and S. Rup.[13]	✓		✓	✓	✓	✓
Proposed method	✓	✓	✓	✓	✓	✓

1.5.2 Shadow Detection and removal

Shadow is usually formed when a non-transparent object prevents the rays of the light from continuing in its path. Although it's a significant impact on the image as seen by humans, the shadow itself doesn't have any wavelength properties. Therefore, it is hard to measure it. To understand the shadows deeper, it is divided into two types.

The first one is Self-shadow. Which is the change in the illumination of the light that lays on the object itself.

The second one is Cast-shadow, which is the change in the illumination of the light that lays on the surrounding area of the object. The Cast-shadow is divided further into two types: Umbra and Penumbra.

Umbra is the central part of the shadow, which is the darker part of the Cast-shadow. On the other hand, the penumbra is the less dark part of the Cast-shadow, that which lays surrounding the Umbra reign usually [15].

Various methodologies were created somewhat recently to handle shadow detection and removal.

Hu et al., in (2016) [16] proposed a technique based on the hypothesis that the shadow region is darker than the corresponding reference but have similar texture and chromacity. The method uses both Hue, saturation, and value (HSV) and RGB colour spaces to obtain spectral information and merge two texture features to extract moving cast shadows. Firstly, candidate shadow regions are detected by spectral analysis. Secondly, the technique is based on the proposed extended local ternary pattern, enhanced edge information and HSV colour space analysis to achieve three shadow detectors which are used to vote to obtain the final shadows, respectively.

D. Carmen and R. Cajote., in (2016) [17] proposed a shadow detection algorithm that combines the usage of texture information using interior edges of objects, HSV colour space information and low-level post-processing. The background and foreground frames are converted to grayscale and their individual HSV colour space channels are equivalent. The candidate shadow pixels are then extracted using an improved form of HSV colour space thresholding which is based on the created background image. An additional HSV threshold independent of the background image, the foreground hue to value ratio is added to improve the shadow detection rate in cases where the background may be changeable.

A. Abdusalomov and T. Whangbo., in (2019) [18] proposed an approach for accurately removing shadows on modern buildings in the presence of a moving object in the scene. the proposed system combined contrast enhancement technique is applied to the input frame sequences. After obtaining suitable enhanced images, segmentation and noise removal filtering are applied to create a foreground mask of the possible candidate moving object shadow regions. Subsequently, geometry and colour information is utilized to remove detected shadow pixels that incorrectly include the foreground mask.

H. Najeeb and R. Ghani., in (2020) [19] proposed a hybrid method that combines two colour spaces (HSV and YCbCr) for the detection of shadows in HSV and YCbCr compared with e thresholds which are selected through trial and error. Shadow is detected when the pixel in both HSV and are equal to thresholds, otherwise, it is a foreground pixel.

Minghu Wu et al., in (2020) [20] proposed a robust shadow elimination method based on colour and texture features. in the beginning, the researchers carry out the analysis in the (HSV) colour space to effectively remove the shadow of the moving target in the video. then, they combine the local variance and OTSU method that is used to perform automatic image thresholding to overcome the disadvantage of the moving target detection error caused by the unstable brightness ratio threshold.

1.6 Layout of Thesis

In addition to Chapter One, the remaining parts of this thesis consist of the following chapters:

- **Chapter Two: Entitled "Theoretical background".**
  This chapter describes the methods which are used for the main steps of the video surveillance system.
- **Chapter Three: Entitled " The Proposed System"**
  In this chapter, the proposed system design and implementation steps are given. The proposed system parts are described in detail.
- **Chapter Four: Entitled "Experimental Results"**
  This chapter is dedicated to presenting the results of each step of the proposed system, as well as comparing those results with the results of benchmark methods.
- **Chapter Five "Conclusions and Future Work"**
  This chapter will discuss any further works related to this project.

CHAPTER TWO

2.1 Introduction

This chapter provides a theoretical background with respect to the research work described in this thesis. It consists of three main parts. The first one explains the video surveillance system and its stages. The second describes the background modelling and foreground detection and the main problems which face it. The third is shadow removal techniques.

2.2 Video surveillance systems

A movie or video is a collection of still pictures or frames linked to audio data. A frame is a single picture or still photo that appears as part of a bigger movie or video. To generate what appears to be seamless video, many discrete pictures are played in rapid succession. As a result, the video is made up of static pictures or frames. Images are collected at a set rate (25-30 fps or even 60 fps) in the movie, depending on the camera's capture speed. A camera with 30fps captures 30 frames per second to create a video stream. These pictures may be extracted from a video stream and processed in the same manner that an image is [٢١].

In many contexts, video surveillance devices are prevalent. Banks, airports, penal facilities, and casinos all used video monitoring to assure security. To improve public security, government institutions, schools, colleges, and even companies have recently turned to video monitoring. The deployment of a large number of security surveillance cameras is both economically and technically feasible, thanks to the widespread availability of low-cost cameras and high-speed broadband wireless networks [٢٢].

A video surveillance system is an important aspect of the research process. In most video surveillance systems, object detection is based on background estimate and subtraction. Object detection is the initial stage in finding instances of semantic objects of a certain class in a video sequence, such as persons, buildings, automobiles, and so on. [٢٣].

Even when a person is taught and assigned the duty of visual observation, human visual attention falls below acceptable levels when it comes to danger detection [٢٤]. Video analysis techniques, on the other hand, may be used to develop intelligent surveillance systems that aid human operators in detecting hazards [٢٥].

The standard surveillance applications consist of two main stages as shown in figure (2.1) moving object detection, shadow removal. Multimedia systems can give surveillance coverage over a vast area, ensure object visibility over a wide range of depths, and be used to distinguish between occlusion and obscurity. People must collaborate in the multi-view system to follow events of special interest, and occurrences of particular interest must be recorded across the scene [٢٦].

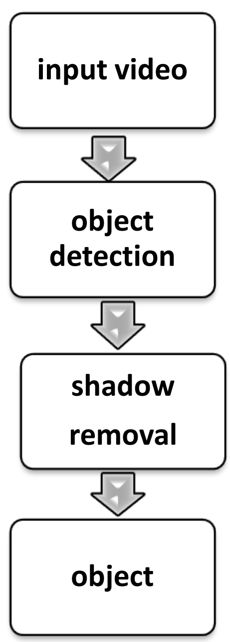


Figure (2.1): Block diagram of standard video surveillance system

2.3 Detection of objects

The stage of object detection consists of two steps, background modelling and detection of moving objects. The first step in object detection is to extract moving objects from the video stream. Most methods rely on background subtraction technology by modelling the background, this lead to the detection of moving objects becoming more efficient and robustness. The main idea is to maintain a parametric or nonparametric model for the background [٢٧]. For a simple example, for each pixels (N) value in the (N) frame, computing the mean by using eqution (2.1) and standard deviation using eqution (2.2) for pixels[28]

$$\mu = \frac{\sum_{i=1}^{n} Xi}{n} \tag{2.1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(Xi-\mu)^2}{n}} \tag{2.2}$$

2.3.1 Essential features for object detection stage

The following categories can be used to categorize features based on their essential properties:

2.3.1.1 Spectral features

these features can be obtained directly from the images such as intensity and colour. Spectral features are possible to identify changes easily if the difference in colour value between the background and foreground is large, so the spectral features lead to detect (false positive) FP due to illumination variations and to detect (false negative) FN when the object has the same colour with background, this is called camouflage problem [٢٩].

Colour characteristics are often weak in the face of changing lighting and shadows created by moving objects. The following are the most often utilized colour attributes in various colour spaces:

- Color space RGB Due to its immediate availability from the sensor or camera, the RGB colour space is the most commonly utilized characteristic. Each pixel's Red, Green, and Blue channels are commonly measured with an 8-bit resolution. However, the RGB colour space has a significant flaw: its three components are interdependent, making it more sensitive to variations in lighting [29]. So a most recent work converted the colour image and video to grey-scale. The images that converted to grey-scale are a mixture of two colours black and white and have many shades of grey in between. Grey-scale images present of only one colour, one refers to mono and colour refers to chrome so, these combinations of greyscale images are called monochrome. The conversion from the colour image to a grey-scale image is done by many methods where one of these methods is the average method that averages the values of Red, Green, and Blue, as shown in equation (2.3).

$$\text{Gray} \ (R+G+B) / 3 \tag{2.3}$$

The advanced version of the average method is called the luminosity method. It also averages the values but adds a weight value to account for human understanding where the weight of the Green value is considered as the largest weight value as shown in Equation[30] (2.4).

$$\text{Gray}=(0.29\times R) + (0.59\times G) + (0.11 \times B) \tag{2.4}$$

- YUV colour space: Because the YUV colour space separates luminance and chroma, it is better for enhancing the model's resilience to variations in light [٣١].

- HSI colour space: In the sense that brightness, or intensity, is separated from the base colour, HSI colour space is closer to human colour interpretation. Polar coordinates are used by HSI [٣١].

- HSV colour space: The HSV colour space is used to enhance shadow and object identification, categorizing shadows as pixels with about the same hue and saturation values as the background but lower brightness. Under low illumination intensity conditions, the HSV colour space proved ideal [32].

The conversion of RGB vectors to the HSV space is commonly denoted as [٣٢]:

$$V = \max(B + G + R) \tag{2.5}$$

$$S = 1 - \frac{3}{B+G+R} \min(B, G, R) \tag{2.6}$$

$$H = \begin{cases} \theta & \text{if } G \geq B \\ 360^0 - \theta & \text{if } G < B \end{cases} \tag{2.7}$$

In which:

$$\theta = \cos^{-1}\left\{\frac{\frac{1}{2}\left((R-G)+(R-B)\right)}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right\} \tag{2.8}$$

2.3.1.2 Spatial features

Edge and texture characteristics are examples of spatial features. These qualities aid in the detection of foreground objects that blend in with the backdrop as well as the suppression of shadows. Spatial features, on the other hand, are not relevant at the pixel level to non-stationary backdrop objects since the related spatial qualities change with time. Texture characteristics are one of the most essential spatial features. They allow for robustness in the presence of shadows and gradual light changes, as well as in the presence of dynamic backdrops. Due to textures caused by local illumination effects such as cast shadows, texture characteristics might produce incorrect detection .Furthermore, a texture-only method may result in identification mistakes in areas with blank textures and varied textures [٣١].

Generally, there are several types of local texture features; in the next sections the important methods to extract local texture features will be explained:

2.3.1.2.1 Local Binary Pattern (LBP)

The LBP is a powerful textured characterization technique that labels every pixel in an image block by comparing it to the image block's middle point and then encoding the outcome in code

$$\text{LBP}\ (i_c, j_c) = \sum_{p=0}^{7} s(\, v_{p\,-}\ v_c\ )2^p \tag{2.8}$$

Where:

v_c: Value of the centre pixel (i_c, j_c)

v_p: Value of the neighbourhood pixels

$x = v_p - v_c$

$$s\ (x) = \begin{cases} 1 & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.9}$$

The original LBP method works with 3*3 block size of pixels which works with square neighbours as shown in figure (2.2) [٣٣].

example

6	5	2
7	6	1
9	8	7

threshold

1	0	0
1		0
1	1	1

weights

1	2	4
128		8
64	32	16

Pattern= 11110001 LBP = 1+16+32+64+128= 241

Figure (2.2): An example for calculating the original LBP code.

2.3.1.2.2 Center - Symmetric Local Binary Pattern (CS-LBP)

The CS-LBP operator produces more compact binary patterns compared with the original LBP descriptor as explained in figure (2.3), the CS-LBP operator produces 16 binary patterns whereas the LBP descriptor produces 265 binary patterns.

neighbourhood

n5	**n6**	**n7**
n4	**nc**	**n0**
n3	**n2**	**n1**

LPB =

$S(n0-nc)2^0+$
$S(n1-nc)2^1+$
$S(n2-nc)2^2+$
$S(n3-nc)2^3+$
$S(n4-nc)2^4+$
$S(n5-nc)2^5+$
$S(n6-nc)2^6+$
$S(n7-nc)2^7$

CS -LPB =

$S(n0-n4)2^0+$
$S(n1-n5)2^1+$
$S(n2-n6)2^2+$
$S(n3-n7)2^3$

Figure (2.3): An example of CS-LBP operator with 8 neighbouring pixels

CS-LBP descriptor is more powerful on the flat area in the image by applying small threshold T for grey level differences, the formula of CS-LBP descriptor shows in the equation below:

$$\mathrm{CS-LBP}_{R,P,T}(i,j) = \sum_{i=0}^{(P/_2)-1} s\left(p_i - (p_{i+(P/_2)})\right) 2^i \qquad (2.10)$$

$$X = p_i - (p_{i+(P/_2)})$$

$$s(x) = \begin{cases} 1 & x > T \\ 0 & \text{otherwise} \end{cases} \qquad (2.11)$$

Where Pi and (pi+(p /₂)) are center symmetric pairs of pixels that are equally spaced on a circle of radius R [٣٤].

2.3.1.2.3 Scale Invariant Local Ternary Pattern (SILTP)

Local Ternary Pattern (LTP) descriptor has robustness against the noise due to using a small acceptance range but it is not invariant to scale of intensity values. The property of intensity scale-invariant is very important because illumination changes either local or global. For the previous reason, the SILTP descriptor is proposed to handle illumination change and noise, the descriptor illustrates in figure (2.4).

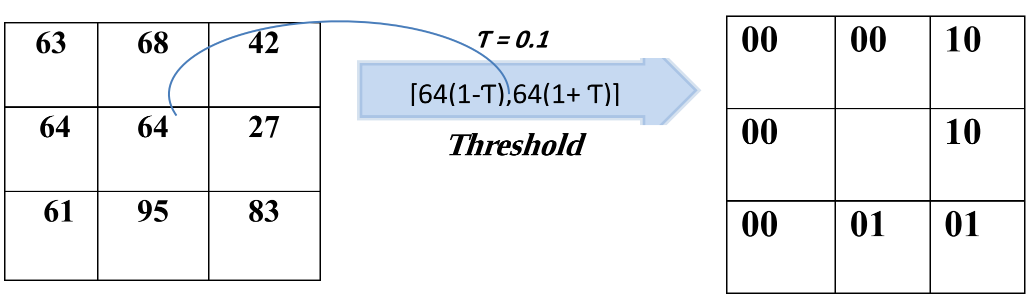


Figure (2.4): An example of SILTP descriptor with 8 neighbouring pixels

Scale-Invariant Local Ternary pattern is computed as follows:

$$SILTP_{P,R,T}(i_c, j_c) = \bigoplus_{p=0}^{P-1} s_T\left(v_c, v_p\right) \tag{2.12}$$

Where v_c represents the grey level of the centre pixel and v_p represents the grey level of pixels evenly spaced around a circle of radius R. $\bigoplus$ refers to concatenation operator, T is a scale factor for comparing ranges, and s_T is a piecewise function that is defined as follows [٣٥]:

$$s\left(v_c, v_p\right) = \begin{cases} 01 & v_p > (T+1)v_c \\ 10 & v_p < (T-1)v_c \\ 00 & otherwise \end{cases} \tag{2.13}$$

2.3.1.3 Temporal features

The motion between subsequent video frames is referred to as temporal characteristics. To cope with extraneous motion in the background, one approach to obtaining temporal characteristics is to assess the consistency of optical flow. To get the optimum performance, the most resilient optical flow approach should be used, according to theory. However, the majority of optical flow methods are computationally inefficient. After that, three different ways of introducing temporal aspects are utilized [٣١]:

1. Those that are solely dependent on the difference between two successive frames. The backdrop model is then calculated just in the scene's stationary portions.
2. Optical flow (calculated across all pixels) is utilized to detect moving objects. The backdrop model is only calculated in stationary regions, like in earlier techniques.

3. After foreground identification, optical flow is solely computed on moving regions. Optical flow helps the program to discriminate between irrelevant moving regions and moving objects in this example.

2.3.2 Background modelling

Background modelling has emerged as an important and efficient method for video surveillance. Since it is practically efficient and important to detect moving objects in videos. The various method has been proposed to the goal of foreground detection, however; most of them are still suffering from some challenges. These challenges include dynamic camera motion, occlusion, cluttered background, and variance of illumination [٣٦].

a. current frame

b. background model

c. background subtraction

Figure (2.5): Background Modeling and Foreground Detection

Background modelling importance comes from its role in detecting foreground objects. Since it is important to detect the background first, to detect the objects that do not belong to the background (i.e. foreground objects) [37].

Many researchers over decades proposed many approaches to detect the background. Each approach has its advantages and disadvantages. Almost all the approaches follow a similar guide to detect the background and distinguish the foreground objects. They first start by using a few frames at the beginning to build the background model. Then they compare the new frames with the background model. A foreground object is defined as anything that does not belong in the background. And, if necessary, the background model is tweaked to fit the new frames [38].

2.3.2.1 Background models initialization

Background model initialization is the process of setting the model of the background without being a foreground object. Therefore, it is the "initial state" of the background model. This is an important step in background modelling, since it is the backbone of the whole process of background modelling, and foreground object detecting. However, it receives poor attention in the literature on digital image processing and computer vision. Where most publications are focused on the problems of representation and adaptation. These early frames that are used in the initialization process must be "empty". In other words, there are no foreground objects in it. Sometimes there are environments, which makes it very difficult to get such empty frames, for example in crowded airports or train stations. Furthermore, some environmental factors may affect the process of getting early empty frames, such as strong illumination change. All these reasons lead to the point of using initial frames with no foreground objects in them [39].

2.3.2.2 Background models maintenance

Background model maintenance is the process of making a stable model of the background. So the model will only consider the true background, without considering the foreground objects. And will not be powerless to any type of changes. The perfect background model will mostly have these properties:

1. The moving objects: The first and obvious criteria of good background, it should not consider any moving object as part of the background model.

2. Illumination changes: Good background model should not be vulnerable to the changes in the illumination. This could happen for various reasons, the most popular one is when the daylight is changing during the day.

3. Waving objects: Some objects which are part of the background have some movement that may confuse the background model. A popular example is waving trees. The trees are usually waving due to the movement of the wind. A good background model should still keep detecting the tree as a background object even with its movement.

4. Shadows: Shadows are a challenging problem. Sometimes when a foreground object is moving, it overthrows a shadow. The shadow is moving with the foreground object. Making changes in the colour of the background. Some models may consider the shadow as part of the foreground object.

In general, there is no perfect background modelling. Up to the present, all the suggested methods and algorithms suffer from some kind of flaws [40].

2.3.2.3 Background modelling methods

Many academics have proposed a wide range of methodologies focusing on object recognition from a video stream. The majority of them employ numerous procedures, and there are overlaps and crossovers between various methodologies.

2.3.2.3.1 Frame differencing

Or sometimes called (Background subtracting), it is a real-time foreground object detection. And it is widely used in both scientific research and industry field. The general idea is to consternate on the interesting places of the video. Which are the moving objects. The input of this approach is the video (sequence of image frames) and optimally will output a signal with the moving objects are detected and highlighted[41].

2.3.2.3.2 Mixture of Gaussians

A Mixture of Gaussians is a statistical-based algorithm. That works by making *k* Gaussian distributions for the initial model of the background model. The distributions are arranged based on the ratio between their peak amplitude and standard deviation. After that, a similar examination is made. When a pixel is characterized, the best identification background distribution is refreshed with the value of the current pixel if the last is viewed as background. Something else, the more vulnerable distribution is supplanted with a new Gaussian dependent on the noticed pixel [42] [43].

2.3.2.3.3 Kernel Density Estimation

Estimation of Kernel Density Is a non-parametric background model with a background subtraction method. The model can handle scenes with a crowded background that isn't fully static but has minor movements like tree branches and plants. Based on a sample of intensity values for each pixel, the model calculates the likelihood of seeing pixel intensity values. The model is particularly sensitive to moving objects since it responds fast to changes in the scene [44][45].

In modelling, the size of the picture element used for interpreting required aspects that accurately describe its properties is critical. A feature value at a specific pixel can be determined by the feature value at that pixel or by the feature values in a predetermined neighbourhood (block or cluster).

A. Pixel level: at the pixel level, the process of involving a certain frame location and pixel coordinates. This level of processing is used in intensity and colour features. It is used to compute statistical measurements of spatial and temporal frame neighbourhoods to consider all these spatial and temporal features. And the calculated statistical value is assigned to the pixel in the centre [31].

B. Block level: in the block level, instead of using single-pixel as in the previous case, here a block of pixels, with *m*n* size will be considered. This will lead to the process a wider area of pixels, so optimally will capture a whole feature. This method will also capture all the spatial and temporal properties of the features in the frame. In practice, these block-based properties might be applied to a block's centre pixel or the whole block[31].

C. Cluster level: sometimes called (Reign based features), which try to capture the feature of an entire element. The advantage of the cluster wise approach is it has less false detection compared with the block-wise approach since in cluster wise approach the foreground objects are detected based on pixel-level accuracy [31].

2.3.3 Object detection challenges

Object detection has many problems that are not completely solved yet. The problem can be formulated as to how to detect and define the object in the image series. Typically, common challenges of object detection in the context of a video surveillance system can be explained as follows [46]:

A. Illumination Changes:

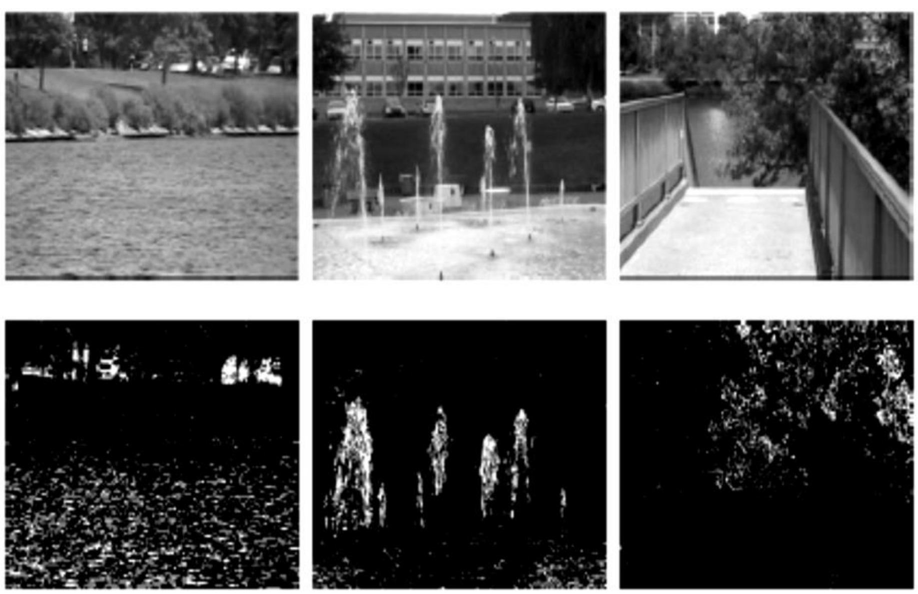
The illumination changes problem happens where there are different brightness values for the same environment during different times as shown in figure (2.6). This will lead to the background having a different appearance. When attempting to identify the background, this might result in a false positive detection mistake. There are two types of illumination changes. The first one is the gradual illumination, where the changes of light brightness in the background are slow. This happens usually in the outdoor scene. Where the sunlight differs gradually during the day. The second illumination change type is the sudden change. This happened where a bright source of light appear suddenly, such as a car light, or switching bulb light on or off.

a. b. c. d.

Figure (2.6): From the left to the right: a) Low illumination. b) Moderate illumination c) High illumination. d) Foreground mask.

B. Dynamic backgrounds:

Sometimes there is a non-static part of the environment. For example, if the background scene has a moving river in it, a fountain, or trees that move during the windstorms. These are all considered part of the background. However since they are moving, it will be difficult to address them as background. Figure (2.7) show an example of dynamic backgrounds.



a. *b.* *c.*

Figure (2.7): Example Water surface, fountain, and waving trees are examples of dynamic background.

C. Occlusion:

Occlusion appears when two objects are passing each other. And one object is positioned behind the other. This will lead to difficult detection and tracking of the hidden object. As shown in figure (2.8).

Figure (2.8): Example of two objects are Occlusion

D.Presence of shadows:

shadows appear when there is a source of light that makes blackish pixels in the image. The presence of shadows makes the process of subtracting the objects from the background challenging task. Since it requires a further process to address this problem. As shown in figure (2.9).

Figure (2.9): Presence of Shadows

E. Motion of the camera:

since the camera may be moving or unstable, it will be harder for classical image processing techniques to distinguish between the foreground objects and background as shown in figure (2.10). Since the jitter magnitude of the videos will be different.

Figure (2.10): Camera jitter

F. Video noise:

Video signal usually contains noises derived from sources such as sensor-generated noises, and compression side effects problems. These types of noise will lead to challenges in background subtraction.

G. Sleeping foreground object:

A static or motionless object should be not recognized as part of the background. This challenging for the system to recognize if the object is part of the background or its part of the foreground objects. Especially if the initial frame of the video is consisting of the motionless foreground object. The system will consider these objects as part of the background since they are not moving.

2.4 Shadow removal

The identification and effective removal of a shadow created by an occluded light source, as well as a photo-realistic restoration of the picture contents, are key computer vision tasks [47].

Shadows are one of the main reasons for moving objects incorrect detection. Since it adds problems in the techniques of object segmentation, and object shape distortion. Moreover, it is one of the main reasons for background incorrect detection since it leads to the loss of the background texture and false connectivity of independent blobs. Figure (2.11) show examples of shadow detection and removing [48].

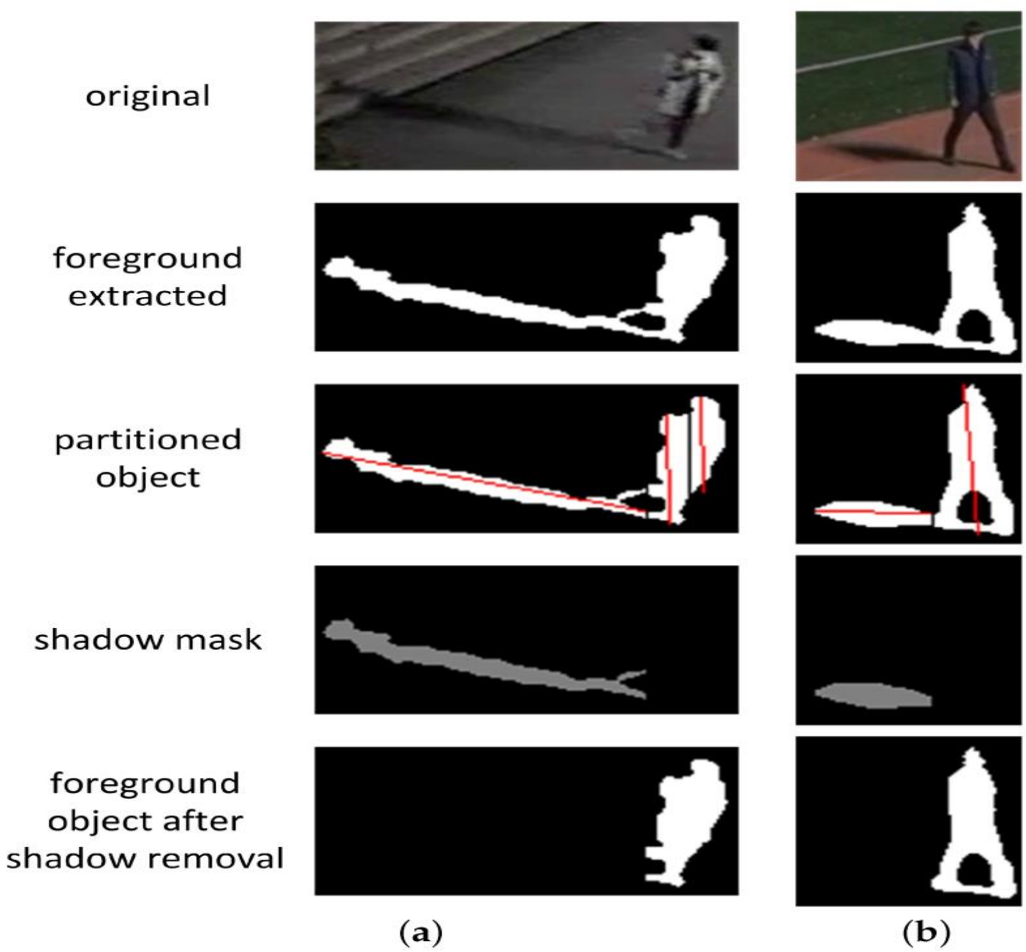


Figure (2.11): Examples (a and b) of shadow detection and removing

2.4.1 Shadow detection techniques

Many shadow detection techniques have been proposed in the last decade. In particular, in the fields of static photos, satellite photographs, and films. However; in the area of detecting shadows of moving objects, the approaches are still working well for well-defined objects such as vehicles but are difficult to use for non-rigid things like human beings. Furthermore, there is no one-size-fits-all solution to the challenge of identifying the shadow of a moving object [49]. In the following, the features that are used to detect shadow will be explained.

2.4.1.1 Intensity features

Intensity features are based on the simple idea that the shadow has a relatively darker colour intensity. In other words, it's based on the intensity of the pixels of the shadow. This method is still effective and fast, however, it sometimes gets confused to distinguish between the shadow and the real object [50].

2.4.1.2 Chromacity features

The motivation behind this method is that the shadow usually lands on a colourful space. So the shadow will have the same colour of that space, but darker. An example of that is supposed the shadow is landed on green grass. The background area is green, and the shadow is darker green. So both the shadow and the background share the same chromacity, but darker [51].

2.4.1.3 Physical features

Physical features include the approaches which try to learn or model the features of the shadow from training datasets (i.e. shadow images). The accuracy of these approaches is usually better than the approaches in which they use chromacity methods. On the other hand, the disadvantage of this method is getting confused when dealing with objects with similar chromonacity to the chromonacity of the background environment. This is because they are relying on spectral characteristics in detection [51][52].

2.4.1.4 Geometry features

In geometry features, the idea is to utilize the geometric shapes of the objects in the foreground, the camera location, and the position of the light source. All of these will lead to the prediction of the shadow shape, size, and orientation. This gives the advantage of working on the pure input frame, without the need for the background estimation. However, the limitation of this method is that there are limited predefined objects shapes, both shadows and foreground objects are requiring different orientations, and this method works only where there is a single light source, or the environment has a flat and stable surface. Besides that, this method can’t deal with more than one shadow to a single foreground object. [51] [52].

2.4.1.5 Textures features

Since the texture of the area under the shadow keeps its texture properties, the motivation of the texture features is that both of the background areas under the shadow and without the shadow are similar. So to detect the shadow, two steps are followed. First step: in this step the pixels which are likely to be shadows are selected. Second step: involve classifying the selected pixels into either shadow reign or background reign based on their texture. If the texture is similar to both the background and the shadow area, it's classified as a shadow area. Many methods are used to detect the texture relationship, including but not limited to, orthogonal transforms, Markov conditional random fields, as well as correlation, which is one of the common methods that give good results, and it can be calculated through the following equation [53].

$$\text{Corr(i,j)}=\sum_{i,j=0}^{n-1} P(i.j)\ \left(\frac{(i-\mu i)(j-\mu j)}{\sqrt{\sigma i^2 \sigma j^2}}\right) \tag{2.14}$$

Because it does not rely on colours and is unaffected by changes in light, the texture features technique is thought to be more powerful than earlier methods. However, texture characteristics have one major drawback: they are computationally costly. Because comparing multiple sections of a picture to see whether they have comparable texture takes a long time [54] [55].

2.4.1.6 Temporal features

Since both the foreground object and its shadow behave the same way. In other words, both the foreground object and its shadow are moving in the same pattern and produce similar temporal consistency, this led to the idea of using the same filters used to detect the foreground object, and detect the shadow. In practice, this method will improve the process of detecting the shadow since it will focus only on the moving objects that may produce shadows. However, there are no approaches that are entirely dependent on the temporal features methods. So the temporal features are a kind of supportive phase in shadow detection [56].

2.5 Evaluation Metrics

Object detection is a binarization technique in which each pixel in the current frame is labelled as either background (black) or foreground (white)

(white). To evaluate the efficiency of such algorithms, the number of correctly identified object pixels (true positives TP). Correctly labelled background pixels (true negatives TN), and pixels that are incorrectly detected as foreground (false positives FP) or wrongly labelled as background are used (false negatives FN) as stated in the table (2.1) these four factors are used to generate a confusion matrix [57] [58].

Table (2.1) Confusion matrix for object detection

		Actual value (Ground truth)	
		Object	Background
Predicted value (by our system)	Object	TP	FP
	Background	FN	TN

The methods that have the entire Receiver Operating Characteristic ROC curve or precision-recall trade would be ideal, but not all approaches allow you to go through the entire series of compromises. In addition, generally, it is not possible rank-order methods depend on a curve. The three metrics will be used as follows:

1. **Recall:** The ratio of the proportion of true positives in the overall number of positive cases labelled in the ground-truth is known as detection rate.

Recall (Re): TP / (TP + FN) (2.15)

2. **Precision:** The fraction of properly detected pixels in the overall foreground identified by any method is known as a positive prediction.

Precision (Pr): TP / (TP + FP) (2.16)

3. **F-measure:** Precision and Recall's harmonic mean, often known as the figure of merit. The higher the score, the more efficient the system is.

F-measure: 2 (Pr*Re) / (Pr+Re) (2.17)

For each video in each category, we first calculated each approach. The recall measure for a certain video v in a specified category c, for example, is computed as follows: TPv,c / (TPv,c + FNv,c) = Rev,c Then, for each category, an average-of-category metric is calculated using the metric's values for all videos it falls into that category, For instance, the overall memory metric for category c is:

$Re_c = 1/ |N_c| \sum Re_{v,c}$ (2.18)

Where $|N_c|$ is the number of videos in category c.

CHAPTER THREE

3.1 Introduction

Several methods associated with the main stages of video surveillance systems (object detection and shadow removing) have been described in the previous chapter. This chapter is specified to explain the design considerations and implementation requirements for each stage of the proposed system, the first stage is foreground detection and the second stage is shadow removal which they consider the crucial stages of video surveillance system due to the positive results of this stages will lead to increase the efficiency of next stages.

3.2 The proposed system description

The overview of the proposed system consists of object detection, shadow removal which is shown in figure (3.1). The proposed system is begun with obtaining the video by a stationary camera which means a camera fixed into a set position and can only be moved manually.Video file consists of a sequence of images (frames) to deal with video firstly the frames are converted to sequence of images (*.bmp) format by using online converter to BMP. The next stage is the image reading process, the general purpose of bitmap BMP format image designed to provide any size of a given image with 24 bit (Red, Green, Blue). The general structure of BMP files has two separate components, file header and actual bitmap (data). The header is the first part of the image file structure, which includes the primary information like the number of bits for each pixel and pointer to the beginning of the pixel data, width, height. The Bitmap (Pixel Data) is the second part, this part is an array of rows of the pixel. Three channels (RGB) represented each pixel of depth 24 bit

The next stage consists of background modelling and detecting moving objects. Background modelling can deal with multi-modal backgrounds based on two types of features (spatial and statistical).

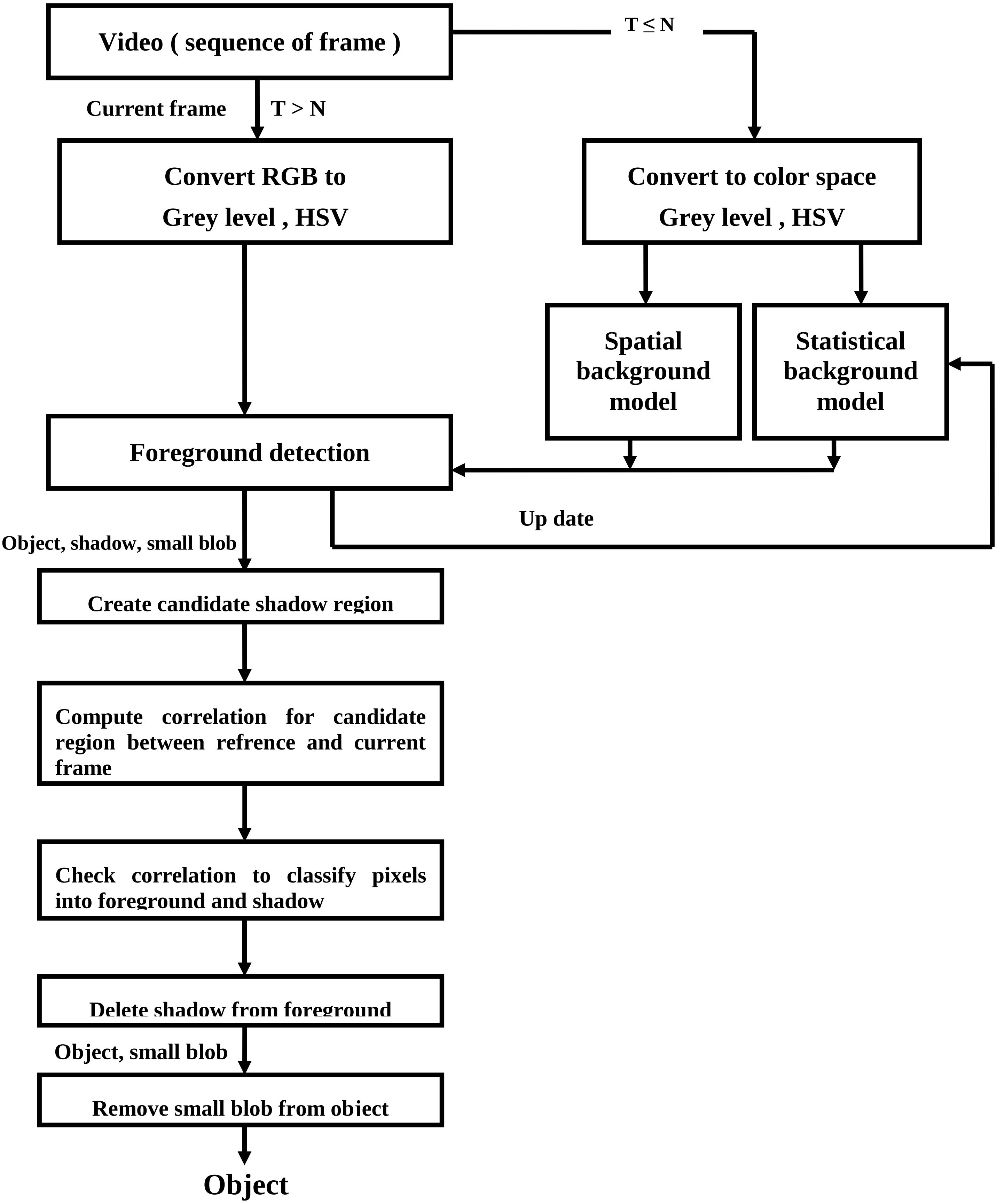


Figure (3.1): Block diagram of the proposed system

In the figure above the (N) represent the number of frames assigned to build models and (T) is the current frame.

The proposed background modelling method can deal efficiently with several challenges such as dynamic background, illumination changes (gradual, sudden), and noise, the main steps of the object detection method can be abstracted by the construction of background models by selecting N clean frames to construct a statistical model and spatial models.

Then moving object detection involves comparing the current image with the background models to label pixels as foreground or background pixels.

Often the detected object consists of the object plus its shadow, in a video surveillance system the shadow represents one of the major challenges, therefore, the shadow removing is done based on two steps, the first one, the candidate shadow pixels is determined using a weak detector, the second one, the candidate shadow pixels are classified into foreground object or shadow using local texture correlation between the current frame and the background models.

To increase the accuracy of the foreground mask, the isolated pixels are deleted also small blobs are removed.

The next sections of this chapter explain the stages of the proposed system in more detail.

3.2.1 Background modeling and foreground detection

This section explain : the background modeling and foreground detection in details.

The algorithm (3.1) illustrate background models construction

First step: Convert every pixel in each of the selected predefined N frames to build the background from RGB colour space to grey level colour space by using equation (2.4).

Second step: construct the statistical model by computing the mean and standard deviation for Pixels for each location using equations (2.1)(2.2).

Third step: construct the spatial model which consists of a set of histogram models built by using the method Center Symmetric Local Binary Patterns (CS-LBP), for each pixel background model are a set of weighted histograms {$\overrightarrow{m1}$, …., $\overrightarrow{mk}$ }, which k user predefined, when k is large this leads to better results in term of accuracy but the time complexity and memory requirements will be raised, the appropriate value of k is 3, which represents a tradeoff between the accuracy on one hand and time complexity and memory requirements on the other hand, figure (3.2) shows the K models for each pixel. In the histogram, the x-axis shows the gray level intensity and the y-axis shows the frequency of that intensity value. In the proposed system the number of models depends on specific condition to remove models with low probability. The spatial model can face illumination changes and noise and is fast in the calculation when compared with another spatial descriptor due to the output value of this descriptor ranging from (0 to 15) while the original descriptor (local binary pattern) the output is range from (0 to 255), the steps below explain the construction of the spatial model.

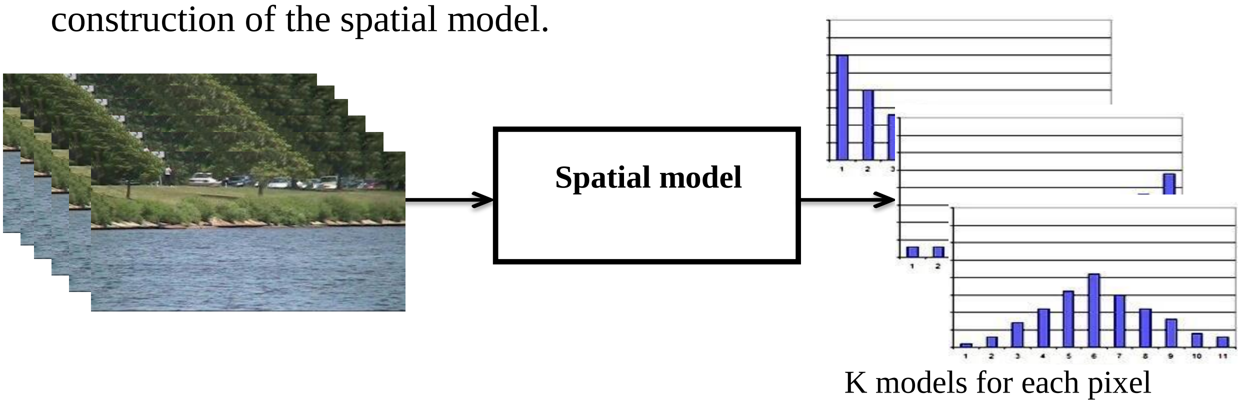


Figure (3.2): The K models for each pixel in background

1- for each pixel, the spatial descriptor (CS-LBP) histogram is computed through the rectangle region around the pixel.

2- the spatial descriptor has weight symbolized with(w_k) and its value between [0...1], the sum of all weights for each (k) model histograms equal to 1.

3- in the beginning, the first five models obtained remain constant, where the comparison starts from the sixth model with the first five models to ensure that in the event of a strong model, we will not lose it, the newly achieved histogram will be set side by side with K model histograms that are associated with the corresponding pixel using histogram intersection as a measure of similarity, table (3.1) shows an example of histogram intersection, histogram intersection calculated using equation (3.1) [59]

$$\cap\,(h,m) = \frac{\sum_{i=0}^{l-1} \min(hi\,,mi)}{\min(\sum_{i=0}^{l-1} hi\,,\sum_{i=0}^{l-1} mi)} \qquad (3.1)$$

Where:

∩: the intersection between two histograms

h: current histogram

m: background model histogram

i: index of histogram bins

L: number of bins of the histogram

table (3.1) an example to explain the histograms intersection

<table>
<tr><th>h (i)</th><th>m (i)</th><th colspan="2">MIN (h(i) , m(i))</th></tr>
<tr><td>26</td><td>18</td><td colspan="2">18</td></tr>
<tr><td>43</td><td>38</td><td colspan="2">38</td></tr>
<tr><td>17</td><td>23</td><td colspan="2">17</td></tr>
<tr><td>5</td><td>9</td><td colspan="2">5</td></tr>
<tr><td colspan="3">Min – sum (h (i), m(i))</td><td>78</td></tr>
<tr><td colspan="2">Min (sum (h), sum (m))</td><td>Min (91 , 88)</td><td>88</td></tr>
<tr><td colspan="4">Intersection = (78/88) = 0.886</td></tr>
</table>

4- if the result of similarity measure does not satisfy the threshold ($T_d = 0.7$) for each one of three histograms models have each pixel then-new histogram model is created instead of histogram model with least weight. small initial weight is assigned for the new histogram model. In the proposed method, the value of the initial weight is 0.005 due to the initial weight must be small value so as not to affect the rest of the models as well as to ensure that this weight is the smallest compared to existing models, then no further processing is needed in term of updating the histogram models and the weights of histogram models.

5- if a new histogram matches one of the histograms models, in this case, processing is necessary. Histogram model with the best matching is adopted with new value by changing its bins using the formula:

$$\overrightarrow{m_k} = \beta \overrightarrow{h} + (1 - \beta) \overrightarrow{m_k} \qquad (3.2)$$

Where $\overrightarrow{m_k}$ background model, β update parameter, $\overrightarrow{h}$ current model, the value of $\beta = 0.01$ where it reflects the amount of update that obtains to the models.

6- The histogram model weights are also updated as follows:

$$w_k = \alpha_w M_k + (1 - \alpha_w)\ w_k \qquad (3.3)$$

Where $\alpha_w \in [0,1]$ is equal to 0.01 where it reflects the amount of updates that obtains to the weights of models. For the best matching histogram, $M_k = 1$ otherwise equal 0. After finishing the updating of the weights of the histograms model. The parameters α_w and β are controlled on the speed of adaptation. The largest values of parameters lead to faster adaptation. In the last step of background modelling, the histogram models are reordered in descending order, and we choose the top three models.

Algorithm (3.1) Construction of Background models

Input: Sequence of frames (image pixel)

Output: statistical model , spatial model

Begin

For frame=1 to N do // N is number of frames.

For i = 0 to frame _ height -1 do

For j = 0 to frame _width -1 do

Begin

Convert RGB_ pixel(i, j) to gray_ pixel(i, j) color space

Construct statistical model by computing mean and standard deviation for pixels of each location using equations (2.1)(2.2)

Construct spatial model over Rregion. By compute (SC-LBP) for pixel // Blok 3*3

K=0 (counter for weighted histograms models)

First five models obtained are incomparable with each other

Calculate similarity S between new spatial histogram with histogram models //using Equation (3.1)

If S >= T_d for one of histogram models then

Begin

Update the best matching histogram model //using Equation (3.2).

Update the weight of histograms model //using Equation (3.3)

End if

Else

Begin

k=k+1

Create new histogram model (with weight is 0.005)

Replace new histogram model with histogram model with least weight.

End Else

End of for j

End of for i

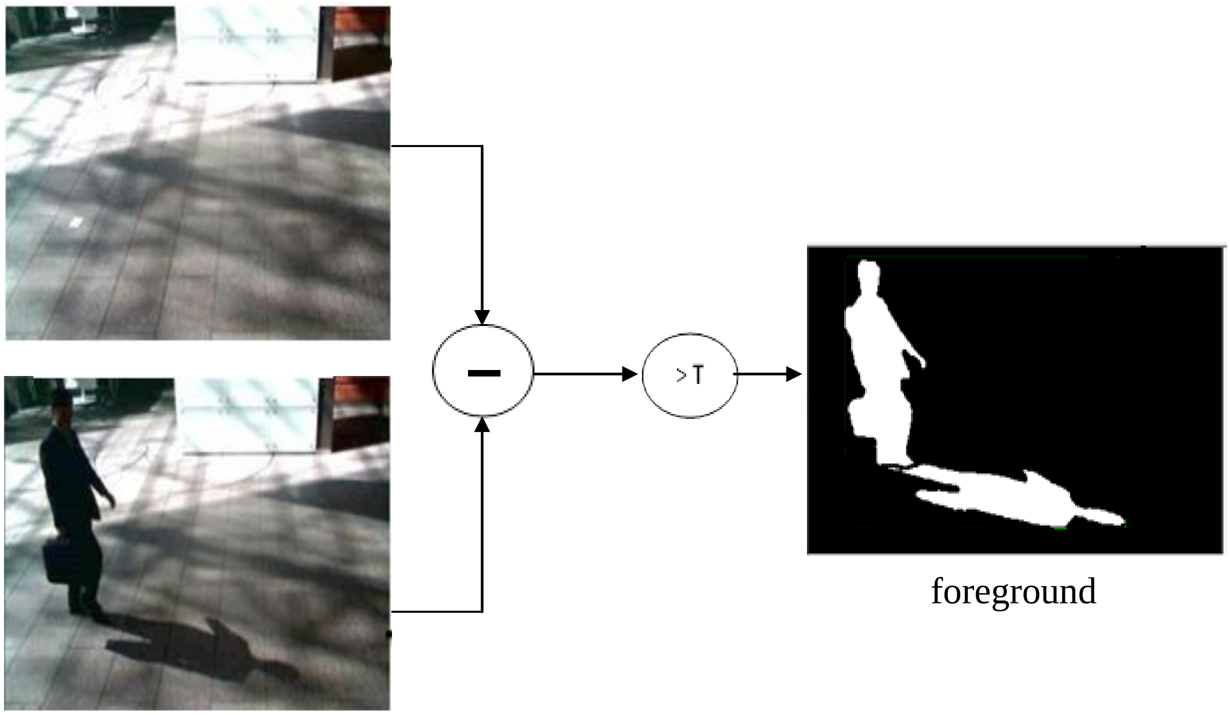
histogram models are reordered in descending order, and choose the top three models

End. //end of algorithm.

3.2.2 Foreground detection stage

In the moving object detection algorithm, the main objective is to evaluate a video sequence to extract moving objects with respect to background models represented by the background scene. Figure (3.3) shows an example of object detection.

Background



foreground

Current frame

Figure (3.3): Foreground detection

First of all, in the present frame, each pixel is checked depending on the statistical model according to the following equations (3.4),(3.7),(3.8) :

$$\text{If } |P_{(x,y,t)} - \mu(_{x,y,t-1})| < 1.5*\sigma \qquad (3.4)$$

where $P_{(x,y,t)}$ and $\mu(_{x,y,t-1})$ represent current and previous pixel value respectively. Then the pixel is considered as background and then update the mean and standard deviation by using estimation method as equation (3.5) (3.6) [60] :

$$\mu_{(x,y,t)} = (1-\varepsilon)\,\mu(_{x,y,t-1}) + \varepsilon\, P_{(x,y,t)} \qquad (3.5)$$

$$\sigma_{(x,y,t)} = (1-\varepsilon)\,\sigma_{(x,y,t-1)} + \frac{\varepsilon}{2}(P_{(x,y,t)} - P_{(x,y,t)})^2 \qquad (3.6)$$

$$\text{If } |P_{(x,y,t)} - \mu(_{x,y,t-1})| > 2.5*\sigma \qquad (3.7)$$

Then the pixel is classified as foreground and then update the mean and standard deviation by using equations (3.5) (3.6).

$$\text{If } 1.5*\sigma < |P_{(x,y,t)} - \mu_{(x,y,t-1)}| \le 2.5*\sigma \qquad (3.8)$$

then the pixel is checked by the spatial model in this case the (CS – LBP) is computed for the current pixel, if the result of similarity measure between the current pixel and the models of the corresponding pixel in the spatial model does not satisfy the threshold ($T_d = 0.7$) for the three histogram models, the pixel will be categorized as foreground. And if the similarity measure satisfies the threshold ($T_d = 0.7$) for one of the three histograms models then the pixel will be categorized as background.

Figure (3.4) illustrate the (foreground detection) process.

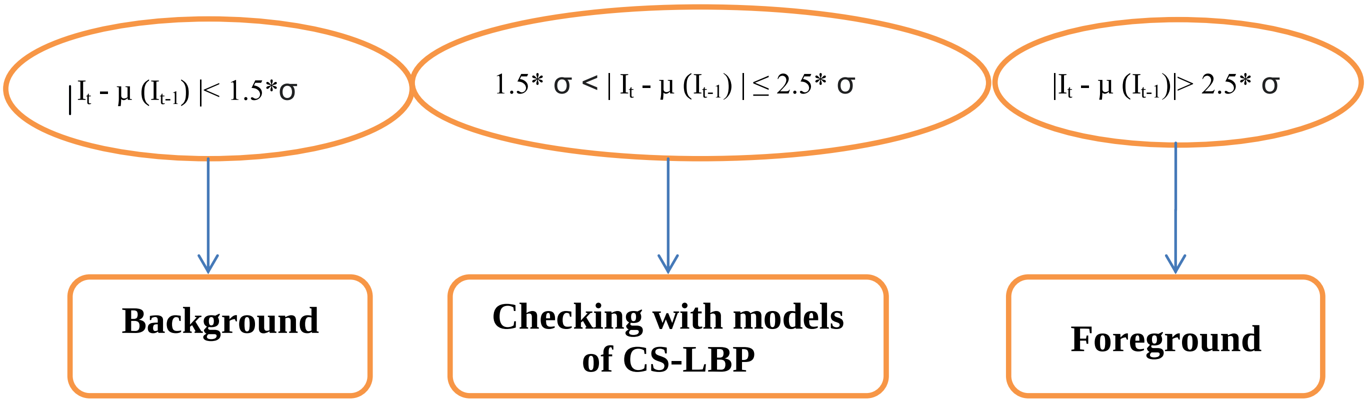


Figure (3.4): The (foreground detection) process

Where I_t and $\mu(I_{t-1})$ represent current and previous pixel respectively.

The algorithm (3.2) illustrates the foreground detection stage.

Algorithm (3.2) Foreground detection

```
Input: frame and background models
Output: binary images(background pixel = 0 and foreground pixel =1)
  Begin
    For frame=1 to N do  // N is number of frames.
      Begin
        For i = 0 to frame_ height -1 do
        For j = 0 to frame_ width -1 do
          Begin
            Convert RGB_pixel(i, j) to Gray level_pixel(i, j) color space
            Compare current pixel with statistical model  // using   equations
              (3.4) (3.7) (3.8)
             If current pixel match equation (3.4) or (3.7) then
                 Begin
               pixel(i,j) = foreground or background
               update mean and standard deviation  // using equations
               (3.5) (3.6)
            End if
           If current pixel match equation (3.8) then
            Begin
              Compare current pixel with spatial model
              Construct spatial histogram over R_region By compute ( SC-LBP)
               for current pixel
               Calculate similarity S between new spatial histogram with
                 histogram models //using Equation (3.1).
              If  S  < T_d for all histograms model then
               Begin
                 pixel(i,j) = foreground.
               End If
             Else
```

```
                  pixel(i,j) = background.
                End Else
              End if
           End for ( i).
           End for ( j).
        End // end of frame sequence
     End. //end of algorithm.
```

3.2.3 Shadow removing stage

There is a truth that the areas under the shadow keep most of their texture

Based on this idea. Texture based method of shadow detection

Follows these two steps :

1- Selecting candidate shadow pixels (which work as simple detector).

2- Classification of the candidate pixels as either foreground or shadow based, on texture correlation.

To select pixels of shadow for a candidate, the weak shadow detector is used based on spectral features then each shadow candidate is classified as either a shadow or an object by computing the correlation of the texture in the frame with the texture in the background models in addition to colour feature. If a candidate's texture plus colour features are correlated in both the frame and the background model, it is classified as a shadow**.**

The next section represents the main steps of the shadow removing stage:

3.2.4.1Create candidate shadow pixels

HSV colour space is used, the images are obtained from the video are RGB colour space, which is converted to HSV. The HSV colour space is computed based on equations (2.5) (2.6) (2.7).

The intensity is measured directly by value (V), pixels that fall under the shadow have a lower value compared with pixels in the background models. Subsequent to the chromacity values, Hue (H) does not change in respect of shadow cast on background, in terms of saturation (S) usually the shadow pixels have a lower saturation. Therefore, the pixel p can be classified as candidate shadow pixel if satisfy three conditions related with chromacity features, when value channel:

$$\boldsymbol{\beta_1} \leq (\mathbf{F}_P^{v} / \boldsymbol{B}_P^{v}) \leq \boldsymbol{\beta_2} \tag{3.9}$$

$$(\mathbf{F}_P^{s} - \boldsymbol{B}_P^{s}) \leq \boldsymbol{\tau_S} \tag{3.10}$$

$$|\mathbf{F}_P^{H} - \boldsymbol{B}_P^{H}| \leq \boldsymbol{\tau_H} \tag{3.11}$$

Where $\boldsymbol{\beta_1}$= 0.3, $\boldsymbol{\beta_2}$ = 0.9 , $\boldsymbol{\tau_S}$ = 40 , $\boldsymbol{\tau_H}$ = 48, these values are selected based on analysis of specific areas with shadow and without shadow.

In the proposed system, the thresholds will be increased to maximize the chance to include all shadow pixels in candidate shadow pixels, the figure (3.5) illustrates an application example of a weak detector of shadow pixels.

111	117	120		139	142	145		95	98	100
125	132	135		138	152	146		97	100	99
114	125	109		140	157	149		83	89	90
Foreground pixels										
120	126	130		99	106	102		149	145	153
121	119	115		110	104	115		141	155	142
117	103	125		107	105	109		147	148	152
Background model pixels										
Hue channel				Saturation channel				Value channel		

The Result		
C	C	F
C	F	C
C	C	C

Figure (3.5): Simple detector example (C: candidate shadow pixel, F: foreground pixels)

3.2.4.2Compute correlation for candidate pixels between background models and current frame

to measure the correlation between background models and candidate pixels in the current frame, initially window with size N*N is selected and the correlation is computed using equation (2.14)

3.2.4.3 Check correlation to classify pixels into foreground or shadow

weighted averaging method of correlation results is calculated (quantized Hue, SC- LBP), The Hue channel is quantized into **u** levels to build Hue histogram using equation (3.12).

New value= new number of levels * (Pixel Value/Old number Levels) (3.12)

the aim of quantization of Hue channel is to reduce the size of histogram that leads to decrease the time complexity. Similar to feature vector of intensity texture, an 5 × 5 pixels structuring element is used to gather the statistic of local color of a pixel. Figure (3.6) shows histogram of quantized hue channel when u=16.

120	200	215	170	185
145	165	222	185	152
136	178	152	125	145
95	74	86	85	63
52	58	76	90	70

A: Original values

7.5	12.5	13.437	10.625	11.562
9.0625	10.312	13.875	11.562	9.5
8.5	11.125	9.5	7.8125	9.0625
5.9375	4.625	5.375	5.3125	3.9375
3.25	3.625	4.75	5.625	4.375

B: New values

8	13	13	11	12
9	10	14	12	10
9	11	10	8	9
6	5	5	5	4
3	4	5	6	4

C: Rounded Values

Figure (3.6): Hue channel quantization, u=16 that is weighting channels according to importance. The texture feature represented by CS-LBP takes 0.65 due to texture feature has great importance in dealing with shadow detection while quantized Hue (H) taking 0.35 due to texture feature has less importance in dealing with shadow detection the equation becomes:

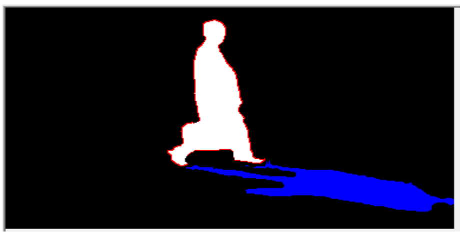
$$corr_t(x,y)=corr_v(x,y)*(0.65) + corr_h(x,y)*(0.35) \quad (3.13)$$

If the value of the $corr_t\ (x,y) > 0.7$ then the pixel is classified as shadow else is classified as foreground object from object detected regions, the value 0.7 is selected by analysis samples of the values of pixels with and without shadow.

3.2.4.4 Delete shadow from foreground

In this step, the shadow pixels are ignored, shadow removal aims to acquire a shadow-free foreground object because remaining the shadow within the foreground object leads to several problems in the next steps such as the difficulty of shape analysis, feature extraction, etc. Figure (3.7) illustrates the shadow detection process. The algorithm (3.٣) illustrates the shadow removing stage

A:The Background frame **B:** The current frame



C: object detection and shadow removing

Figure (3.7): Shadow detection, A: The Background frame, B: The Current frame and C: object detection and shadow removing

Algorithm (3.٣) the shadow removing stage

Input: frame contains objects with shadow.

Output: objects free shadow.

Begin

For m=1 to num_detected_objects

Begin

For i=low to high // height of detected object

For j= left to right // width of detected object

Convert RGB_pixel(i,j) to HSV_pixel(i,j) color space using equations (2.5)(2.6)(2.7).

Quantize Hue_pixel(i,j) channel to 16 levels //using Equation (3.12).

Begin

If Pixel(i,j) =1 then

Begin

If (F(i, j) $_p^v$ / B (i , j)$_p^v$) $>=$ 0.3 and (F(i,j) $_p^v$ / B (i, j)$_p^v$) < 0.9 then

Begin

If(F(i, j) $_p^s$− B (i , j)$_p^s$) $<=$ 40 and

Abs (F(i, j)$_p^H$− B (i, j)$_p^H$) < = 48 then

Begin

Pixel (i, j) = shadow_candidate

For Bk=1 to K // number of spatial background models.

Select background model with highest weight

Begin

```
If Pixel (i, j) = shadow_candidate then
  Begin
    Compute texture_correlation (F_P^{CS-LBP}, B_P^{CS-LBP}) using equation (2.14)
    Compute color_correlation ( F_P^H , B_P^H ) using equation (2.14)
    Weighted_correlation using equation (3.13)
   If Weighted_correlation >= 0.7 then
        Begin
          Pixel(i,j) = shadow
        End if
          Else
           Begin
             Pixel(i,j) = foreground
          End else
    End if
End // for Bk
End if
End if
End if
End  // for j
End  // for i
End  // for m
End. // end of algorithm.
```

3.2.4.5 Remove small blob from shadow free object

After completion of treatment of object detection and shadow removal from foreground object stages, a number of small blobs or noise may appear, This happens for two reasons, the first is the false detection of the object, for example, an area that is actually a background and is classified as a foreground, and the second reason is that when the shadow is detected, there is a misclassification for a part of shadow pixels, and when the shadow is removed, these pixels are remain as a small blob.

The noise will be treated as a small blob because the condition is any size of a detected object less than 1% of the total number of pixels will consider as a small blob and will be ignored.

The algorithm (3.4) illustrate the small blob removing .

Algorithm (3.4) of small blob removing

```
Input: binary image (foreground = 1, background = 0).
Output: objects without small blob.
 Begin
   counter=0
   For I = 0 to height - 1
   For j = 0 to width - 1
     Begin
       If pixel(i,j) =1 then
         Begin
           Loc_x(x)=k
           Loc_y(y)=l
           counter=counter + 1
           For k = i-1 to i +1
           For m = j-1 to j +1  // check 8 neighbors of pixel(i,j).
```

```
        Begin
        If pixel(k,l) =1 then
         Begin
           Loc_x(x)=k
           Loc_y(y)=l
           counter=counter + 1
           Add_to_buffer(pixel(k,l))
         End if
      End  // for m
      End  // for k
    While buffer() not empty do
      Begin
        Pixel(x,y)= Remove_from_buffer()
        For k= x - 1 to x + 1
        For l= y - 1 to y + 1
          Begin
            If pixel(k,l)=1 then
              Begin
                Loc_x(x)=k
                Loc_y(y)=l
                 counter=counter + 1
                Add_to_buffer(pixel(k,l))
              End if
            End  // for l
            End  // for k
        End // while
  If counter  <  (h*w*0.01) then
    Begin
```

```
            Replace the value of (loc(x),loc(y)) with zero
          End if
      End // for j
      End // for i
   End. //end of algorithm.
```

CHAPTER Four

4.1 Introduction

To evaluate the performance of the proposed system and compare it with benchmark algorithms, ChangeDetection.net Video Database is used which represents a standard and most common available dataset [57] [58], Figure (4.1) explains samples of the dataset as well as the corresponding website www.changedetection.net. This dataset contains some video categories from four to six video sequences for each category. The dataset thereby includes a wide variety of challenging, real-world scenarios. All videos are existing with manually constructed ground truth that identifies change relative to a training portion of the video. The images of ground truth are made with 5 labels: moving, static, shadow, unknown and non-ROI. Shadow and static labels are backgrounds, moving labels are foreground and other labels are not involved in the statistics computation.

4.2 CDnet 2014 Dataset description

The dataset of videos has multi videos which provide a factual, camera-captured (without CGI - Computer Generated Imagery), various videos set. These videos have been selected to cover a wide range of detection challenges and are a delegate of exemplary indoor and outdoor visual data that are captured in monitoring, video database scenarios, and smart environment [57] [58]

4.3 Ground Truth Labels

All frames have been manually marked at the pixel level, as follows [57] [58]:

1. Value of 0 represents the static pixels (background).

2. Value of 50 represents shadow pixels. The shadow label is related to clear and well-defined moving shadows such as the shadow area as explained in Figure (4.1).

3. Non-ROI (Region of Interest) pixels (i.e. outside of the ROI) are allocated

a greyscale value of 85. In each video sequence, the first hundred frames are characterised as Non-ROI to avoid the inaccuracy of evaluation metrics because of the errors that happen during the initialization process. Also, the Non-ROI label avoids the metrics from being corrupted by activities unconnected to the considered category.

4. Pixels that are half occluded or corrupted by motion blur are allocated with the unknown grayscale value of 170

5. Value of 255 represents moving pixels (foreground).

Figure (4.1): Two video frame samples with associated manually-labelled ground truth.

4.4 Experimental results of object detection and shadow removing

CVPR 2014 Change Detection and evaluation metrics (Recall, F_ Measure, Precision) are used to evaluate the proposed system. The next tables and figures show the comparison result of:

(1) Object detection stage (2) Shadow removal stage

The proposed system is compared with benchmark methods. In our use case, the evaluation scores were obtained from the Change detection website (www.changedetecition.net). The top result of each metric is highlighted. The proposed method will be compared with the following benchmark methods: Kernel Density Estimation (KDE) [61], Gaussian Mixture Model (GMM) [62], SWCD [63].

The tables from (4.1) to (4.8) include the following columns.

1- TP: true positives is the number of pixels that are correctly labeled as foreground pixels.

2- FP: false positives the number of pixels that are mistakely labelled as foreground.

3- FN: false negative represents the number of background pixels that are incorrectly classified as background.

4- TN: true negatives the number of pixels that are correctly classified as background.

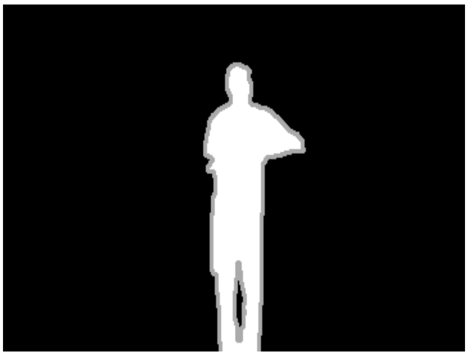
5- Recall: is the ratio of the number of true positives to the total number of positive cases labelled in the ground truth. It is computed by equation (2.15).

6- Precision: is the fraction of the number of true positives to the total number of foreground pixels detected. It is computed by equation (2.16).

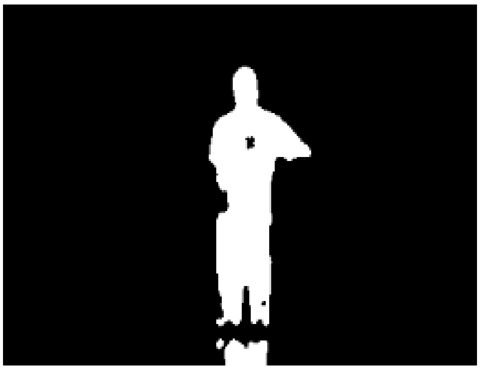
7- F-measure: is the consonant mean of Precision and Recall. It is computed by equation (2.17).

The columns values from (1 to 4) for the benchmark methods used to compare with the proposed system are obtained from the used database and for the proposed system is computed by the program.

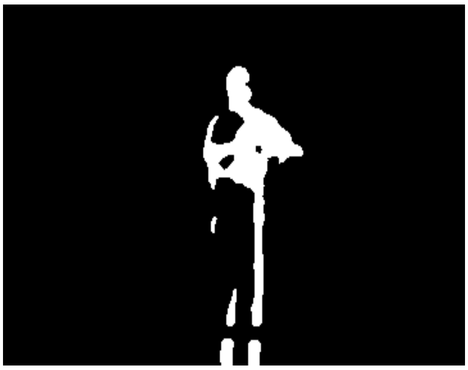
A: input frame



B: ground truth

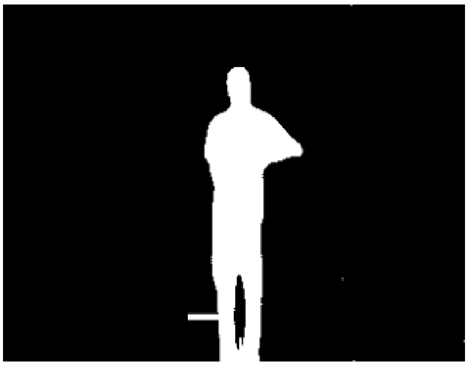


C: KDE



D: GMM

E: SWCD



F: Proposed method

Figure (4.2): Object detection of office video

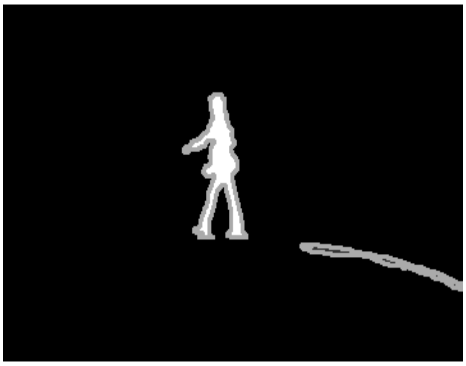
Figure (4.2) contains the frame 1100 that represents the ROI of the office video from the database, as well as the ground truth for the ROI, as well as the images of the results of the proposed method for the object detection stage and the benchmark methods that are compared with them.

Table (4.1) Result analysis office video, Num. of pixels=125180780

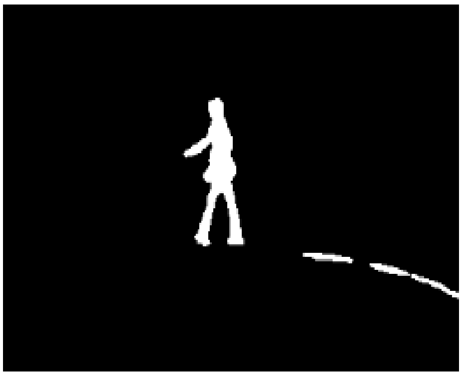
Method	TP	FP	FN	TN	Recall	F_ Measure	Precision
KDE	7825053	262239	817041	116276447	0.9054579	0.93548598	0.967574
GMM	4238342	1440907	4403752	115097779	0.4904299	0.5918917	0.746286
SWCD	8419127	869508	222967	115669178	0.9741998	0.93907247	0.90639
Proposed method	8727388	289233	224560	115939599	**0.9749149**	**0.9714060**	**0.9679222**

Table (4.1) shows the values of (TP, FP, FN, TN) and the evaluation metrics (Recall, F_Mesure, Precision) for the office video. The experimental result in this table shows that the proposed system gives the highest results in terms of (Recall, F_Mesure, Precision).

A: input frame

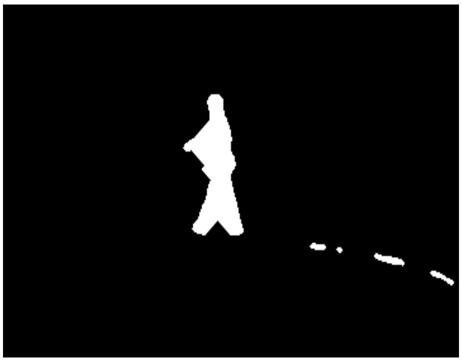


B: ground truth

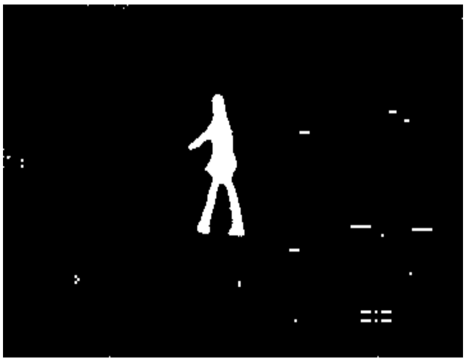


C: KDE

D: GMM



E: SWCD



F: Proposed method

Figure (4.3): Object detection of Pedestrians video

Figure (4.3) contains frame 950 that represents the ROI of the Pedestrians video from the database, and the ground truth for the ROI, as well as the images of the results of the proposed method for the object detection stage and the benchmark methods that are compared with them.

Table(4.2) Result analysis of Pedestrians video, Num. of pixels=68239969

Method	TP	FP	FN	TN	Recall	F-Measure	Precision
KDE	640022	26329	30840	67542778	0.9540293	0.95724765	0.960488
GMM	661976	55587	8886	67513520	**0.9867544**	0.9535639	0.922534
SWCD	646369	70629	24493	67498478	0.9634903	0.9314613	0.901493
Proposed Method	672185	26570	9352	67531862	0.9862781	**0.9739750**	**0.9619752**

Table (4.2) shows the values of (TP, FP, FN, TN) and the evaluation metrics (Recall, F_Mesure, Precision) for the Pedestrians video. The experimental result in this table shows that the proposed system gives the highest results in terms of (F_Mesure, Precision) and the GMM gives the highest result in terms of (Recall).

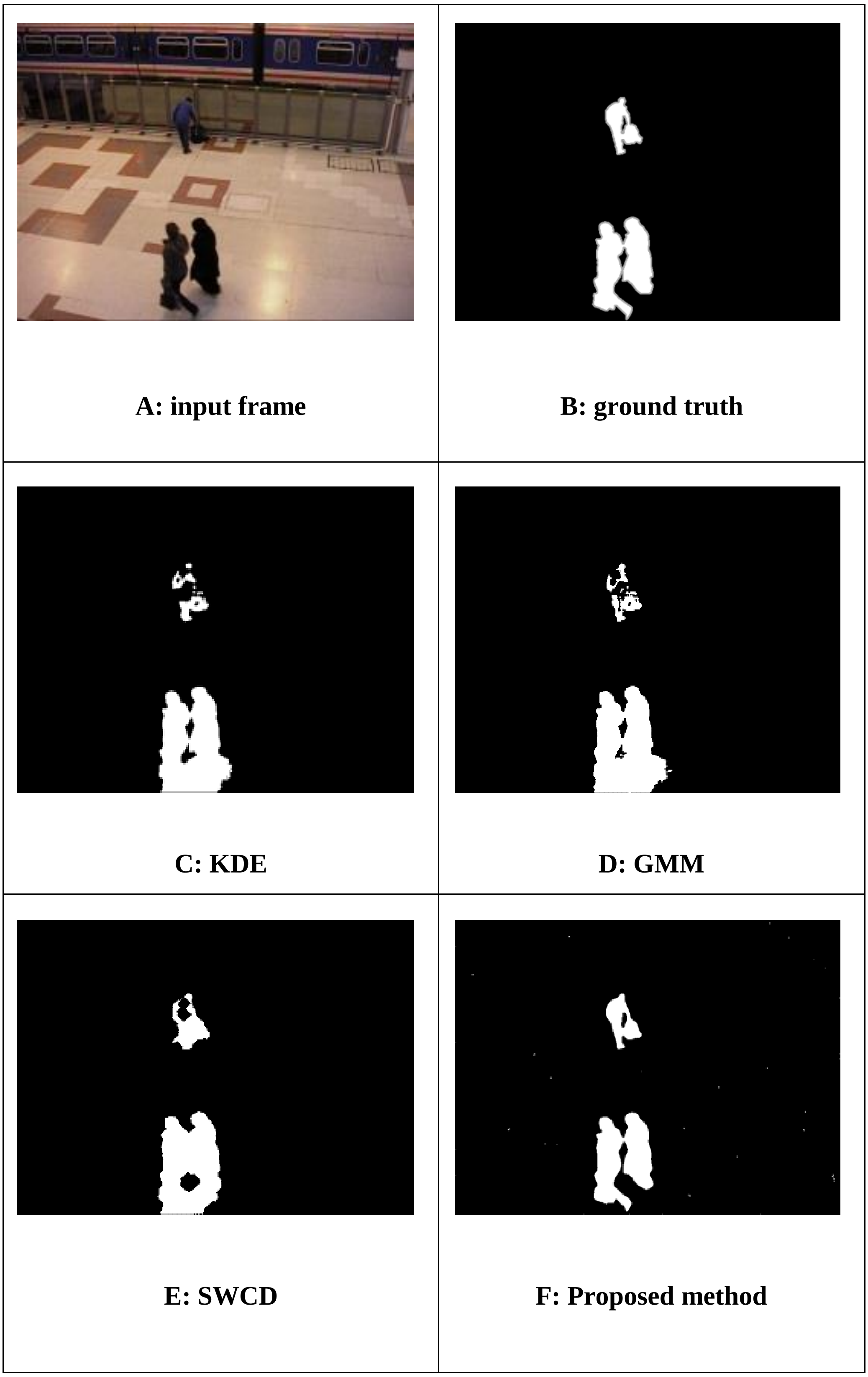


Figure (4.4): Object detection of PETS2006 video

Figure (4.4) contains frame 950 that represents the ROI of the PETS2006 video from the database, and the ground truth for the ROI, as well as the images of the results of the proposed method for the object detection stage and the benchmark methods that are compared with them.

Table (4.3) Result analysis of PETS2006 video, Num. of pixels=371541494

Method	TP	FP	FN	TN	Recall	F-Measure	Precision
KDE	3816139	790623	1012051	365922681	0.790387	0.8089366	0.828378
GMM	4232121	1154934	596069	365558370	0.876544	0.8285892	0.785609
SWCD	4624026	746572	204164	365966732	**0.957714**	**0.9067795**	0.860989
Proposed Method	4666663	748810	230288	365895733	0.952973	0.9050565	**0.861728**

Table (4.3) shows the values of (TP, FP, FN, TN) and the evaluation metrics (Recall, F_Mesure, Precision) for the PETS2006 video. The experimental result in this table shows that the proposed system gives the highest results in terms of (Precision) and the SWCD gives the highest result in terms of (Recall, F_Mesure).

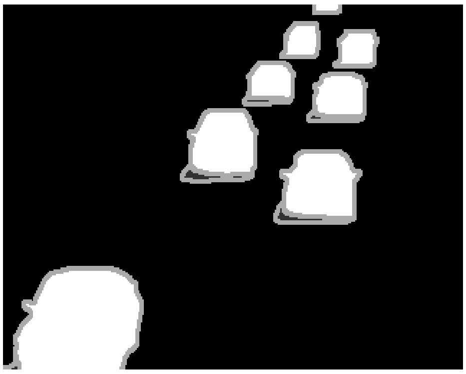
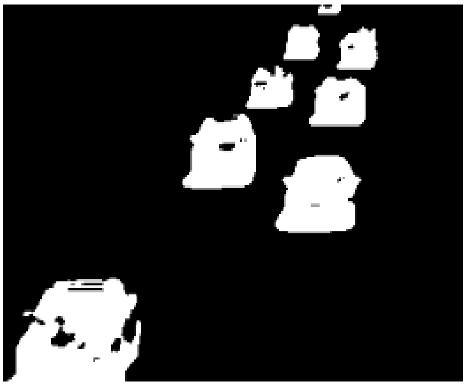
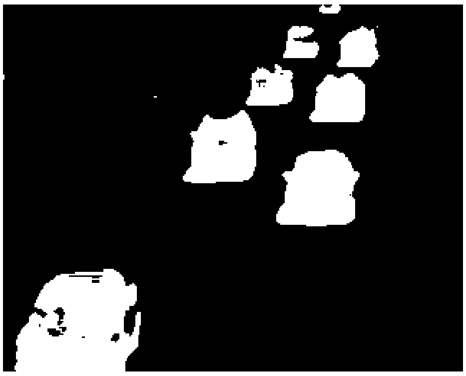
 **A: input frame**	 **B: ground truth**
 **C: KDE**	 **D: GMM**
 **E: SWCD**	 **F: Proposed method**

Figure (4.5): Object detection of Highway video

Figure (4.5) contains the frame 800 that represents the ROI of the Highway video from the database, and the images of the results of the proposed method for object detection stage as well as the ground truth for the ROI, and the benchmark methods that are compared with them.

Table (4.4)Result analysis of Highway video, Num. of pixels =92112458

Method	TP	FP	FN	TN	Recall	F-Measure	Precision
KDE	5118955	368979	339034	86285490	0.937883	0.9353172	**0.932765**
GMM	5011645	378316	446344	86276153	0.9182219	0.9239801	0.929811
SWCD	5176973	762799	281016	85891670	0.9485129	0.908419	0.871578
Proposed Method	5118312	377240	237100	86379806	**0.955727**	**0.943383**	0.931355

Table (4.4) shows the values of (TP, FP, FN, TN) and the evaluation metrics (Recall, F_Mesure, Precision) for the highway video. The experimental result in this table shows that the proposed system gives the highest results in terms of (Recall, F_Mesure) and the KDE gives the highest result in terms of (Precision).

A: input frame

B: ground truth

C: KDE

D: GMM

E: SWCD

F: Proposed method

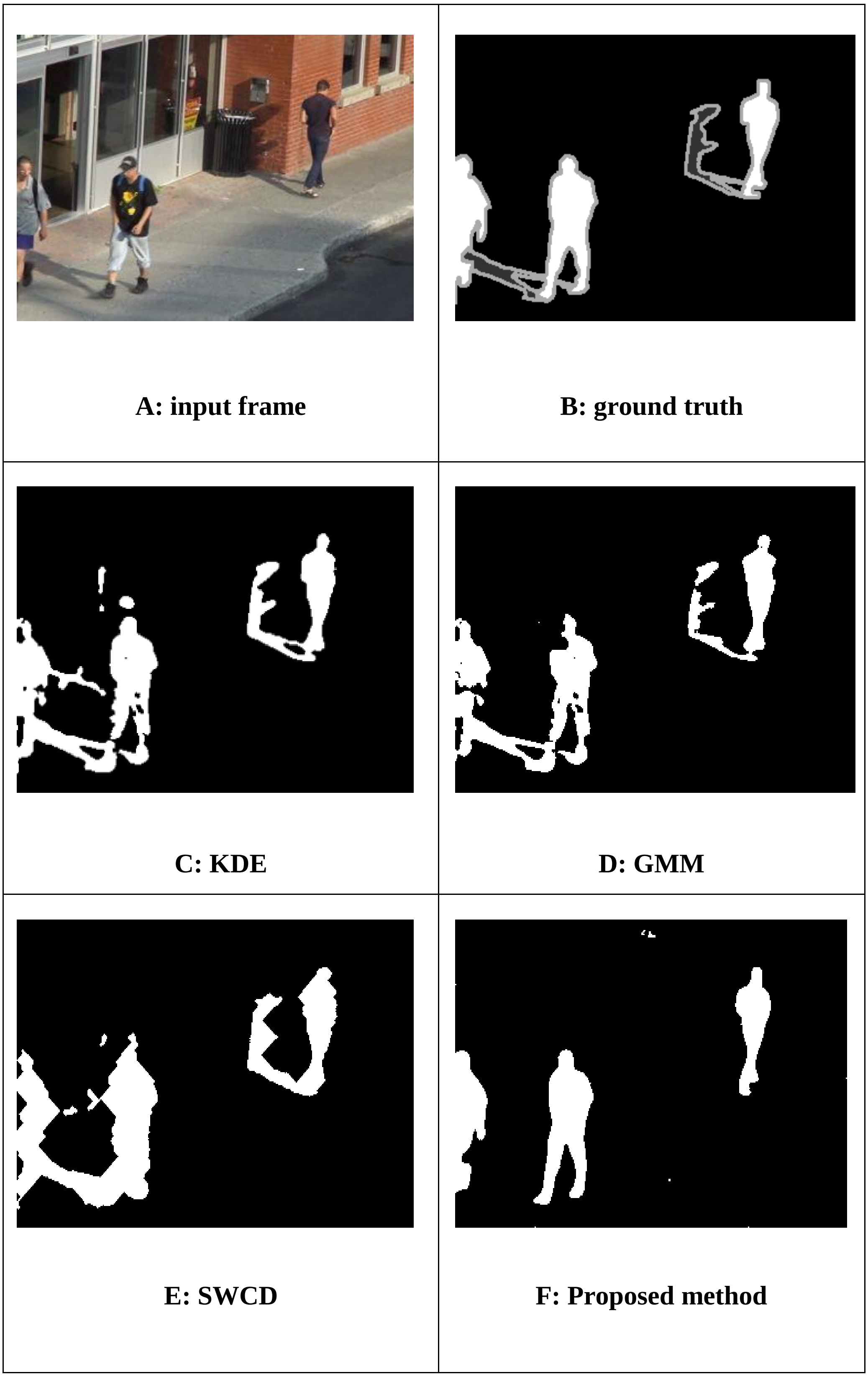
Figure (4.6): Object detection of fountain02 video

Figure (4.6) contains frame 1279 that represents the ROI of the fountain02 video from the database, and the images of the results of the proposed method for object detection stage as well as the ground truth for the ROI, and the benchmark methods that are compared with them.

Table (4.5) Result analysis of fountain02 video, Num. of pixels= 93360558

Method	TP	FP	FN	TN	Recall	F-Measure	Precision
KDE	227895	58572	39334	123903759	0.852807891	0.823177339	0.795537
GMM	232951	79682	34278	123882649	0.871727994	0.803470481	0.745126
SWCD	249240	19188	17989	123943143	**0.932683204**	0.930595512	0.928517
Proposed Method	266692	17805	20778	123924285	0.927721154	**0.932543311**	**0.937416**

Table (4.5) shows the values of (TP, FP, FN, TN) and the evaluation metrics (Recall, F_Mesure, Precision) for the fountain02 video. The experimental result in this table shows that the proposed system gives the highest results in terms of (F_Mesure, Precision) and the SWCD gives the highest result in terms of (Recall).



Figure(4.7): Shadow removal of busStation video

Figure (4.7) contains frame 1057 that represents the ROI of the busStation video from the database, and the images of the results of the proposed method for the shadow removing stage as well as the ground truth for the ROI, and the benchmark methods that are compared with them.

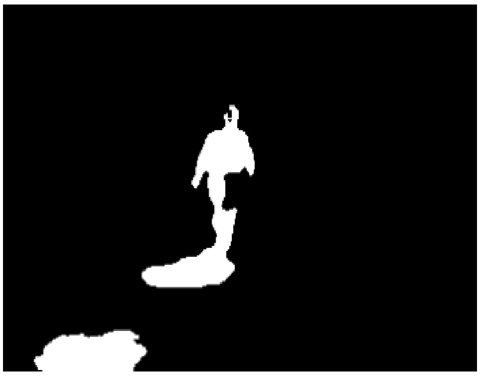
Table (4.6) Result analysis of busStation video, Num. of pixels= 79966711

Method	TP	FP	FN	TN	Recall	F-Measure	Precision
KDE	2194698	417374	756511	76598128	0.743660649	0.788994121	0.840213
GMM	2166043	287639	785166	76727863	0.733951069	0.801512186	0.882773
SWCD	2693481	656616	257728	76358886	0.912670367	0.854896112	0.804001
Proposed Method	2882296	301797	199965	76582653	**0.935123924**	**0.919927601**	**0.905217279**

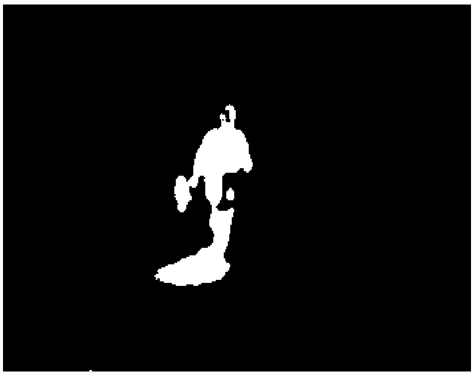
Table (4.6) shows the values of (TP,FP,FN,TN) and the evaluation metrics (Recall ,F_Mesure ,Precision) for the busStation video. The experimental result in this table shows the proposed method gives the highest results in terms of (Recall, F_Mesure, Precision).

A: input frame

B: ground truth



C: KDE



D: GMM

E: SWCD

F: Proposed method

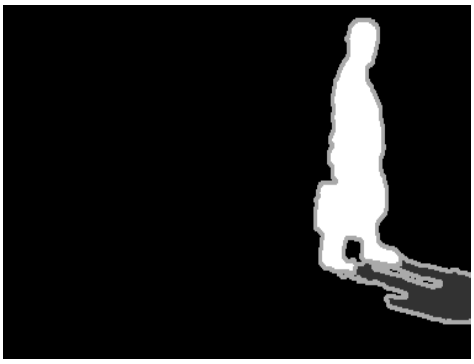
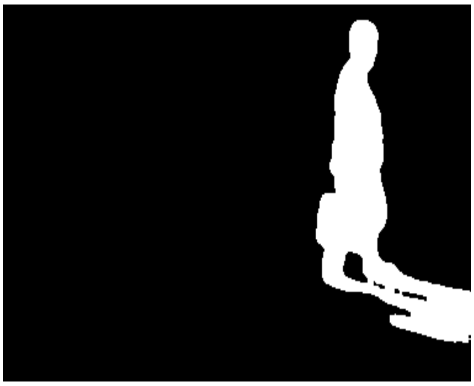
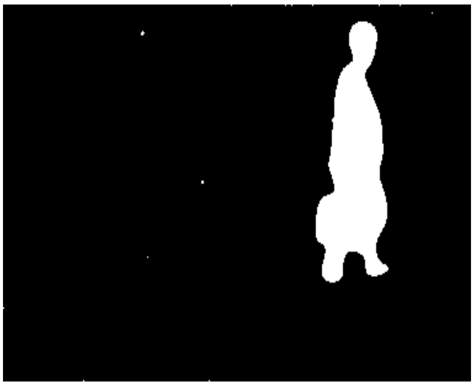
Figure(4.8): Shadow removal of cubicle video

Figure (4.8) contains frame 1233 that represents the ROI of the cubicle video from the database, and the images of the results of the proposed method for the shadow removing stage as well as the ground truth for the ROI, and the benchmark methods that are compared with them.

Table (4.7) Result analysis of cubicle video, Num. of pixels = 528277968

Method	TP	FP	FN	TN	Recall	F-Measure	Precision
KDE	8660750	6997267	1714443	510905508	0.834755556	0.665361667	0.553119
GMM	8420397	6940898	1954796	510961877	0.811589433	0.654354782	0.548157
SWCD	9250238	1503754	1124955	516399021	0.891572619	0.875588718	0.860168
Proposed Method	9477374	1066787	673943	517059864	**0.933610289**	**0.915888389**	**0.898826753**

Table (4.7) shows the values of (TP, FP, FN, TN) and the evaluation metrics (Recall, F_Mesure, Precision) for the cubicle video. The experimental result in this table shows the proposed method gives the highest results in terms of (Recall, F_Mesure, Precision).

 **A: input frame**	 **B: ground truth**
 **C: KDE**	 **D: GMM**
 **E: SWCD**	 **F: Proposed method**

Figure(4.9): Shadow removal of peopleInShade video

Figure (4.9) contains frame 489 that represents the ROI of the peopleInShade video from the database, and the images of the results of the proposed method for the shadow removing stage as well as the ground truth for the ROI, and the benchmark methods that are compared with them.

Table (4.8) Result analysis of peopleInShade video, Num. of pixels = 86906016

Method	TP	FP	FN	TN	Recall	F-Measure	Precision
KDE	4709517	882798	195839	81117862	0.960076496	0.897249876	0.842141000
GMM	4630042	878690	275314	81121970	0.943874818	0.889188184	0.840491000
SWCD	4800534	1142382	104822	80858278	**0.978631113**	0.885032012	0.807774000
Proposed Method	4789574	562567	112471	81441404	0.977056310	**0.934169519**	**0.894889354**

Table (4.8) shows the values of (TP,FP,FN,TN) and the evaluation metrics (Recall ,F_Mesure ,Precision) for the peopleInShade video. The experimental result in this table shows the proposed method gives the highest results in terms of (F_Mesure, Precision) and the SWCD gives the highest result in terms of (Recall).

Table (4.9) contains the total precision rate for sequence videos used in the object detection stage for proposed methods and benchmark methods that compare with them. The best result of each precision metric is highlighted.

Table (4.9) Comparison of average Precision Rate for object detection stage

Sequence name	KDE	GMM	SWCD	Proposed method
Office	0.967574	0.746286	0.90639	**0.967922**
Pedestrians	0.960488	0.922534	0.901493	**0.9619752**
PETS2006	0.828378	0.785609	0.860989	**0.861728**
Highway	**0.932765**	0.929811	0.871578	0.931355
fountain02	0.795537	0.745126	0.928517	**0.937416**
Total	0.8969484	0.8258732	0.8937934	**0.9320792**

Table (4.9) shows the proposed method gives the highest results in terms of precision in most of the rating scale used in the videos sequences and in total the proposed method show the best result among them.

Table (4.10) contains the total precision rate for sequence videos used in the shadow removing stage for proposed methods and benchmark methods that compare with them. The best result of each precision metric is highlighted.

Table 10. Comparison of average Precision Rate for shadow removal stage

Sequence name	KDE	GMM	SWCD	Proposed method
peopleInShade	0.842141	0.840491	0.807774	**0.89488935**
busStation	0.840213	0.882773	0.804001	**0.90521727**
cubicle	0.553119	0.548157	0.860168	**0.89882675**
Total	0.74516	0.75714	0.82398	**0.89964**

Table (4.10) shows the proposed method gives the highest results for the precision of the rating scale used in the videos sequences and in total the proposed method shows the best result among them.

Figure (4.10) shows the Precision,Recall,F-measure Relationship of office video of the object detection method.

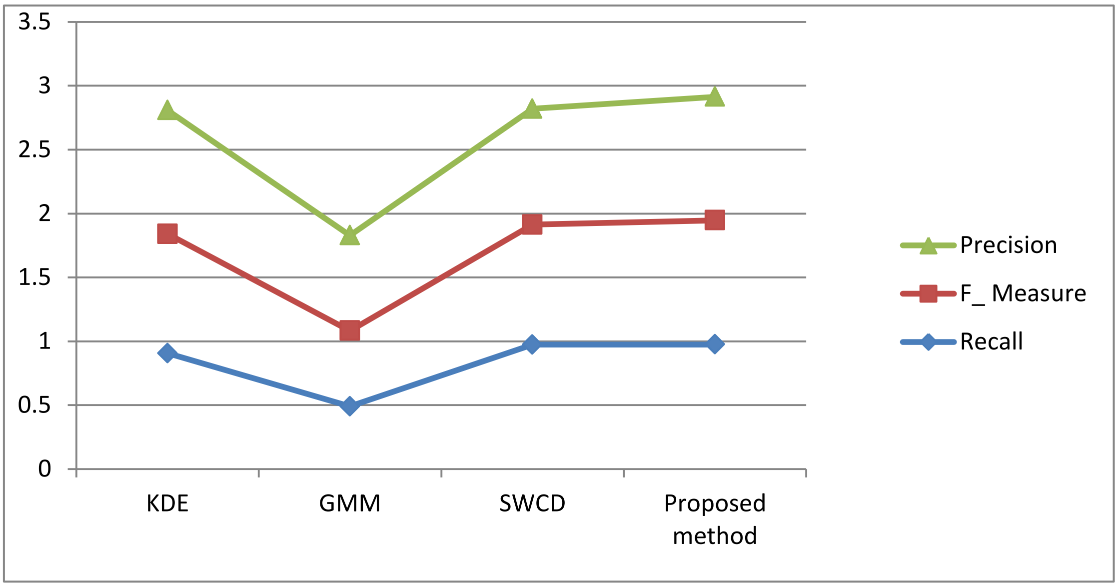


Figure (4.10): Precision,Recall,F-measure Relationship of Baseline category,office video

Figure (4.11) shows the Precision,Recall,F-measure Relationship of Pedestrians video of the object detection method.

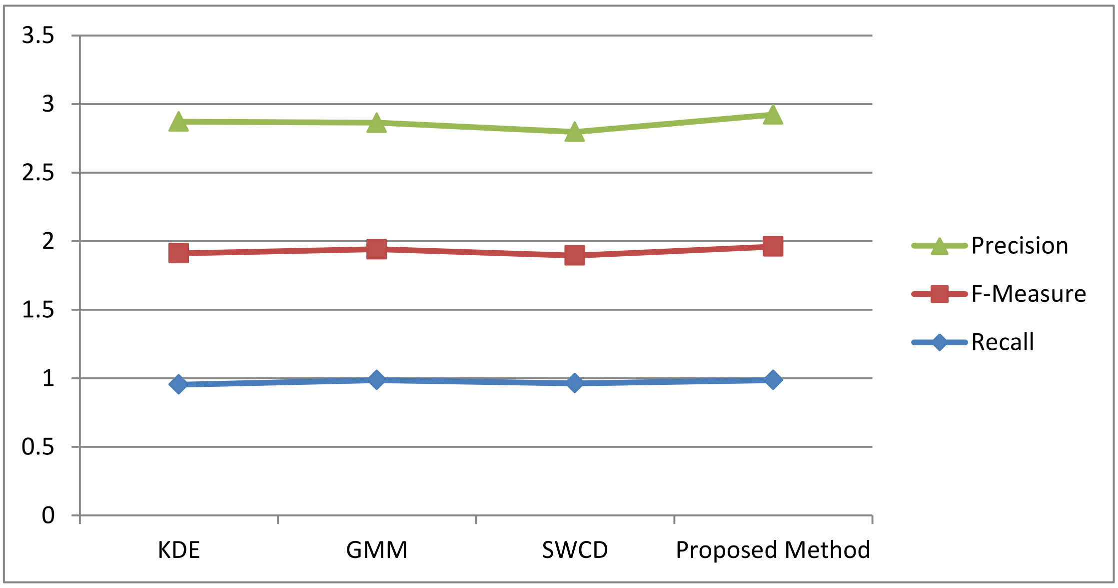


Figure (4.11)Precision,Recall,F-measure Relationship of Baseline category,Pedestrians video

Figure (4.12) shows the Precision,Recall,F-measure Relationship of PETS2006 video of the object detection method.

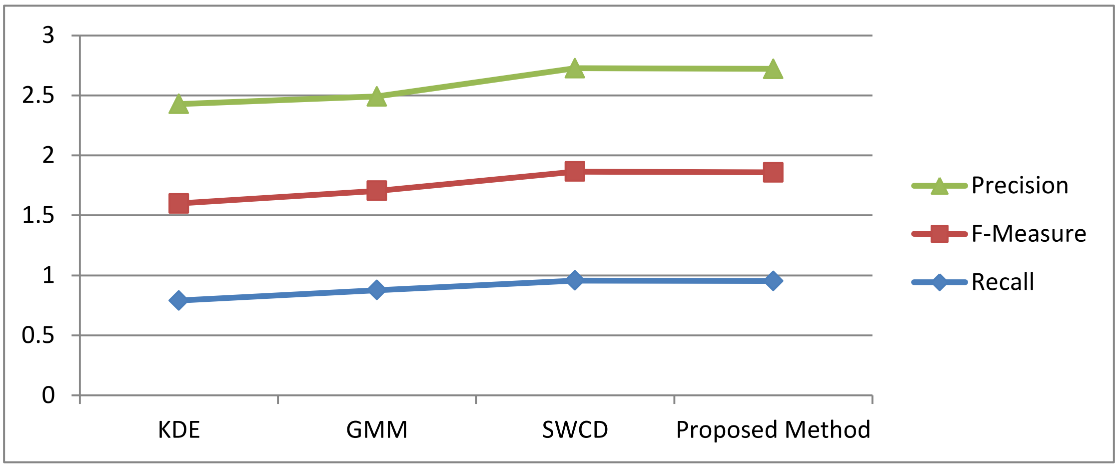


Figure (4.12)Precision,Recall,F-measure Relationship of Baseline category, PETS2006 video

Figure (4.13) shows the Precision,Recall,F-measure Relationship of Highway video of the object detection method.

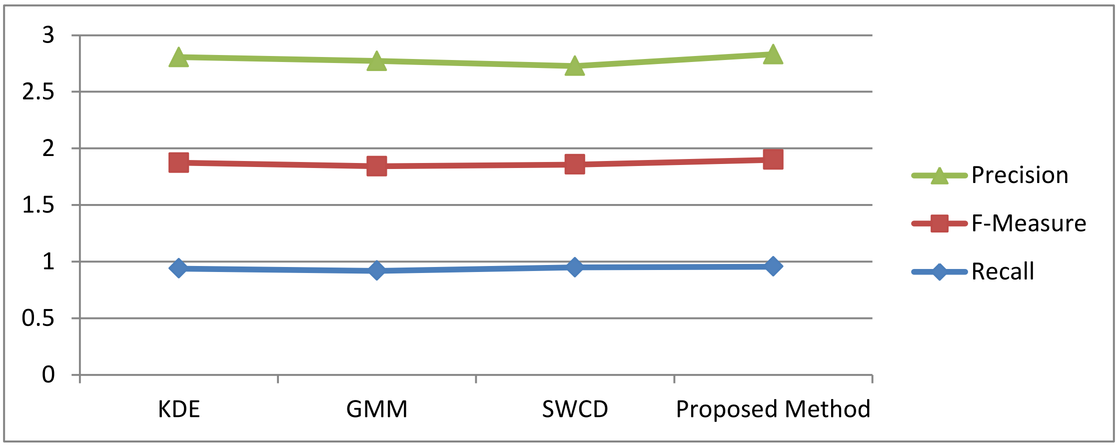


Figure (4.13)Precision,Recall,F-measure Relationship of Baseline category, Highway video

Figure (4.14) shows the Precision,Recall,F-measure Relationship of fountain02 video of the object detection method.

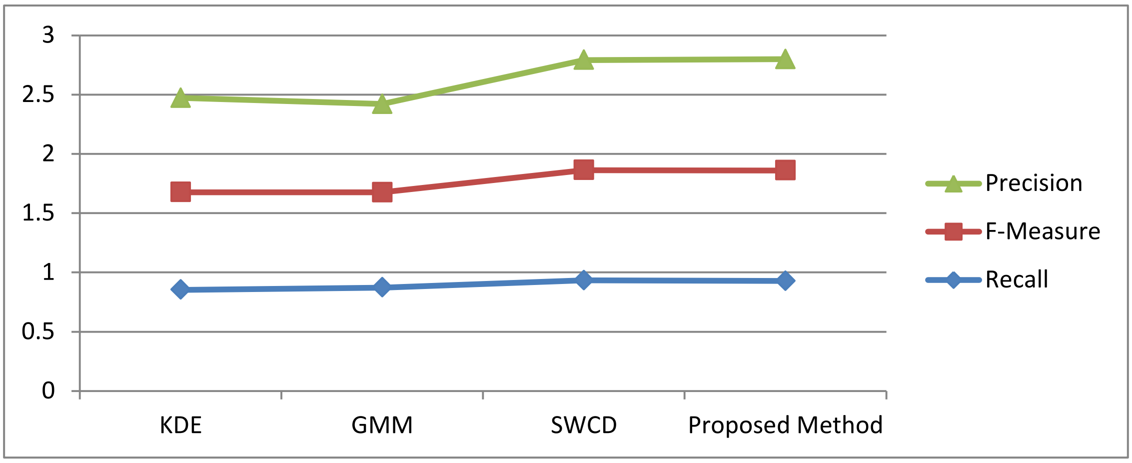


Figure (4.14)Precision,Recall,F-measure Relationship of fountain02 video

Figure (4.15) shows the Precision,Recall,F-measure Relationship of Shadow removal of busStation video.

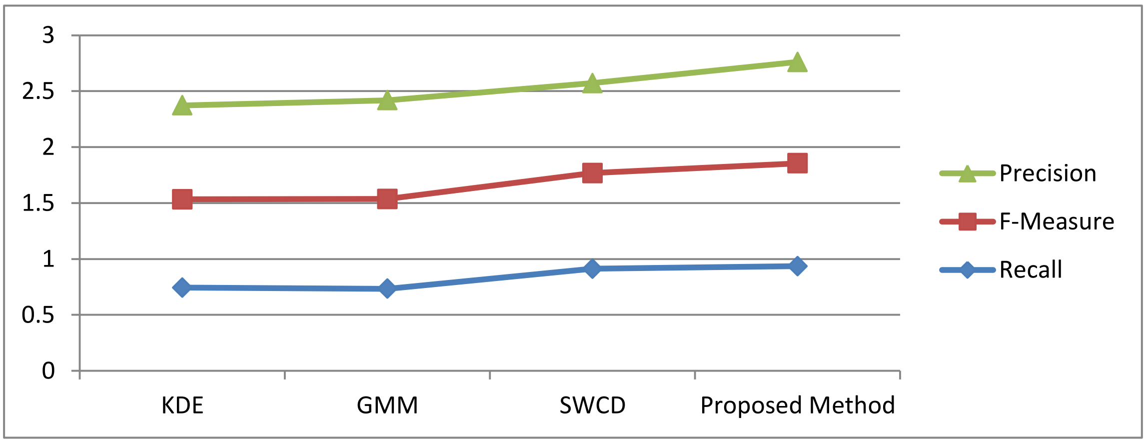


Figure (4.15)Precision,Recall,F-measure Relationship of busStation video

Figure (4.16) shows the Precision,Recall,F-measure Relationship of Shadow removal of cubicle video .

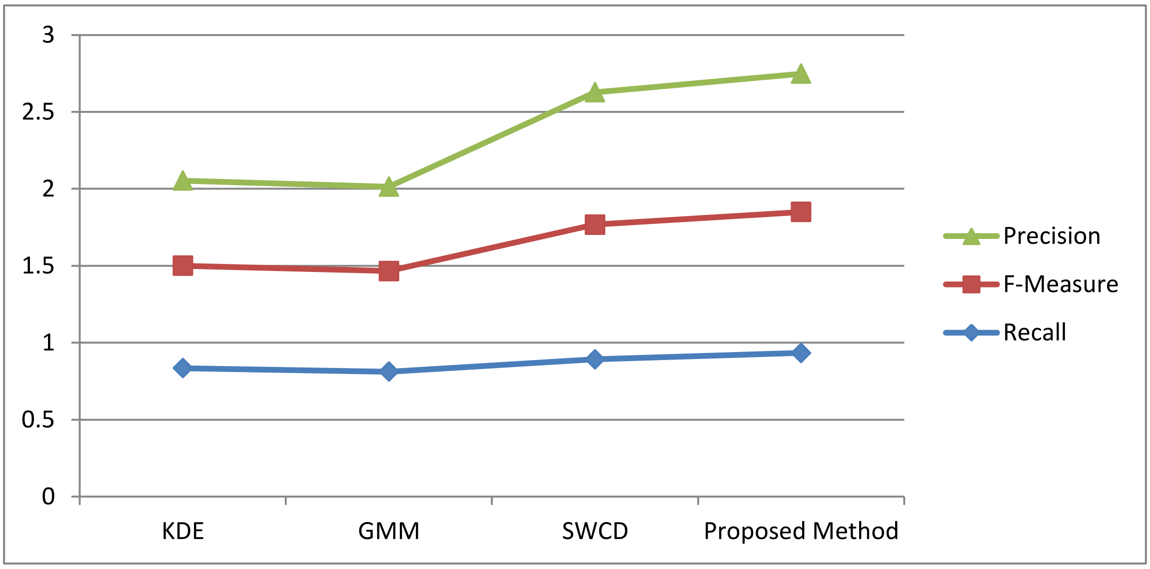


Figure (4.16): Precision,Recall,F-measure Relationship of cubicle video

Figure (4.17) shows the Precision,Recall,F-measure Relationship of Shadow removal of peopleInShade video .

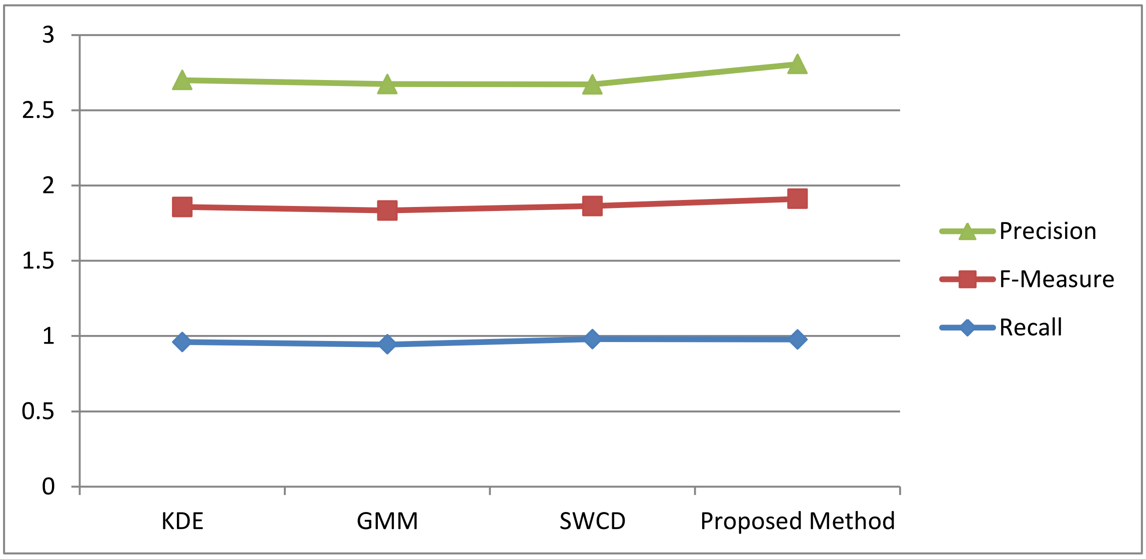


Figure (4.17): Precision,Recall,F-measure Relationship of peopleInShade video

Figure (4.18) shows the total Precision Relationship of sequence of videos that used in object detection methods.

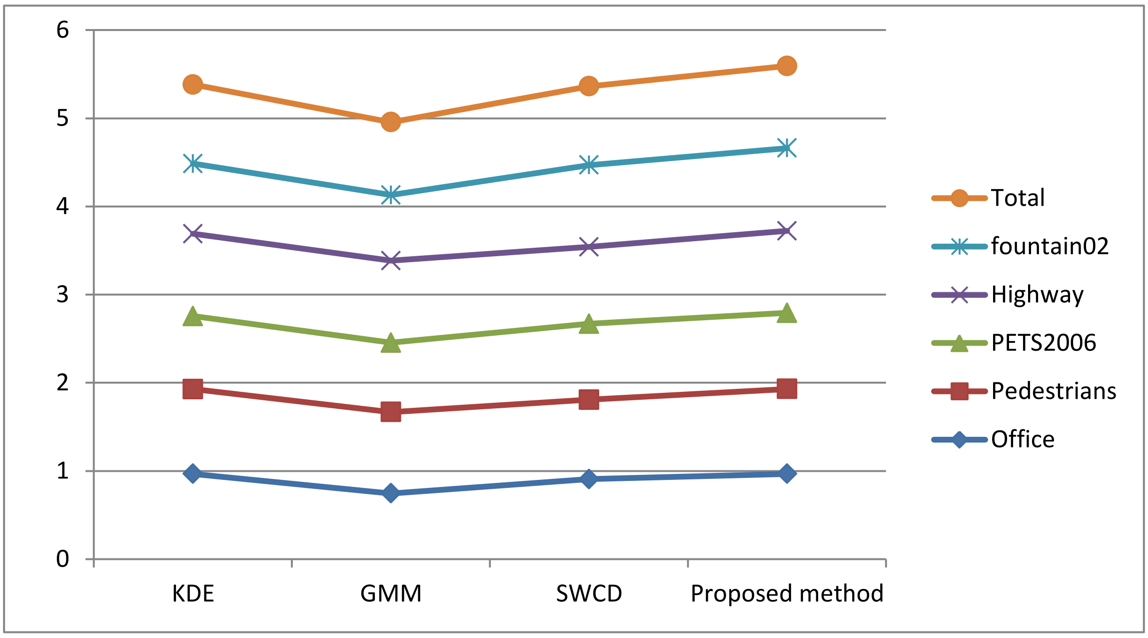


Figure (4.18): Precision Relationship of total object detection methods

Figure (4.19) shows the total Precision Relationship of sequence of videos that used in shadow removing methods.

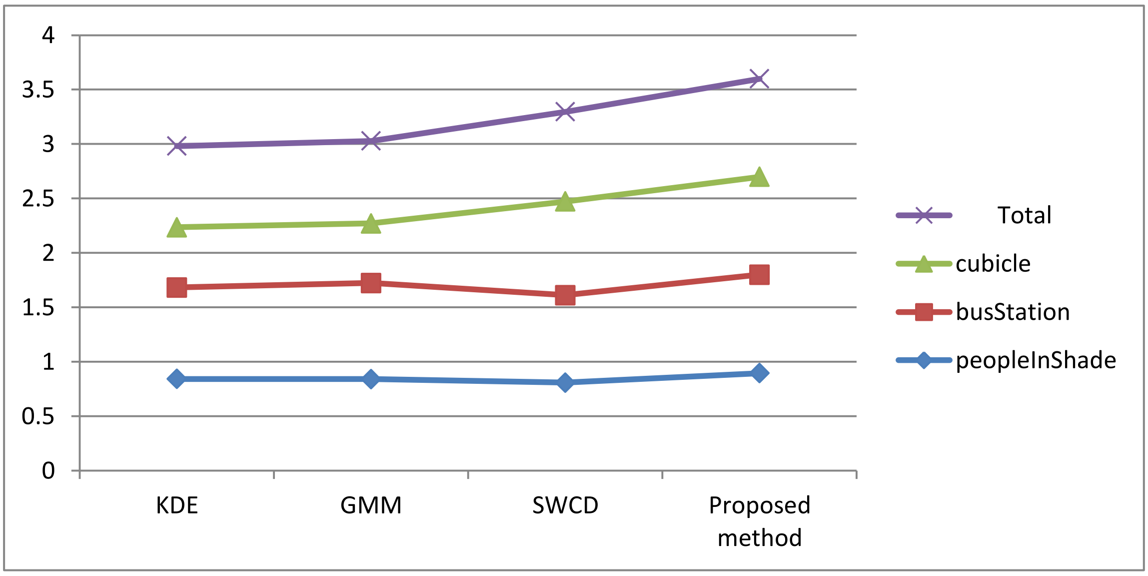


Figure (4.19): Precision Relationship of total shadow removing methods

4.5 Time complexity

The software of proposed system applied under Windows 7 operating system and Visual basic 6.0 language using computer with processor core i7, RAM 8 G.B, VGA card 1 G.B, table (4.11) shows a comparison of the avarge rate for the number of frames per second that was processed in the proposed system with benchmark methods.

Table (4.11) : comparison of the avarge rate for the number of frames

No.	Method	Average Time Complexity	Average Precision
1	**GMM**	**11.6 fps**	**0.7915088**
2	**KDE**	**7.4 fps**	**0.8210542**
3	**SWCD**	**6.7 fps**	**0.8588867**
4	**Proposed method**	**10.2 fps**	**0.9158596**

CHAPTER FIVE

5.1 Conclusions

From the design and implementation of the proposed system and discussing its results, the following points are concluded:

1. The proposed system addresses the problems of the statistical model in terms of accuracy when the pixel is located within the critical region it is difficult for the statistical model to classify it (foreground or background) this pixel is compared with the spatial model that is accurate in classification.

2. The proposed system addresses the problems of the spatial model which is characterized by the complexity of time as most of the pixels are compared with the statistical model which is fast in terms of time.

3. The statistical model has good features in terms of the ability to face dynamic background throw updating the mean and standard deviation to update the statistical background model.

4. The spatial model can face illumination changes because the use of texture-dependent features leads to effective algorithms towards illumination change because the relationships between adjacent pixels remain the same even when the illumination changes.

5. Texture features are suitable to address a range of problems such as illumination changes in local level (linear and monotonic), global level, noise, dynamic background and shadow.

6. The proposed shadow detection and removal technique is a hybrid technique of chromacity that represents a weak detector to select candidate shadow regions and an enhanced form of the correlation method.

7. The use of a weak detector in the shadow removing algorithm leads to reducing the time complexity for the algorithm due to avoiding checking all pixels of detected objects and increasing the efficiency of the proposed method.

5.2 Open Problems and Future Works

It is also suitable to suggest the following future works:

1. Objects classification is an important research area the detected moving area may be different objects such as vehicles, swaying trees, humans, floating clouds, birds and other moving objects.

2. Face detection and recognition is another important research area in the field of video surveillance systems. Detection and recognition methods of human faces in the video surveillance system are difficult because of several problems such as face blurriness, face darkness due to illumination direction, human not facing the camera, self-shadow, different facial expressions, small human face size because the long view of the camera and faces that be similar to each other very closely.

3. Human behaviour analysis is another important researchable to recognize and predict human behaviour, as well as describe it in natural language. It can be used in the fields of intelligent surveillance, protection of accidents, marketing, psychology, etc.

References

References
[1] Vezzani, R.: Computer Vision for People Video Surveillance Ph.D. Thesis. (2006) Available via Internet. http://imagelab.ing.unimo.it/Pubblicazioni/ publications.
[2] M. Narayana, "Automatic Tracking of Moving Objects in Video for Surveillance Applications," MSc thesis, University of Kansas, 2007.
[3] S., Manjula & Tamilselvan, Lakshmi & Ravichandran, Manjula. (2016). A STUDY ON OBJECT DETECTION.
[4] Shaikh, Soharab & Saeed, Khalid & Chaki, Nabendu. (2014). Moving Object Detection Using Background Subtraction. 10.1007/978-3-319-07386-6.
[5] Shaikh, Soharab & Saeed, Khalid & Chaki, Nabendu. (2014). Moving Object Detection Using Background Subtraction. 10.1007/978-3-319-07386-6.
[6] Prati, A., Mikic, I., Trivedi, M. M., & Cucchiara, R. (2003). Detecting moving shadows: algorithms and evaluation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(7), 918–923. doi:10.1109/tpami.2003.1206520.
[7] Kar, A., Deb, K., 2015. Moving cast shadow detection and removal from Video based on HSV color space, in: 2nd International Conference on Electrical Engineering and Information and Communication Technology, ICEEiCT 2015. Institute of Electrical and Electronics Engineers Inc. doi:10.1109/ICEEICT.2015.7307443.
[8] Sanin, A., Sanderson, C., Lovell, B.C., 2012. Shadow detection: A survey and comparative evaluation of recent methods. Pattern Recognition. doi:10.1016/j.patcog.2011.10.001.
[9] Y.-T. Zhang, J.-Y. Bae, and W.-Y. Kim, "Multi-Layer Multi-Feature Background Subtraction Using Codebook Model Framework," vol. p.63, no. 1, pp. 23–27, 2016.
[10] A. M. Rahma and N.B. Abd," Detection of a Moving Car Based on Invariant Moments", Journal of Computer Sciences, 2017.
[11] S. Shams, "Foreground Object Detection and Separation Based on Region Contrast," Iraqi Journal of Science 2017, Volume 58, Issue 4A, Pages 1963-1969.
[12] M. Ghazi, "Foreground Object Detection Based on Chrominance and Texture Features with Enhancement by Canny Filter," Iraqi Journal of Information Technology. V.9 N.2. 2018 , ISSN: 2664-0600.

[13] C. Y. Lin, K. Muchtar, W. Y. Lin, and Z. Y. Jian, “Moving Object Detection through Image Bit-Planes Representation without Thresholding,” IEEE Trans. Intell. Transp. Syst., vol. 21, no. 4, pp. 1404–1414, 2020, doi: 10.1109/TITS.2019.2909915.
[14] Mohanty Subrata Kumar; Rup, Suvendu (2020). An adaptive background modeling for foreground detection using spatio-temporal features. Multimedia Tools and Applications, (), –. doi:10.1007/s11042-020-09552-8.
[15] S. H. Khan, M. Bennamoun, F. Sohel, and R. Togneri, “Automatic Shadow Detection and Removal from a Single Image,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 3, pp. 431–446, Mar. 2016, doi: 10.1109/TPAMI.2015.2462355.
[16] Hu, Y., B. Warig, Y. Warig and Y. Zhao, 2016. Moving shadows removal using HSV color space and texture analysis. Adv. Sci. Technol. Lett., 122: 240-246.
[17] Carmen, D.D.J.R. and R.D. Cajote, 2016. Moving shadow detection and removal for video-based traffic monitoring. Proceedings of the Annual Summit and Conference on Asia-Pacific Signal and Information Processing Association (APSIPA), December 13-16, 2016, IEEE, Jeju, South Korea, ISBN: 978-1-5090-2401-8, pp: I -7.
[18] Abdusalomov, A.; Whangbo, T.K. Detection and Removal of Moving Object Shadows Using Geometry and Color Information for Indoor Video Streams. Appl. Sci. 2019, 9, 5165.
[19] H. D. Najeeb and R. F. Ghani, “Shadow Elimination in Soccer Game Video using Background Subtraction and Sobel Operators”, eijs, vol. 61, no. 9, pp. 2426-2433, Sep. 2020.
[20] Minghu Wu, Rui Chen, Ying Tong, "Shadow Elimination Algorithm Using Color and Texture Features", Computational Intelligence and Neuroscience, vol. 2020, Article ID 2075781, 10 pages, 2020. doi.org/10.1155/2020/2075781 .
[21] Kokaram, A.C.,2010. n.d. Image and Video Processing Course Notes for Integrated Systems Design 1–22.
[22] Dimitrijević, Z., Wu, G., Chang, E.Y., 2003. SFINX: A multi-sensor fusion and mining system, in: ICICS-PCM 2003 - Proceedings of the 2003 Joint Conference of the 4th International Conference on Information, Communications and Signal Processing and 4th Pacific-Rim Conference on Multimedia. Institute of Electrical and Electronics Engineers Inc., pp. 1128–1132. doi:10.1109/ICICS.2003.1292636.
[23] Gawande, U., Hajari, K., & Golhar, Y. (2020). Pedestrian Detection and Tracking in Video Surveillance System: Issues, Comprehensive Review, and Challenges.Recent Trends in Computational Intelligence. doi:10.5772/intechopen.90810 .

[24] Paul, M., Haque, S.M.E., Chakraborty, S., 2013. Human detection in surveillance videos and its applications - a review. EURASIP Journal on Advances in Signal Processing 2013, 176. doi:10.1186/1687-6180-2013-176.
[25] Araki, S., Matsuoka, T., Yokoya, N., Takemura, H., 2000. Real-time tracking of multiple moving object contours in a moving camera image sequence. IEICE TRANSACTIONS on Information and Systems 83, 1583–1591.
[26] Luboš OVSENÍK, Anna KAŽIMÍROVÁ KOLESÁROVÁ, Ján TURÁN, " VIDEO SURVEILLANCE SYSTEMS", Act Electrotechnicaet Informatica, Vol. 10, No. 4, 2010, pp. 46–53, ISSN 1335- 8243.
[27] Mohanty, A., Sanjivani, S., 2015. A Survey on Moving Object Detection using Background Subtraction Methods in Video. International Journal of Computer Applications 5–10.
[28] Fedias, M., Saigaa, D., 2011. A new approach based in mean and standard deviation for authentication system of face. International Review on Computers and Software (I.RE.CO.S.), Vol. 5, n. 3.
[29] Li, Shuai & Florencio, Dinei & Zhao, Yaqin & Cook, Chris & Li, Wanqing. (2017). Foreground Detection in Camouflaged Scenes. 10.1109/ICIP.2017.8297083.
[30] Bala, R., Braun ,K. M., Color to grayscale conversion to maintain discriminability, Xerox Innovation Group, 800 Phillips Rd, 128-27E, Webster, NY 14580.
[31] Bouwmans, T., Silva, C., Marghes, C., Zitouni, M.S., Bhaskar, H., Frelicot, C., 2018. On the role and the importance of features for background modeling and foreground detection. Computer Science Review. doi:10.1016/j.cosrev.2018.01.004.
[32] M. Loesdau, S. Chabrier, A. Gabillon, "Hue and Saturation in the RGB Color Space", Proc. of the 6th International Conference ICISP 2014 LNCS, vol. 8509, pp. 203-212, 2014.
[33] Heikkilä, M., Pietikäinen, M., 2006. A texture-based method for modeling the background and detecting moving objects. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 657–662. doi:10.1109/TPAMI.2006.68.
[34] Tan Wenming, Li Bin, Z.W., 2010. Research on background modeling method based on center-symmetric local binary patterns, Journal of University of Science and Technology of China.
[35] Zhang, Z., Wang, C., Xiao, B., Liu, S., Zhou, W., 2012. Multi-scale fusion of texture and color for background modeling, in: Proceedings - 2012 IEEE 9th International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2012. pp. 154–159. doi:10.1109/AVSS.2012.48.

[36] Bouwmans, T., Silva, C., Marghes, C., Zitouni, M.S., Bhaskar, H., Frelicot, C., 2016. On the role and the importance of features for background modeling and foreground detection. arXiv:1611.09099v1.
[37] Rout, R.K., 2013. A Survey on Object Detection and Tracking Algorithms 75.
[38] Xu, Y., Dong, J., Zhang, B., Xu, D., 2016. Background modeling methods in video analysis: A review and comparative evaluation. CAAI Transactions on Intelligence Technology 1, 43–60. doi:10.1016/j.trit.2016.03.005.
[39] Bouwmans, T., Maddalena, L., Petrosino, A., 2017. Scene background initialization: A taxonomy. Pattern Recognition Letters 96, 3–11. doi:10.1016/j.patrec.2016.12.024.
[40] Toyama, K., Krumm, J., Brumitt, B., Meyers, B., 1999. Wallflower: principles and practice of background maintenance, in: Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, pp. 255–261 vol.1. doi:10.1109/ICCV.1999.791228.
[41] Bouwmans, T. (2014-07-25). "Traditional Approaches in Background Modeling for Static Cameras". Background Modeling and Foreground Detection for Video Surveillance. CRC Press. ISBN 9781482205374.
[42] Thangarajah, Akilan & Wu, Q. M. Jonathan & Singh, Ashutosh & Madon, Bakri & Chowdhury, Adib. (2015). Video Foreground Detection in Non-static Background Using Multi-dimensional Color Space. Procedia Computer Science. 70. 55-61. 10.1016/j.procs.2015.10.030.
[41] I. Setitra, S. Larabi, 2014. Background subtraction algorithms with post-processing: A review, Proc. Int. Conf. Pattern Recognit., pp. 2436-2441.
[42] A. Elgammal and L. Davis. Non-parametric model for background subtraction. European Conference on Computer Vision, ECCV 2000, pages 751–767, June 2000.
[43] Bouwmans, T., El Baf, F., Vachon, B., 2008. Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey. Recent Patents on Computer Sciencee 1, 219–237. doi:10.2174/2213275910801030219.
[44] Zivkovic, Z., & Van Der Heijden, F. (2006). Efficient adaptive density estimation per image pixel for the task of background subtraction. Pattern Recognition Letters, 27(7), 773–780. doi.org/10.1016/j.patrec.2005.11.005 .
[45] Bo Li , Yuhong Li , Han Zhou, 2014, An Improved Kernel Density Estimation Approach for Moving Objects Detection, The Open Automation and Control Systems Journal, 6, 768-781.

[46] Shaikh, S.H., Saeed, K., Chaki, N., 2014. Moving Object Detection Using Background Subtraction. Springer International Publishing, Cham. doi:10.1007/978-3-319-07386-6.
[47] Self-Supervised Shadow Removal .2020 (Florin-Alexandru VasluianuETH Zurich , Andr´es RomeroETH Zurich , Luc Van GoolETH Zurich.
[48] Kar, A., Deb, K., 2015. Moving cast shadow detection and removal from Video based on HSV color space, in: 2nd International Conference on Electrical Engineering and Information and Communication Technology, ICEEiCT 2015. Institute of Electrical and Electronics Engineers Inc. doi:10.1109/ICEEICT.2015.7307443.
[49] Subramanyam, M., Nallaperumal, K., Ravi, S., Perumalsamy, P., Durairaj, S., S, S.K., S, G.D., 2014. A Review on Survey and Analysis of Shadow Detection Techniques. Proc. of Int. Conf. on Adcances in Communication.
[50] Katarina Berg, Jonna Hellstr¨O M, 2013, Shadow Detection For Improved Object Tracking In Surveillance Cameras, Master Thesis.
[51] Jyoti Bala, Ritika, 2014, Techniques and Algorithms of Shadow Detection in Images , International Journal Of Core Engineering & Management (IJCEM), Volume 1, Issue 7.
[52] Mosin Russell, 2014, Moving Shadow Detection Based on Colour-Intensity Analysis, master thesis.
[53] Ramola, Ayushman; Shakya, Amit Kumar; Van Pham, Dai (2020). Study of statistical methods for texture analysis and their modern evolutions. Engineering Reports, 2(4), –. doi:10.1002/eng2.12149
[54] Leone, A., Distante, C., 2007. Shadow detection for moving objects based on texture analysis. Pattern Recognition 40, 1222–1233. doi:10.1016/j.patcog.2006.09.017.
[55] Sanin, A., Sanderson, C., Lovell, B.C., 2010. Improved shadow removal for robust person tracking in surveillance scenarios, in: Proceedings - International Conference on Pattern Recognition. pp. 141–144. doi:10.1109/ICPR.2010.43.
[56] Zhang, Z., 2005. Background noise distribution before and after high-resolution processing in ship-borne radar. Chinese Journal of Electronics 14, 115–118. doi:10.1049/ip-vis.
[57] Goyette, N., Jodoin, P.M., Porikli, F., Konrad, J., Ishwar, P., 2012. changedetection.net: A new change detection benchmark dataset, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. pp. 1–8. doi:10.1109/CVPRW.2012.6238919.
[58] Wang, Y., Jodoin, P.M., Porikli, F., Konrad, J., Benezeth, Y., Ishwar, P., 2014. CDnet 2014: An expanded change detection benchmark dataset, in: IEEE Computer Society Conference on Computer Vision and

Pattern Recognition Workshops. IEEE Computer Society, pp. 393–400. doi:10.1109/CVPRW.2014.126.
[59] S. Lee, J. Xin, S. Wesland, Evaluation of image similarity by histogram intersection, Color Research and Applications 30 (2005) 265–274.
[60] P. Gil-Jiménez, H. Gómez-Moreno, J. Acevedo-Rodríguez, and S. M. Bascón, “Continuous variance estimation in video surveillance sequences with high illumination changes,” Signal Processing, vol. 89, no. 7, pp. 1412–1416, 2009, doi: 10.1016/j.sigpro.2009.01.013.
[61] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," inProc. Int. Conf. on Computer Vi-sion and Pattern Recognition, Vol. 2, IEEE, Piscataway.
[62] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," in Proc. Eur. Conf. on Computer Vision, Lect. Notes Comput. Sci. 1843, 751-767 2000.
[63] Işık, Şahin, Kemal Özkan, Serkan Günal, and Ömer Nezih Gerek.2018, "SWCD: a sliding window and self-regulated learning-based background updating method for change detection in videos." Journal of Electronic Imaging 27, no. 2 (2018): 023002.

Republic of Iraq
Ministry of Higher Education
And Scientific Research
University of Kerbala
Dean Office

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة كربلاء كلية العلوم
مكتب العميد

No : 73 / 5 / م
date : 2021 / 6 / 28

العدد : 73 / 5 / م
التاريخ : 2021 / 6 / 28

Subject: Paper Publishing Acceptance

Dear Mohamed Qasim Mohamed, Tawfiq A. Al- Asadi

We are delighted to inform you that your manuscript entitled "**Statistical - Spatial Technique for Video Object Detection**" has been accepted in the 9th International Conference of Applied Science and Technology (ICAST2021) Conference Proceeding.

The accepted manuscripts will be published in the AIP conference proceedings which is Scopus indexed and has an impact factor.

Once your article is uploaded online, a notification email will be sent to you.
If you have any enquiry, please, do not hesitate to contact us.

Yours Sincerely,

Karbala University
Dean Office
Science College

Jasim Alawadi

Assist. Prof. Dr. Jasim Hanoon Hashim Al-Awadi
Chairman of ICAST 2021
Dean of the Faculty of Science, University of Kerbala, Karbala, Iraq.
Phone: +9647831556219
E-mail: jasem.h@uokerbala.edu.iq

ICAST 2021

College of Science

College of Computer Science & Information Technology

Certification of Participation

Awarded to

Mohamed Qasim Mohamed, Tawfiq A. Al- Asadi

For their participation in the 9th International Conference on Applied Science and Technology which was held on 16th-17th June 2021, College of Science & Collage of Computer Science & Information Technology, University of Karbala

To verify the content please scan the QR Code

Karbala University Dean Office Science College

Jasim Alawadi

Assist. Prof. Dr.Jasim Hanoon Hashim Al-Awadi

Conference Chairman

Statistical - Spatial Technique For Video object Detection

MOHAMED QASIM MOHAMED AND TAWFIQ A. AL-ASADI

Department of Software, Faculty of Information Technology, University of Babylon, Hillah, Iraq
Corresponding author: moh.swd1988@gmail.com

Abstract. Object detection is the most important stage due to its great impact on next stages. Moving objects detections in real environment have several challenges such as dynamic background, illumination changes (gradual, sudden) and noise etc. Traditional techniques of background modeling do not have the ability to face these challenges but can deal with the static background. Based on statistical , spatial and temporal features we have proposed hybrid technique to model background and detecting moving object, it can deal efficiently with challenges in real environment. The proposed system consists of two stages, the first one is construction of statistical model by computing mean and standard deviation for each pixel and spatial model that are created by calculation Center Symmetric Local Binary Patterns which consists of a group of histograms for each pixel. The second one is foreground detection by checking each pixel in current frame with predefined thresholds if the value of pixel raises out of these thresholds then the statistical model will be applied otherwise the spatial model is applied. The contribution of this paper is creating model hybrid between statistical and spatial features which lead speed and accuracy cause statistical features are fast in calculation and spatial features characterized by accuracy. Experimental results proved that our proposed method lead to approximate 2% increase in the accuracy comparing with the traditional benchmark methods by using the exact dataset.

Keywords: Background modeling , foreground detection, texture, histogram, local binary pattern , benchmark methods .

1. INTRODUCTION

With no doubt, the video surveillance has become an important part of urban and home security, and traffic nowadays. Due to the massive increase in video data nowadays, automatic video analysis is necessary [1]. Background modeling and foreground detection are essential steps for multiple video processing practical and industrial applications. For example the intelligent video surveillance which is able to detect some patterns of criminal activities and accidents. Or in optical motion capture. And content based video retrieval, which is able to classify and store videos based on the content of the video. Another prominent application is human computer interaction, activity recognition and many other applications [2].

Strategies of conventional background modeling concentrate on the temporal changes between sequential frames to detect moving objects. Average, median and standard deviation represent traditional methods that rely on thresholding the difference between reference model and current frame such as these methods fail to address some critical situations. For instance the illumination changes, dynamic background and noise do not have the ability to construct multi-modal background. Background model must be strong to overcome a number of challenges by building a multi-model and be adapting with challenging situations[2].

The methods of background model are classified into two types, the first one is the pixel based method which is considered with the lower pixel level. The second one is the region (block) based method, which is considered with a higher reign level. The classification depends on the method is used to build background model. In pixel based method, each pixel remains independent when building the model. This case is the most common one, but does not take into account temporal and spatial features. These methods are not effective for environments that have rapid changes in the background[3].

In region (block) based methods, the frames are divided into overlapped or non-over lapped blocks, the features associated with the block are calculated such as correlation, covariance and histogram to model the block. The block based methods supply high performance against dynamic background. In non-overlapping methods, the moving object is segmented and produced gruff shape than the overlapping methods. In overlapping methods, each pixel is modeled based on features of block, the detection of moving object is done at pixel level, so, the object's shape is better than non-overlapping methods [4].

Depending on the most important properties and characteristics of the image, features could be grouped into the following types [5]:

نبذة مختصرة

إن أنظمة المراقبة بالفيديو لها أهمية كبيرة في مراقبة المناطق الأمنية الحساسة مثل الحدود ، والمصارف ، والطرق السريعة ، والأماكن العامة ، وما إلى ذلك. ادى التقدم في تخزين السعة الكبيرة ، وقوة الحوسبة والتوافر إلى التطور في هذا المجال. لجعل أنظمة المراقبة بالفيديو "تعمل بشكل آلي" يتطلب هذا خوارزميات قوية وموثوقة بها للكشف عن الأجسام المتحركة ، وإزالة الظل.

في هذه الرسالة ، اقترحنا نظام اكتشاف بشكل تلقائي والذي يتكون من مرحلتان. المرحلة الأولى هي اكتشاف الكائن المتحرك، تتضمن هذه المرحلة خطوتين ، الأولى هي بناء نموذج الخلفية عن طريق تحديد N من الصور الفيديوية لإنشاء موديل احصائي و موديل نسيجي الموديل الاحصائي يبنى بحساب المتوسط الحسابي والانحراف المعياري لكل بكسل اما الموديل النسيجي فيتكون مجموعة من المدرجات التكرارية (الموديلات)، لكل بكسل التي يتم بنائها باستخدام(CS-LBP). الخطوة الثانية هي اكتشاف الكائنات اذا كانت قيمة البكسل المحدد في الاطار الحالي ليست ضمن العتبة المخصصة لنطاق الاطار الاحصائي هذا يعني انه من الصعوبة تصنيف البكسل هل ينتمي الى الكائن او ينتمي الى الخلفية لذلك يصنف عن طريق الموديل النسيجي .

المرحلة الثانية هي اكتشاف الظل وازالته ، طريقة كشف الظل تتحقق بخطوتين ، الخطوة الاولى اختيار البكسلات المرشحة كظل بواسطة مكتشف الظل الضعيف والذي يعتمد على الخصائص الطيفية، الخطوة الثانية تصنيف كل بكسل مرشح ظل على أنه إما كائن أو ظل من خلال قياس الترابط النسيجي في الإطار الحالي مع النسيج في موديلات الخلفية بالإضافة إلى خاصية الألوان التي تعتبر مكملة للخصائص النسيجية.

توضح التجارب على الفيديوهات كفاءة الطريقة المقترحة لاكتشاف الكائن مقارنة بطرق قياسية عن طريق زيادة precision إلى (3.6%). فيما يتعلق بطريقة ازالة الظل المقترحة فان النتائج التجريبية تشير إلى كفاءة الطريقة المقترحة مقارنة بالطرق القياسية عن طريق زيادة precision إلى (7.6%).

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جـامعـــة بـابـــل
كلية تكنولوجيا المعلومات
قسم البرمجيات

اكتشاف كائنات الفيديو وازالة الظل بالاعتماد على الخصائص الاحصائية والمكانية

رسالة مقدمة إلى
مجلس كلية تكنولوجيا المعلومات - جامعة بابل كجزء من متطلبات
نيل درجة الماجستير في تكنولوجيا المعلومات / البرمجيات

من قبل

محمد قاسم محمد جودة

بإشـــراف

أ.د. توفيق عبد الخالق عباس الاسدي

1444 هـ   2022 م