Review on Real Time Background Extraction: Models, Applications, Environments, Challenges and Evaluation Approaches

Maryam A. Yasir [✉]
University of Baghdad, Baghdad, Iraq
maryam.a@sc.uobaghdad.edu.iq

Yossra Hussain Ali
University of Technology, Baghdad, Iraq

Abstract—In the computer vision, background extraction is a promising technique. It is characterized by being applied in many different real time applications in diverse environments and with variety of challenges. Background extraction is the most popular technique employed in the domain of detecting moving foreground objects taken by stationary surveillance cameras. Achieving high performance is required with many perspectives and demands. Choosing the suitable background extraction model plays the major role in affecting the performance matrices of time, memory, and accuracy.

In this article we present an extensive review on background extraction in which we attempt to cover all the related topics. We list the four process stages of background extraction and we consider several well-known models starting with the conventional models and ending up with the state-of-the art models. This review also focuses on the model environments whether it is human activities, Nature or sport environments and illuminates on some of the real time applications where background extraction method is adopted. Many challenges are addressed in respect to environment, camera, foreground objects, background, and computation time.

In addition, this article provides handy tables containing different common datasets and libraries used in the field of background extraction experiments. Eventually, we illustrate the performance evaluation with a table of the set performance metrics to measure the robustness of the background extraction model against other models in terms of time, accurate performance and required memory.

Keywords—Video surveillance, Background extraction, Foreground object, Evaluation metrics

1 Introduction

In the last two decades and with the rapid development of sensors and the increasing safety concerns, detecting moving objects has been one of the most essential topics in

the computer vision field [1]. Real time applications with different environment and many challenges are available with diverse interest where the foreground data can be used in tracking, synopsis and anomaly detection [2][3][4].

Foreground detection has been applied to vast scope of smart video surveillance systems of human activities like road and airport traffic surveillance, human detection and tracking, analysis of sports video surveillance, and recognition of gesture in the interface of human-machine etc. [5][6]. Background extraction is the most popular technique employed in this domain to extract the foreground moving objects taken by stationary surveillance camera [7].

In this article, we shed light on the publications of different background extraction techniques. In section 2.1, we discuss the four stages of background extraction process. In section 2.2, we review the background extraction models and describing how the model works and how it reacts according to different environments and challenges. In section 3 and 4, we focus on reviewing the implementation of background extraction in video surveillance applications and the environments of background extraction-based video surveillance. While section 5, is about the challenges of background extraction models applied on video surveillance.

Furthermore, in section 6&7 we list tables of datasets and libraries used in the field of background extraction model and finally, section 8 is illustrating the performance evaluation and the set performance metrics to assess the background extraction models in terms of time, accuracy and memory.

2   Background Extraction

Background extraction remains an active interesting area in the domain of computer vision. Many models have been developed for extracting the background to detect the foreground. Background extraction models are going through ordered stages to obtain the foreground masks.

2.1   Background extraction stages

In surveillance systems, where background extraction model is applied, the background extraction model is usually comprising of sequenced stages as the following:

1. **Background initialization**: We generate or construct the first background scene among numbers of video frames.
2. **Background modeling**: We describe a representation model for the background scene.
3. **Background maintenance**: We update the background model according to changes that occur within time. The update is done according to the prior scene, the recent scene and foreground mask.
4. **Foreground detection**: It is the final stage, where we classify the pixel into background or foreground object, and this is done by comparing the background scene with the recent scene.

In each stage, different algorithms are employed; some algorithms are offline where all data is required at once like in background initialization stage. On the other hand, the online algorithms require to take data one after another incrementally like in background maintenance stage.

Some pre-processing operations might be taken, including framing the video or changing the color space. Post-processing also could be done by applying various algorithms to overcome a specific challenge in background extraction process [8]. Figure1 Illustrates the overview of background extraction stages.
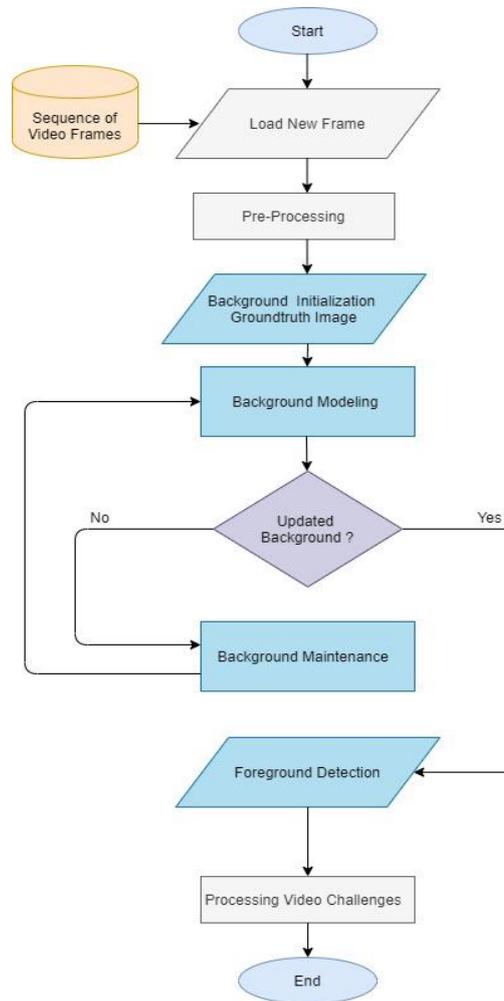

Fig. 1. The overview of background extraction stages

2.2   Models

Background extraction models have been classified in many works; the following is the classification of these extraction models:

1. Basic models

The basic models are the conventional methods in which used earlier for the background extraction. In basic models, each pixel of the current frame is classified either as a background or foreground according to a threshold differences with the background model frame. These are quite simple and fast techniques, but they come with low performance, especially with complex and challenging background images [9]. These models are dependable on threshold and are not supporting the multiple model background distributions [10]. An example of basic models is mean model [11], median model [12][13][14][15][16], and analysis of histogram model [17][18].

2. Mathematical models

The mathematical models comprise a verity of models like statistical (probabilistic) models, dempster– Schafer model [19] and fuzzy models [20]. Statistical models, are either parametrical or non-parametrical .In paramedical model, the main method used is the Gaussian Mixture Model (GMM) [21], it is one of the most robust algorithms in respect to different background variation such as lighting changes and multi-modal changes [22]. In GMM, there are three important parameters to be considered: the component's number, detection threshold and learning rate [6]. Each pixel in GMM is represented in multiple Gaussian components [5]. However, the GMM suffers from slow recovery and poor performance with unexpected lighting changes and background abnormal motions [23]. Moreover, due to the parametrical nature of GMM model, the efficiency can be diminished either by selecting inaccurate parameters or by the time consumed in setting the parameters in which makes GMM unpreferable with real time applications [24]. The parameters are: group of components, learning rate and threshold of classification.

For years, researchers work on enhancing the GMM to prevail the challenges, like in Boosted Gaussian Mixture Model where this method boosts the performance of GMM using both color space classification and dynamic learning to update the background model[25]. Improved Gaussian background modeling (GBM) is another enhanced model in which wavelet denoising is applied on foreground object, this model achieved a better performance in respect to shadow and lighting changes challenges [26]. On the other hand a Self-Adaptive Gaussian Mixture Mode improved the speed performance of GMM four times by using a block of pixels instead of considering one pixel value[27].

Visual Background extractor (ViBe) is also introduced as a background extraction algorithm, which uses a novel techniques that considering the effect of a value in a multicolor space to be restricted to a local neighborhood[28][29][30]. In ViBe classification is done by comparing a value to its closest values in the set of samples instead of using the probability density function (pdf) to update the background model for

obtaining the better results [31]. Substance Sensitivity Segmenter (SuBSENSE) algorithm is influenced by ViBe algorithm, it fuses color-LBSP with local adaptive sensitivity to perform much robust performance[32] in CDnet 2014 dataset [33].

Another method used as statistical model is the non-parametric model for kernel density estimation (KDE) [34][35]. In this model the pixels density is being estimated by kernel function which denoted by histogram of pixels with most recent values [10]. This non-parametric method is directly estimating the intensity of pixels using the density function in contract with a parametric method. This leads to a clearer representation of pixels intensity[36]. KDE performs better in static background environment for long period of time and a moderate speed motions for the objects. While, KDE fails to solve the challenges of multi background model, multiple moving objects, and slow movement of the objects [21]. It is not suitable also with real time applications, due to high computation time consumed for estimating the density of pixels from many needed samples over time [36]. Moreover, Pixel-Based Adaptive Segmenter (PBAS) is another non-parametric method in which performs well with shadow and thermal scenes [37][38].

Many methods are classified under mathematical models like mean shift which, performs well with multiple background models[10][39], dempster – Schafer [19] and fuzzy models [20][40] where both could deal with ambiguity, incompetence and inaccuracy of data caused by many challenges [41].

3. Single processing models (Filters models)

Filter model is a predicting model which estimates the background by expecting the information of a pixel or a block of pixels from the previous pixel, this information could be the orientation or the value of the pixel [6]. Many examples of filters have been applied in this approach such as Kalman filter [42][43], Tchebychev filter [44], Wiener filter [45], particle filters [46], Correntropy filter [47], and optical flow [48]. This model is not applied in real time applications because of its complexity, inaccuracy detection and highly prone to noise, even though it could be a good choice for images with progressive lighting changes [6].

4. Clustering algorithms models

Clustering model uses the color or intensity of the pixel in detecting the foreground object. In the background frame each pixel is characterized by its correspondent cluster of pixels. Pixels for the current frame are compared to clusters, to decide whether it is a background pixel or not.

Many algorithms have been used in this model such as K-means algorithm [49], background reconstruction algorithm [50], and Codebook model [51][52].

Codebook model is used as a compression method to represent the value of background pixel in a long period of frequent image sequence and this method uses the intensity of color as information to detect the foreground objects [51][52][7]. Codebook model performs better with noise and dynamic backgrounds, but fails with unexpected lighting changes, shadow movements, similarity of foreground color to background and

slow objects movements [36][52]. Mandy improvements have been done to codebook model in [53][54][55][56].

5. Machine learning models

Machine learning models include many models like, subspace learning which is either reconstructive unsupervised techniques that provide a good data estimation [57][58][59][60] or discriminative supervised techniques that provide a good data separation classification [61][62] or mixed [63].

Another examples of machine learning concept are: robust subspace tracking [64], support vector machines (SVM) [65][66][67][68][69][69], robust subspace learning [70][71][72][73], tensor decomposition [74][75][76], deep learning [77][78][79][80][81][82] and neural networks [83][84][85][86][87] which have been widely adopted since 2014 [88], mainly because of the great increase of processing power in new hardware as well as the availability of many training datasets [89].

Whilst machine learning provides the concepts that enable us to acquire the pixel representation of the background in both supervised and unsupervised behaviors; the extraordinary breakthrough began with the supervised training of an 8-layer neural network using a dataset containing millions of images [90][91][92] for the task of foreground/background separation, the application of DNNs was successful in enhancing foreground detection, background subtraction, generating the ground-truth, as well as for learning deep spatial features.

Recently, in the area of computer vision, particularly in object detection and segmentation, the use of deep convolutional neural networks (CNNs) has produced remarkable results [5][93]. The use of CNNs in foreground detection was initially suggested by Braham and Droogenbroeck [77]. Background images produced by median filtering along with patches of frames, were adopted as input for the CNN. The CNN, which has multiple convolutional layers and fully connected layers, provides a map that represents the likelihood of every pixel being a foreground pixel. CNNs easily achieve the distinction between the foreground and the background. However, given the small scale of the dataset, as only dozens of videos are available in CDnet 2014 [33], it can be challenging to learn the high-level information.

In recent years, significant advancements were made in the field of computer vision due to the utilization of deep learning. Even though deep convolutional neural networks (CNNs) were originally devised for the task of image classification, they were utilized in computer vision for their robust performance in the extraction of high-level characteristics. For many years, the use of CNNs in computer vision was restricted due to the limited computing power, and the small size of accessible datasets [90][91]. The machine learning methods achieve good results to overcome the challenge of lighting changes. However, these methods do not work properly in scenes with abnormal dynamic background movements and non-static shadows as well as an extended processing period [94].

Many scholars have been studied the background extraction since the early convolutional models up till the state-of-the-art models like deep learning and CNN models. Regardless of various background extraction models, there is no unique model that solves all the challenges in respect to real time applications. Fusion of different models

and strategies have been experienced in numerous papers to achieve a better performance [7].

Table 1 summarizes the background extraction models, the model environments, the model strength, and the model weakness.

Table 1. The background extraction models, environments, strength, and weakness.

Background Extraction Models	Environment	Model strength	Model weakness
Basic Models: -Mean [11] -Median [12] [13][14][15][16] -Histogram [17][18]	Traffic surveillance in road and highways and Airport surveillance	Simple and fast techniques which still dominant in the field of background extraction with real time applications	Low performance with challenging background images and these models are dependable on threshold and are not supporting the multiple model background distributions.
Mathematical Models: **-*Statistical Models*** Parametric models -GMM[21][22][23][24][25][26][27] -ViBe [28][29][30][31] -SuBSENSE [32] **Non-parametric models** -KDE [34][35] -PBAS [37][38] *Mean shift [10][39]* *Dempster – Schafer [19]* *Fuzzy models [20][40]*	Nature environments especially with a very challenging environment like maritime & submarine	-**Parametric Models** performing well in respect to different background variation such as lighting changes and multi-modal changes -**Nonparametric models** perform better in static background environment for long period and a moderate speed motions for the objects Mean-shift performs well with multiple background models Both Dempster-Schafer and Fuzzy models could deal with ambiguity, incompetence and inaccuracy of data caused by many challenges and they consume less time	-**Parametric Models** are not performing well with sudden lighting changes and background abnormal motions -**Nonparametric models** fail to solve the challenges of multi – background model, multiple moving object and slow movement of object It is not suitable also, with real time applications due to high computation time Mean-shift does not perform well with multi-dimensions of the space. For both Dempster-Schafer and Fuzzy models, they still need to be experimented on various datasets.
Single Processing Models: -Kalman filter [42][43]	Human activities	This model is a good choice	This model is not applied

	Environments	for images with progressive lighting changes	in real time application because of its complexity, inaccuracy detection and highly prone to noise
-Tchebychev filter [44] -Wiener filter [45] -Particle filters [46] -Correntropy filter [47] -Optical flow [48]			
Clustering Algorithms Models: - K-means algorithm[49] - Codebook model [51][52][53] [54][55][56]	Human activities environments and Nature environments	This model performs better with noise and dynamic backgrounds.	This model fails with unexpected lighting changes, shadow movements, similarity of foreground color to background and slow objects movements
Machine Learning Models: -Reconstructive [57][58][59][60] -Discriminative [61][62] -Mixed [63] -Robust subspace learning [70][71][72][73] Support vector machines [65][66][67][68][69] -Tensor decomposition [74][75][76] -Robust subspace tracking [64] -Neural networks [83][84][85] [86][87] -Deep learning [77][78][79][80] [81][82]	Human activities environments like vehicle counting, vehicle and persons detections and Nature environments	This model performs better results for lighting changes.	This model does not work properly in scenes with abnormal dynamic background movements and non-static shadows and have an extended processing period

3 Implementation of Background Extraction in Applications

Background extraction models have been applied in many surveillance application systems. The following are background extraction-based surveillance applications utilized in various realms:

3.1 Surveillance system of human activities

In the last decades, surveillance systems of human activities have gained a remarkable interest due to security concerns [95]. Many Surveillance application systems are exploited in terms of human activities [96]. The most common environment is transportations environment which comprises, traffic sense for roads, highways, and airports. The other environments include supermarkets and warehouses, Automated Teller Machine (ATM), and military environment [41][96]. The main target for such surveillance applications is to detect and track the interested objects [8][97]. In traffic scenes, it is more common to detect the stopped vehicles or estimate the traffic density, whether it is empty or jam [14][98][99] and sometimes its required to count the vehicles number [100].

Background extraction applications can also detect the congestion and can be applied to parking detection for free places [101][102][103] or legal parking [104][105]. Moreover, Human activities applications are applied for security purpose in trains and airports to track the lost baggage. Marina [106][107][108][109][110] and stores [111][112][113][114][115] are other environments where human activities-based background extraction are used.

3.2 Surveillance system of Nature

In Nature surveillance system, the target is to detect or track an animal, insect or even an unwanted object in the nature environment like forests, rivers, and lakes [116][117]. The object detection and tracking are required in many studies that are conducted for analyzing the behavior of these animals and insects. Different behaviors could be observed to evaluate the influence of animals to their places of plants or water [118][119][120][121] and to what extent can be affected by weather conditions and climate changes [122][123][124][125]. It is also critical for estimating the interaction among insects inside the same group, like in honeybee swarms [126] and mice[127][128]. Surveillance of endangered animals is one of the applications that is widely employed to monitor, track, and record their movements [129][130]. Automated nature surveillance system is playing an essential role for data collection used in the field of biodiversity. There is therefore a need for a convenient background extractions system to detect the foreground target which eventually enrich studies and help to protect Nature.

3.3 Human Machine Interaction (HMI) applications

Many applications are classified under HMI applications using background extraction, some are related with games where player can use a color webcam with PlayStation to detect motion, color, and gestures, which deals with virtual image or silhouette [41]. Kinect by Microsoft also uses different background extraction techniques like GMM, KDE, codebook and others[131]. HMI is also applicable in art where graphic designers can use mocap sensors to track and record actor and object motions, then use the information to animate a digital model for the actor/object in 2D or 3D animations [132].

In virtual environment, many gesture recognition applications are applied, like sign language translation, human-machine interface, behavior analysis, robotics, teleseminars, object manipulation and games [133]. The aforementioned applications require detecting and tracking the gesture of the hand as an example of human machine interaction applications using the background extraction techniques to detect the aimed objects [41][132].

3.4 Video coding and matting

In teleseminars using digital movies and video phones, we only send the moving foreground with the key frame as content based [134]. The content usually consists of

semantic information like shape and texture where they called video object plane in MPEG-4. Background extraction is used in content-based video coding for a faster transmission in wireless networks [41]. For video matting the target is different, the background is subtracted from the input scene and substituted with another background. Thus, background extraction used as a first step in this process[135].

There are numbers of applications using the background extraction technique to detect the foreground objects like in carried baggage detection [136], fire detection [137], organic light-emitting diode (OLED) defect detection [138] and obtaining approximated geometric model for the object [139]. All background extraction-based applications involve moving foreground detection and have their features that are related to challenges of camera location, foreground object type, and the environment.

4　Background Extraction Based Environments

Background extraction is applied in a variety of environments where video surveillance system is employed. In each environment, there are many characteristics and aspects to be considered, different background scenes and different moving objects to be detected and eventually different challenges to be solved related to each individual environment. Video surveillance with stationary camera can be involved in the following environments:

4.1　Human activities surveillance

Background extraction and foreground detection has been employed in human activities surveillance since the last two decades. Detecting the moving objects is our major focus in this review, as for surveillance of human activities, the target is to automatically detect the foreground objects and to compute the required statistics. Various appealing foreground objects can be detected like, individuals [140], vehicles [141], baggage [136], planes [16], canoes [142] and products in warehouses [143]. Obviously, there are many examples of human activities environments and they are as follows:

1. Transportation scenes

Traffic surveillance [144] is one of the most exposed transportation scenes, where background extraction is used to detect or track the foreground moving objects. Traffic surveillance includes highways and roads environments, the foreground detection can be employed to count stopped vehicles [14][15][98][99][141], estimating the road congestion [104][105], or parking surveillance [101][102][103][145][146][147][148].

Traffic surveillance is affected by different factors, for example the camera location factor where most of the time it is a stationary camera. Sometimes traffic senses are aerial videos taken by a drone [149][150] or even high-resolution videos taken by a satellite [151][152][153]. Also, the quality of the camera is highly considered in terms of the video characteristics in traffic scenes. Most of the time, low-quality CCTV camera comes with low number of frames which eventually affects the whole process of detection. Another factor is the environment itself and its challenges, like (shadows, lighting changes, dynamic background, etc.) these challenges can vary from one

environment to another. Foreground object of interest are different in their structure, actions, and colors. Automobiles, motorbikes, trucks, walkers, bicyclist are examples of these objects.

Airport surveillance is an additional transportation scene where security is the main aspect of this environment. It is important to track and monitor the planes parking, fueling cars, luggage vehicles and workers. Different challenges arise due to environments like bad weather, lighting changes and non-uniform lighting. The color and shape of the foreground object is another challenge in airport surveillance scene. Marine surveillance is part of the transportation surveillance scenes. This surveillance is meant to track and monitor boats that normally observed in rivers, seas, and canals. Many papers have been done to address this environment using different approaches to detect the foreground objects based on background extraction techniques. Examples of these approaches are: Bayesian combined to segmentation method, independent background extraction to overcome the dynamic background challenge in water [154][155][156], GMM [157], median [158] and many others. The most challenging factor in the marine environments is the dynamic background due to water movements [41].

2. Warehouse scenes

Warehouse surveillance is being one of the common demands in the marketing field, more specifically to study the customer behavior [159][160][112][113][114]. During shopping, customers may stay for a long time in a specific place checking a specific product. Sometimes it's needed to check on the numbers products left on the shelves [143], such information could be provided by the Radio Frequency Identification (RFID) tag [160] which normally attached to the product itself, but it is not always the best choice in terms of cost and time needed to attach the tag to each product. Thus, applying the background extraction techniques on videos taken by stationary camera can be an alternative solution used for that purpose [143].

3. Military scenes

Military surveillance is crucial for security reasons, it is supporting the protection of country borders, bases, security checkpoints, and many others military sites. Military surveillance is into using sensors to detect the target in the battlefield. These sensors are either infrared camera [161] or a radar [162] to detect and monitor a vehicle, person, or both as a target. Many challenges arise here due to environment like (desert, forest, maritime, etc.) or even due to bad weather like (snowing, raining, etc.) and lighting changes like night scenes. In military surveillance, different background extraction models have been used to handle the foreground detection such as, median, temporal median and type2 fuzzy GMM [161][162].

4.2   Nature surveillance

Nature is another surveillance environment where it covers the animals and insects' surveillance and other natural environmental scenes. Following is a brief description of the two types:

1. Animals and insects' scenes

Plenty of researches have been done to study the behavior of animals and insects in a variety of environments. Background extraction technique is widely employed to monitor and track the detected foreground objects for analytical purposes. The analysis may consider the interaction of insects between each other in the same community like in honeybees [126] and the interaction with environments, like in birds with weather and climate changes [118][119][120][121][122]. Another purpose could be to study the behavior of animals that endangered, like in many species of wolf, fox, and castor [129][130].

2. Natural environments scenes

Nature surveillance sometime involve detecting the foreign objects in a specific nature like in seas, rivers, forests, and oceans. The main purpose can be to protect that environment and its biodiversity. The foreground target could be any floating object like wood or bottles [116] [117][163][164].

4.3 Sport events surveillance

Sport is another domain where the background extraction is applied. Decisions must be taken in many games like using the Hawk Eye system in tennis, surfing actions [165] and the recently used Video Assistant Referee (VAR) system in football. Background extraction is also applied in surveillance systems to follow up the athletes and body-builders routines [166][167].

5 Background Extraction Based Challenges

Background extraction has experienced many challenges due to implementations in various environments and due to different foreground targets. Background extraction challenges has been addressed in many papers, and can be classified into the following main challenges categories:

5.1 Environment challenges

Each environment has its characteristics and challenges, like in traffic and nature surveillance, the scenes are daily exposed to different weather and lighting changes during the day. Various challenges arise like shadow, illumination, and baseline. Whether in outdoor scenes like highway roads and pedestrian scenes or indoor scenes like office scenes [168]. Another most common challenge is the bad weather, it includes blizzard, skating issues, snowfall, and wet snow [33]. In marine surveillance environment, scenes face an additional challenge, they normally suffer from dynamic background caused by the movement of water like in fountains, lakes, and rivers [33][169]. All aforementioned challenges and many others are more associated with each individual environment making the detection of foreground objects harder.

5.2 Foreground target challenges

The foreground target could be a challenge itself in many cases. One of the most common challenges is when the color of the target is similar or close enough to the color of the background which leads to false detections. Infrequent stops of the foreground target (which is called sleeping foreground) is also another challenge [168]. The target becomes immobile as a part of the background for a certain period of time, like when vehicles stop in the road intersections or in heavy roads and traffic jams [170].

5.3 Background challenges

The background scene encounters some challenges like intensity variation which leads to inaccurate histogram estimation [41]. Dynamic background is another common challenge to extract the foreground objects like in fountains, lakes and falls scenes.

5.4 Camera challenges

Many challenges are a result of the surveillance cameras issues. The position of the camera could be one of the issues where an environmental cause like wind can move the camera [171][172]. Another case would be the shadow, bad contrast, and long distance of the object [149][150] [173][174][175][176][177]. The video quality is also affected directly by the camera low quality, low frame rate [178] and camera jitter [168].

5.5 Computation time challenges

In the real time applications, reducing the computation time is one of the main metrics to be considered. Therefore, it is important to choose the suitable background extraction model for each individual environment. For example, with traffic surveillance it is common to apply uni-models like basic methods of histogram and median due to their speed.

6 Datasets

The selection of a convenient dataset plays a crucial role in the evaluation of the various background extraction models. There are two types of datasets, the public and private. Private dataset is more specific to a single research where it is collected by the author him/herself. While public datasets are regularly available on the internet to be downloaded and used in research experiments. In this section, we list some of the most common public datasets with environments of visual surveillance of human activities and visual surveillance of animals and insects. In table 2 and table 3, we list the datasets with provided years, scenarios and scenes included, and reference links.

Table 2. Visual surveillance of human activities datasets

Visual Surveillance of Human Activities			
Dataset	*Dataset Scenario*	*Limitation/ Privileges*	*Reference Link*
Wallflower dataset named after wallflower Algorithm for background maintenance provided in **1999** [45]	-Moved object -Time of day -Light switch -Waving trees -Camouflage -Bootstrapping -Foreground aperture	This dataset is limited to a person detection with one Ground-Truth (GT) image by video	https://www.re-searchgate.net/figure/The-Wallflower-dataset-scenarios_tbl2_263618322
Institute for infocomm Research (**I2R**) dataset provided in **2004** [179]	-Bootstrap -Campus -Curtain -Escalator -Fountain -Hall -Lobby -Shopping mall -Water surface	This dataset did not cover a huge spectrum of challenges and the GTs are also limited to 20 by video	https://github.com/I2RDL2/ASTAR-3D
Multicamera Human Action Video (**MuHAVi**) dataset provided in **2010** [180]	**Human actions** -Walk turn back -Run stop -Punch -Kick -Shot gun collapse -Pull heavy object -Walk fall -Look in car -Crawl on knees -Wave arms -Draw graffiti -Jump over fence -Drunk walk -Climb ladder -Smash object -Jump over gap	This dataset consists of human action video (Mu-HAVi) ,8 cameras used, and 17 action classes performed by 14 actors. It contains frames before and after the original action. This dataset is suitable for background extraction and tracking approaches	http://velas-tin.dynu.com/MuHAVi-MAS/
Change Detection **CDnet 2012** dataset provided in **2012** [169]	-Baseline -Camera jitter -Dynamic background -Intermittent object motion -Shadow -Thermal	This dataset comes with less scenarios.	http://changedetection.net/
Change Detection **CDnet 2014** dataset provided in **2014** [33]	-Bad weather -Baseline -Camera jitter -Dynamic background -Intermittent object motion -Shadow -Thermal -Low frame rate -Night videos - PTZ -Turbulence	It is the most applicable dataset due to extended scenarios.	http://changedetection.net/

Scene Background Initialization dataset (**SBI**) provided in **2015** [181]	-Board -Candela_m1.10 -CAVIAR1 -CAVIAR2 -CaVignal -Foliage -Hall & monitor -Highway I -Highway II -Human body2 -IBMtest2 -People & foliage -Snellen -Toscana	This dataset comprises various scenes which included 14 images sequences with ground-truth and MATLAB scripts	http://sbmi2015.na.icar.cnr.it/SBIdataset.html
Labeled and Annotated Sequences for Integral Segmentation Algorithm (LASIESTA) dataset provided in 2016 [182]	-Bootstrap -Moving camera -Simulated motion -Illumination changes -Occlusions -Camouflage -Simple sequences -Modified background -Weather conditions	This dataset consists of indoor and outdoor scenes, it uses in both pixel-level and object-level, it is suitable for both moving object detection and tracking objects approaches. It contains scenes of static and moving cameras.	https://www.gti.ssr.upm.es/data/lasiesta_database.html
Scene Background Modeling dataset (**SBMnet**) provided in **2016** [183]	-Basic -Intermittent motion -Clutter -Jitter -Illumination changes -Background motion -Very long frame rate -Very short frame rate	This dataset covers many indoor and outdoor challenges for surveillance, video database scenarios and smart environment. The dataset contains both personal videos and public collections.	http://scenebackground-modeling.net/
Scene Background Modelling for moving object detection on RGBD videos (**SBM-RGBD**) dataset provided in **2017** [184]	-Illumination changes -Color camouflage -Depth camouflage -Intermittent motion -Out of sensor range -Shadows -Bootstrapping	This dataset is collected for moving object detection on RGBD videos. It provides scenes, ground-truths, and evaluation scripts for various challenges	http://rgbd2017.na.icar.cnr.it/SBM-RGBDdataset.html

Table 3. Visual surveillance of animals and insects

Visual Surveillance of Animals and Insects		
Dataset	*Dataset Scenes*	*Reference Link*
Aqu@theque dataset provided in **2007** [185]	Fish in tank	https://sites.google.com/site/thierrybouwmans/recher-che---aqu-theque-dataset
Caltech resident intruder mice dataset provided in **2012** [186]	Social behavior recognition of mice	http://people.vision.caltech.edu/~dhall/projects/Merg-ingPoseEstimates/
e Mammal dataset provided in **2013** [187]	Camera trap sequences	https://emammal.si.edu/
Fish4knowledge dataset provided in **2014** [188]	Fish in open sea	http://groups.inf.ed.ac.uk/f4k/
Caltech Camera Traps (CCT) dataset provided in **2018** [189]	Sequences of images taken at approximately one frame per second for census and recognition of species	https://beerys.github.io/CaltechCameraTraps/

7   Libraries

Many libraries are employed in the field of background extraction and exploited for evaluations; we list the most well-known libraries in the following table 4.

Table 4. Background extraction libraries

Libraries	Target	Reference Link
OpenCV [190] in **2018** [191] in **2018**	It is cross-platform C and C++ Library, it has more than 2500 conventional and state-of-the-art computer vision and machine learning algorithms.	https://opencv.org/
Background Subtraction (BGS) [192] in **2013** [193] in **2014**	It is an extended OpenCV library with C++ framework, it has more background extraction algorithms.	https://github.com/andrewssobral/bgsli-brary
Low-Rank and Sparse (LRS) [194] in **2016**	It is designed for video movement segmentation. It has various algo-rithms of sparse and low rank decomposi-tion written in MATLAB. LRS is also employed in computer vision with many tensor and matrix based algorithms.	https://github.com/andrewssobral/lrslibrary
Independent Multimodal Background Subtraction (IMBS) [155] in **2012** [154][195] in **2014**	It is C++ library builds a multi-modal method to overcome the lighting changes, dynamic background, and camera jitter, IMBS uses OpenCv functions.	https://github.com/imbs-hl/ranger

Background Subtractor CNT [196] in **2019**	It is a library written in C++ and based On OpenCV, it applies the most rapid and high qualified algorithms. It is best suited for low hardware Specification.	https://github.com/sagi-z/BackgroundSubtractorCNT

8   Performance Evaluation and its Metrics

Performance evaluation is classified into qualitative and quantitative evaluations. In qualitative, the visual results from interested dataset are compared by different background extraction algorithms, for example we evaluate the results of algorithms in respect to some challenges like (shadow, camera jitter and dynamic background, etc.) and then decide the best performance visually. The qualitative evaluation gives a better perspective about the evaluation and shed the light on the subjective elements.

On the other hand, the quantitative evaluation is done by using statistical measures, we compare the results of the performance metrics from background extraction algorithms over a wide range of dataset, for example, we compare the algorithms F-measures results in respect to some challenges of the dataset [7].

In qualitative evaluation, many performance metrics are applied to assess the robustness of the background extraction algorithms to different challenges. True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), Positive samples (P), Negative samples (N), Accuracy (recognition rate), F-measures(F1), sensitivity (recall), precision, Fβ, specificity, and error rate (misclassification rate) are used as evaluation measures.

The F-measure, sensitivity (recall) and precision depend on the accuracy (recognition rate) to detect the foreground and background pixels. This is done by calculating the True Positive (TP) number which is the number of foreground pixels that is categorized as a foreground. While False Positive (FP) is for background pixels which are categorized as a foreground. True Negative (TN) is for background pixels that are categorized as a background. Yet, False Negative (FN) is for foreground pixels that are categorized as a background.

Recall is also used to calculate the truly categorized foreground pixels with respect to the number of all pixels categorized as a foreground. Precision on the other hand is measuring the truly categorized foreground pixels as a foreground to the total number of foreground pixels in the ground truth. Normally we measure the relevance by recall and precision.

A low recall is an indication of over segmentation of the foreground objects, where a low precision is an indication of under segmentation of the foreground objects. High F-measures is an indication of a robust background extraction algorithm [7][197]. The set performance metrics are expressed in table 5.

Table 5. The set performance metrics

Performance Metrics	Formula
F-Measures	$\dfrac{2 * Precision * Sensitivity}{Precision + Sensitivity}$
Sensitivity(recall)	$\dfrac{TP}{P}$
Precision	$\dfrac{TP}{TP + FP}$
Fβ, β is non-negative Real number	$\dfrac{(1 + \beta)^2 * Precision * Senstivity}{\beta^2 * Precision + Sensitivity}$
Accuracy	$\dfrac{TP + TN}{P + N}$
Specificity	$\dfrac{TN}{N}$
Error rate	$\dfrac{FP + FN}{P + N}$

9   Conclusion

In this article, we reviewed the background extraction models used to detect the moving foreground target in a video taken by a surveillance stationary camera. The review started with a brief introduction to the process stages of background extraction models (background initialization, background modeling, background maintenance, and foreground detection).

We also reviewed the real-time applications in which background extraction models are applied, and we came across various environments with different challenges. The importance of choosing the suitable background extraction model for a unique environment or a challenge is highlighted, as in real-time applications, it is required to detect the foreground mask accurately and quickly.

Moreover, in real-time applications, we concluded that the conventional and some statistical models for the background extraction models like (mean, median, GMM, and histogram) are still dominant in the foreground detection field, and this is due to the fact that these approaches are requiring less time, less memory and less computation consumption. Conventional models are also well-known and well-experimented approaches that make them easier to adapt compared to state-of-the-art models. On the other hand, state-of-the-art models like deep learning and CNN models achieve good results in terms of accuracy and solving challenges. Therefore, there is no single background extraction model that can handle all the challenges in different environments, so there are numerous works that fused more than one approach to overcome the challenges and develop the performance.

Furthermore, we presented helpful tables of datasets and libraries used in implementing background extraction models. Eventually, we elaborated on the performance evaluation (qualitative and quantitative), so as the set performance metrics used to evaluate the background extraction models.

10   References

[1] N. K. Jain, R. K. Saini, and P. Mittal, "A review on traffic monitoring system techniques," in Soft Computing: Theories and Applications, Springer, 2019, pp. 569–577. https://doi.org/10.1007/978-981-13-0589-4_53

[2] D. Gutchess, M. Trajkovics, E. Cohen-Solal, D. Lyons, and A. K. Jain, "A background model initialization algorithm for video surveillance," in Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, 2001, vol. 1, pp. 733–740. https://doi.org/10.1109/iccv.2001.937598

[3] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," ACM Comput. Surv., vol. 41, no. 3, pp. 1–58, 2009. https://doi.org/10.1145/1541880.1541882

[4] Y. Pritch, A. Rav-Acha, and S. Peleg, "Nonchronological video synopsis and indexing," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 11, pp. 1971–1984, 2008. https://doi.org/10.1109/tpami.2008.29

[5] C. Lin, B. Yan, and W. Tan, "Foreground detection in surveillance video with fully convolutional semantic network," in 2018 25th IEEE International Conference on Image Processing (ICIP), 2018, pp. 4118–4122. https://doi.org/10.1109/icip.2018.8451816

[6] K. Goyal and J. Singhai, "Review of background subtraction methods using Gaussian mixture model for video surveillance systems," Artif. Intell. Rev., vol. 50, no. 2, pp. 241–259, 2018. https://doi.org/10.1007/s10462-017-9542-x

[7] D. Zeng, M. Zhu, and A. Kuijper, "Combining background subtraction algorithms with convolutional neural network," J. Electron. Imaging, vol. 28, no. 1, p. 13011, 2019. https://doi.org/10.1117/1.jei.28.1.013011

[8] R. A. Hadi, L. E. George, and M. J. Mohammed, "A computationally economic novel approach for real-time moving multi-vehicle detection and tracking toward efficient traffic surveillance," Arab. J. Sci. Eng., vol. 42, no. 2, pp. 817–831, 2017. https://doi.org/10.1007/s13369-016-2351-8

[9] T. Bouwmans, "Recent advanced statistical background modeling for foreground detection-a systematic survey," Recent Patents Comput. Sci., vol. 4, no. 3, pp. 147–176, 2011. https://doi.org/10.2174/2213275911104030147

[10] S. Lee and D. Kim, "Background subtraction using the factored 3-way restricted Boltzmann machines," arXiv Prepr. arXiv1802.01522, 2018.

[11] P. W. Power and J. A. Schoonees, "Understanding background mixture models for foreground segmentation," in Proceedings image and vision computing New Zealand, 2002, vol. 2002, pp. 10–11.

[12] N. J. B. McFarlane and C. P. Schofield, "Segmentation and tracking of piglets in images," Mach. Vis. Appl., vol. 8, no. 3, pp. 187–193, 1995.

[13] M. Hadiuzzaman, N. Haque, F. Rahman, S. Hossain, M. R. K. Siam, and T. Z. Qiu, "Pixel-based heterogeneous traffic measurement considering shadow and illumination variation," Signal, Image Video Process., vol. 11, no. 7, pp. 1245–1252, 2017. https://doi.org/10.1007/s11760-017-1081-z

[14] G. Monteiro, J. Marcos, M. Ribeiro, and J. Batista, "Robust segmentation for outdoor traffic surveillance," in 2008 15th IEEE International Conference on Image Processing, 2008, pp. 2652–2655. https://doi.org/10.1109/icip.2008.4712339

[15] G. Monteiro, J. Marcos, M. Ribeiro, and J. Batista, "Robust segmentation process to detect incidents on highways," in International Conference Image Analysis and Recognition, 2008, pp. 110–121. https://doi.org/10.1007/978-3-540-69812-8_11

[16] P. Blauensteiner and M. Kampel, Visual surveillance of an airport's apron-An overview of the AVITRACK project. na, 2004.

[17] Z. Zivkovic and F. Van Der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," Pattern Recognit. Lett., vol. 27, no. 7, pp. 773–780, 2006. https://doi.org/10.1016/j.patrec.2005.11.005

[18] J. Zheng, Y. Wang, N. L. Nihan, and M. E. Hallenbeck, "Extracting roadway background image: Mode-based approach," Transp. Res. Rec., vol. 1944, no. 1, pp. 82–88, 2006. https://doi.org/10.1177/0361198106194400111

[19] O. Munteanu, T. Bouwmans, E. Zahzah, and R. Vasiu, "The detection of moving objects in video by background subtraction using Dempster-Shafer theory," Trans. Electron. Commun., vol. 60, no. 1, 2015.

[20] F. El Baf, T. Bouwmans, and B. Vachon, "Type-2 fuzzy mixture of Gaussians model: application to background modeling," in International Symposium on Visual Computing, 2008, pp. 772–781. https://doi.org/10.1007/978-3-540-89639-5_74

[21] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149), 1999, vol. 2, pp. 246–252. https://doi.org/10.1109/cvpr.1999.784637

[22] J. D. Pulgarin-Giraldo, A. Alvarez-Meza, D. Insuasti-Ceballos, T. Bouwmans, and G. Castellanos-Dominguez, "GMM background modeling using divergence-based weight updating," in Iberoamerican Congress on Pattern Recognition, 2016, pp. 282–290. https://doi.org/10.1007/978-3-319-52277-7_35

[23] T. Bouwmans, F. El Baf, and B. Vachon, "Statistical background modeling for foreground detection: A survey," in Handbook of pattern recognition and computer vision, World Scientific, 2010, pp. 181–199. https://doi.org/10.1142/9789814273398_0008

[24] B. White and M. Shah, "Automatically tuning background subtraction parameters using particle swarm optimization," in 2007 IEEE International Conference on Multimedia and Expo, 2007, pp. 1826–1829. https://doi.org/10.1109/icme.2007.4285028

[25] I. Martins, P. Carvalho, L. Corte-Real, and J. L. Alba-Castro, "BMOG: boosted Gaussian mixture model with controlled complexity," in Iberian Conference on Pattern Recognition and Image Analysis, 2017, pp. 50–57. https://doi.org/10.1007/978-3-319-58838-4_6

[26] X. Lu, C. Xu, L. Wang, and L. Teng, "Improved background subtraction method for detecting moving objects based on GMM," IEEJ Trans. Electr. Electron. Eng., vol. 13, no. 11, pp. 1540–1550, 2018. https://doi.org/10.1002/tee.22718

[27] S. T. Ali, K. Goyal, and J. Singhai, "Moving object detection using self-adaptive Gaussian Mixture Model for real time applications," in 2017 International Conference on Recent Innovations in Signal processing and Embedded Systems (RISE), 2017, pp. 153–156. https://doi.org/10.1109/rise.2017.8378144

[28] O. Barnich and M. Van Droogenbroeck, "ViBe: a powerful random technique to estimate the background in video sequences," in 2009 IEEE international conference on acoustics, speech and signal processing, 2009, pp. 945–948. https://doi.org/10.1109/icassp.2009.4959741

[29] L. F. Huang, Q. Y. Chen, J. F. Lin, and H. Z. Lin, "Block background subtraction method based on ViBe," in Applied Mechanics and Materials, 2014, vol. 556, pp. 3549–3552. https://doi.org/10.4028/www.scientific.net/amm.556-562.3549

[30] X. Zhou, X. Liu, A. Jiang, B. Yan, and C. Yang, "Improving video segmentation by fusing depth cues and the visual background extractor (ViBe) algorithm," Sensors, vol. 17, no. 5, p. 1177, 2017. https://doi.org/10.3390/s17051177

[31] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," IEEE Trans. Image Process., vol. 20, no. 6, pp. 1709–1724, 2010. https://doi.org/10.1109/tip.2010.2101613

[32] P.-L. St-Charles, G.-A. Bilodeau, and R. Bergevin, "Subsense: A universal change detection method with local adaptive sensitivity," IEEE Trans. Image Process., vol. 24, no. 1, pp. 359–373, 2014. https://doi.org/10.1109/tip.2014.2378053

[33] Y. Wang, P.-M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, "CDnet 2014: An expanded change detection benchmark dataset," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2014, pp. 387–394. https://doi.org/10.1109/cvprw.2014.126

[34] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," in European conference on computer vision, 2000, pp. 751–767. https://doi.org/10.1007/3-540-45053-x_48

[35] M. Hua, Y. Li, and Y. Luo, "Robust Background Modeling with Kernel Density Estimation.," Int. J. Online Eng., vol. 11, no. 8, 2015.

[36] M. Shah, J. D. Deng, and B. J. Woodford, "Video background modeling: recent approaches, issues and our proposed techniques," Mach. Vis. Appl., vol. 25, no. 5, pp. 1105–1119, 2014. https://doi.org/10.1007/s00138-013-0552-7

[37] S. Javed, S. H. Oh, and S. K. Jung, "An Improved Pixel-Based Adaptive Background Segmenter for Visual Surveillance System," pp. 579–582, 2014.

[38] M. Hofmann, P. Tiefenbacher, and G. Rigoll, "Background segmentation with feedback: The pixel-based adaptive segmenter," in 2012 IEEE computer society conference on computer vision and pattern recognition workshops, 2012, pp. 38–43. https://doi.org/10.1109/cvprw.2012.6238925

[39] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603–619, 2002. https://doi.org/10.1109/34.1000236

[40] F. El Baf, T. Bouwmans, and B. Vachon, "Fuzzy statistical modeling of dynamic backgrounds for moving object detection in infrared videos," in 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2009, pp. 60–65. https://doi.org/10.1109/cvprw.2009.5204109

[41] T. Bouwmans and B. Garcia-Garcia, "Background subtraction in real applications: Challenges, current models and future directions," arXiv Prepr. arXiv1901.03577, 2020. https://doi.org/10.1016/j.cosrev.2019.100204

[42] S. Messelodi, C. M. Modena, N. Segata, and M. Zanin, "A kalman filter based background updating algorithm robust to sharp illumination changes," in International Conference on Image Analysis and Processing, 2005, pp. 163–170. https://doi.org/10.1007/11553595_20

[43] F. Lei and X. Zhao, "Adaptive background estimation of underwater using Kalman-Filtering," in 2010 3rd International Congress on Image and Signal Processing, 2010, vol. 1, pp. 64–67. https://doi.org/10.1109/cisp.2010.5647080

[44] R. Chang, T. Gandhi, and M. M. Trivedi, "Vision modules for a multi-sensory bridge monitoring approach," in Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No. 04TH8749), 2004, pp. 971–976. https://doi.org/10.1109/itsc.2004.1399038

[45] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: Principles and practice of background maintenance," in Proceedings of the seventh IEEE international conference on computer vision, 1999, vol. 1, pp. 255–261. https://doi.org/10.1109/iccv.1999.791228

[46] E. Zhang, F. Chen, and W. Zhang, "A novel particle filter based background subtraction method," in 2006 International Conference on Computational Intelligence and Security, 2006, vol. 2, pp. 1837–1840. https://doi.org/10.1109/iccias.2006.295382

[47] G. T. Cinar and J. C. Pr\'\incipe, "Adaptive background estimation using an information theoretic cost for hidden state estimation," in the 2011 International Joint Conference on Neural Networks, 2011, pp. 489–494. https://doi.org/10.1109/ijcnn.2011.6033261

[48] X. Yu, X. Chen, and H. Zhang, "Accurate motion detection in dynamic scenes based on ego-motion estimation and optical flow segmentation combined method," in 2011 Symposium on Photonics and Optoelectronics (SOPO), 2011, pp. 1–4. https://doi.org/10.1109/sopo.2011.5780637

[49] D. Butler, S. Sridharan, and V. M. J. Bove, "Real-time adaptive background segmentation," in 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings (ICASSP'03)., 2003, vol. 3, pp. III--349. https://doi.org/10.1109/icassp.2003.1199481

[50] M. Xiao, C. Han, and X. Kang, "A background reconstruction for dynamic scenes," in 2006 9th International Conference on Information Fusion, 2006, pp. 1–7. https://doi.org/10.1109/icif.2006.301727

[51] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis, "Real-time foreground--background segmentation using codebook model," Real-time imaging, vol. 11, no. 3, pp. 172–185, 2005. https://doi.org/10.1016/j.rti.2004.12.004

[52] L. Wang and C. Pan, "Effective multi-resolution background subtraction," in 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011, pp. 909–912. https://doi.org/10.1109/icassp.2011.5946552

[53] W. Su, H. Zhuang, and X. Qiu, "Moving targets detection and tracking based on improved codebook algorithm and Kalman filtering," in 2017 36th Chinese Control Conference (CCC), 2017, pp. 11494–11498. https://doi.org/10.23919/chicc.2017.8029193

[54] D. K. Panda and S. Meher, "A new Wronskian change detection model based codebook background subtraction for visual surveillance applications," J. Vis. Commun. Image Represent., vol. 56, pp. 52–72, 2018. https://doi.org/10.1016/j.jvcir.2018.07.014

[55] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis, "Background modeling and subtraction by codebook construction," in 2004 International Conference on Image Processing, 2004. ICIP'04., 2004, vol. 5, pp. 3061–3064. https://doi.org/10.1109/icip.2004.1421759

[56] A. Ilyas, M. Scuturici, and S. Miguet, "Real time foreground-background segmentation using a modified codebook model," in 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, 2009, pp. 454–459. https://doi.org/10.1109/avss.2009.85

[57] N. M. Oliver, B. Rosario, and A. P. Pentland, "A Bayesian computer vision system for modeling human interactions," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 831–843, 2000. https://doi.org/10.1109/34.868684

[58] Y. Dong, T. X. Han, and G. N. DeSouza, "Illumination invariant foreground detection using multi-subspace learning," Int. J. knowledge-based Intell. Eng. Syst., vol. 14, no. 1, pp. 31–41, 2010. https://doi.org/10.3233/kes-2010-0188

[59] Y. Kawanishi, I. Mitsugami, M. Mukunoki, and M. Minoh, "Background image generation by preserving lighting condition of outdoor scenes," Procedia-Social Behav. Sci., vol. 2, no. 1, pp. 129–136, 2010. https://doi.org/10.1016/j.sbspro.2010.01.026

[60] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" J. ACM, vol. 58, no. 3, pp. 1–37, 2011.

[61] C. Marghes and T. Bouwman, "Background modeling via incremental maximum margin criterion," in Asian Conference on Computer Vision, 2010, pp. 394–403. https://doi.org/10.1007/978-3-642-22819-3_40

[62] D. Farcas, C. Marghes, and T. Bouwmans, "Background subtraction via incremental maximum margin criterion: a discriminative subspace approach," Mach. Vis. Appl., vol. 23, no. 6, pp. 1083–1101, 2012. https://doi.org/10.1007/s00138-012-0421-9

[63] C. Marghes, T. Bouwmans, and R. Vasiu, "Background modeling and foreground detection via a reconstructive and discriminative subspace learning approach," in International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV 2012, 2012. https://doi.org/10.1007/s00138-012-0421-9

[64] N. Vaswani, T. Bouwmans, S. Javed, and P. Narayanamurthy, "Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery," IEEE Signal Process. Mag., vol. 35, no. 4, pp. 32–55, 2018. https://doi.org/10.1109/msp.2018.2826566

[65] H.-H. Lin, T.-L. Liu, and J.-H. Chuang, "A probabilistic SVM approach for background scene initialization," in Proceedings. International Conference on Image Processing, 2002, vol. 3, pp. 893–896. https://doi.org/10.1109/icip.2002.1039116

[66] I. Junejo, A. Bhutta, and H. Foroosh, "Dynamic scene modeling for object detection using single-class SVM," in Proc. of IEEE International Conference on Image Processing (ICIP), 2010, vol. 1, pp. 1541–1544. https://doi.org/10.1007/s11760-011-0230-z

[67] J. Wang, X. Xiang, and W. Wei, "Existence of periodic solutions for integrodifferential impulsive periodic system on Banach space," in Abstract and Applied Analysis, 2008, vol. 2008. https://doi.org/10.1155/2008/939062

[68] A. Tavakkoli, M. Nicolescu, and G. Bebis, "A novelty detection approach for foreground region detection in videos with quasi-stationary backgrounds," in International symposium on visual computing, 2006, pp. 40–49. https://doi.org/10.1007/11919476_5

[69] A. Tavakkoli, A. Ambardekar, M. Nicolescu, and S. Louis, "A genetic approach to training support vector data descriptors for background modeling in video data," in International Symposium on Visual Computing, 2007, pp. 318–327. https://doi.org/10.1007/978-3-540-76856-2_31

[70] S. Javed, S. K. Jung, A. Mahmood, and T. Bouwmans, "Motion-aware graph regularized RPCA for background modeling of complex scenes," in 2016 23rd International Conference on Pattern Recognition (ICPR), 2016, pp. 120–125. https://doi.org/10.1109/icpr.2016.7899619

[71] S. Javed, A. Mahmood, T. Bouwmans, and S. K. Jung, "Spatiotemporal low-rank modeling for complex scene background initialization," IEEE Trans. Circuits Syst. Video Technol., vol. 28, no. 6, pp. 1315–1329, 2016. https://doi.org/10.1109/tcsvt.2016.2632302

[72] S. Javed, S. Ho Oh, A. Sobral, T. Bouwmans, and S. Ki Jung, "Background subtraction via superpixel-based online matrix decomposition with structured foreground constraints," in Proceedings of the IEEE International Conference on Computer Vision Workshops, 2015, pp. 90–98. https://doi.org/10.1109/iccvw.2015.123

[73] S. Javed, T. Bouwmans, and S. K. Jung, "Combining arf and or-pca for robust background subtraction of noisy videos," in International Conference on Image Analysis and Processing, 2015, pp. 340–351. https://doi.org/10.1007/978-3-319-23234-8_32

[74] A. Sobral, C. G. Baker, T. Bouwmans, and E. Zahzah, "Incremental and multi-feature tensor subspace learning applied for background modeling and subtraction," in International Conference Image Analysis and Recognition, 2014, pp. 94–103. https://doi.org/10.1007/978-3-319-11758-4_11

[75] S. Javed, T. Bouwmans, and S. K. Jung, "Stochastic decomposition into low rank and sparse tensor for robust background subtraction," 2015. https://doi.org/10.1049/ic.2015.0105

[76] A. Sobral, S. Javed, S. Ki Jung, T. Bouwmans, and E. Zahzah, "Online stochastic tensor decomposition for background subtraction in multispectral video sequences," in Proceedings

of the IEEE International Conference on Computer Vision Workshops, 2015, pp. 106–113. https://doi.org/10.1109/iccvw.2015.125

[77] M. Braham and M. Van Droogenbroeck, "Deep background subtraction with scene-specific convolutional neural networks," in 2016 international conference on systems, signals and image processing (IWSSIP), 2016, pp. 1–4. https://doi.org/10.1109/iwssip.2016.7502717

[78] Y. Wang, Z. Luo, and P.-M. Jodoin, "Interactive deep learning method for segmenting moving objects," Pattern Recognit. Lett., vol. 96, pp. 66–75, 2017. https://doi.org/10.1016/j.patrec.2016.09.014

[79] K. Lim, W.-D. Jang, and C.-S. Kim, "Background subtraction using encoder-decoder structured convolutional neural network," in 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2017, pp. 1–6. https://doi.org/10.1109/avss.2017.8078547

[80] X. Li, M. Ye, Y. Liu, and C. Zhu, "Adaptive deep convolutional neural networks for scene-specific object detection," IEEE Trans. Circuits Syst. Video Technol., 2017. https://doi.org/10.1109/tcsvt.2017.2749620

[81] T. Minematsu, A. Shimada, and R. Taniguchi, "Analytics of deep neural network in change detection," in 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2017, pp. 1–6. https://doi.org/10.1109/avss.2017.8078550

[82] T. Minematsu, A. Shimada, H. Uchiyama, and R. Taniguchi, "Analytics of deep neural network-based background subtraction," J. Imaging, vol. 4, no. 6, p. 78, 2018. https://doi.org/10.3390/jimaging4060078

[83] L. Maddalena and A. Petrosino, "The 3dSOBS+ algorithm for moving object detection," Comput. Vis. Image Underst., vol. 122, pp. 65–73, 2014. https://doi.org/10.1016/j.cviu.2013.11.006

[84] L. Maddalena and A. Petrosino, "Self-organizing background subtraction using color and depth data," Multimed. Tools Appl., vol. 78, no. 9, pp. 11927–11948, 2019. https://doi.org/10.1007/s11042-018-6741-7

[85] G. D. Sergio, V. P. Javier, and others, "Simplified SOM-neural model for video segmentation of moving objects," in 2009 International Joint Conference on Neural Networks, 2009, pp. 474–480. https://doi.org/10.1109/ijcnn.2009.5178632

[86] M. I. Chacon-Murguia and S. Gonzalez-Duarte, "An adaptive neural-fuzzy approach for object detection in dynamic backgrounds for surveillance systems," IEEE Trans. Ind. Electron., vol. 59, no. 8, pp. 3286–3298, 2011. https://doi.org/10.1109/tie.2011.2106093

[87] M. I. Chacon-Murgu\'\ia, G. Ramirez-Alonso, and S. Gonzalez-Duarte, "Improvement of a neural-fuzzy motion detection vision model for complex scenario conditions," in the 2013 International Joint Conference on Neural Networks (IJCNN), 2013, pp. 1–8. https://doi.org/10.1109/ijcnn.2013.6706734

[88] P. Xu, M. Ye, Q. Liu, X. Li, L. Pei, and J. Ding, "Motion detection via a couple of auto-encoder networks," in 2014 IEEE International Conference on Multimedia and Expo (ICME), 2014, pp. 1–6. https://doi.org/10.1109/icme.2014.6890140

[89] T. Bouwmans, S. Javed, M. Sultana, and S. K. Jung, "Deep neural network concepts for background subtraction: A systematic review and comparative evaluation," Neural Networks, 2019. https://doi.org/10.1016/j.neunet.2019.04.024

[90] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105. https://doi.org/10.1145/3065386

[91] P. Xu, M. Ye, X. Li, Q. Liu, Y. Yang, and J. Ding, "Dynamic background learning through deep auto-encoder networks," in Proceedings of the 22nd ACM international conference on Multimedia, 2014, pp. 107–116. https://doi.org/10.1145/2647868.2654914

[92] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in 2009 IEEE conference on computer vision and pattern recognition, 2009, pp. 248–255. https://doi.org/10.1109/cvprw.2009.5206848

[93] W. Qiu and A. Yuille, "Unrealcv: Connecting computer vision to unreal engine," in European Conference on Computer Vision, 2016, pp. 909–916. https://doi.org/10.1007/978-3-319-49409-8_75

[94] T. Bouwmans, "Subspace learning for background modeling: A survey," Recent Patents Comput. Sci., vol. 2, no. 3, pp. 223–234, 2009. https://doi.org/10.2174/2213275910902030223

[95] J. Du, M. G. Meng, and C. Yang, "Design of Wireless Security Surveillance System in Remote Unattended Wells," Int. J. Online Biomed. Eng., vol. 15, no. 07, pp. 4–15, 2019. https://doi.org/10.3991/ijoe.v15i07.7837

[96] M. Chowdhury, J. Gao, and R. Islam, "Human surveillance system for security application," in International Conference on Security and Privacy in Communication Systems, 2015, pp. 711–724.

[97] J. Batista, P. Peixoto, C. Fernandes, and M. Ribeiro, "A dual-stage robust vehicle detection and tracking for real-time traffic monitoring," in 2006 IEEE Intelligent Transportation Systems Conference, 2006, pp. 528–535. https://doi.org/10.1109/itsc.2006.1706795

[98] G. Monteiro, J. Marcos, and J. Batista, "Stopped vehicle detection system for outdoor traffic surveillance," RECPAD 2008, 2008.

[99] G. Monteiro, "Traffic video surveillance for automatic incident detection on highways," Master Thesis, 2009.

[100] Á. Virginás-Tar, M. Baba, V. Gui, D. Pescaru, and I. Jian, "Vehicle counting and classification for traffic surveillance using wireless video sensor networks," in 2014 22nd Telecommunications Forum Telfor (TELFOR), 2014, pp. 1019–1022. https://doi.org/10.1109/telfor.2014.7034580

[101] C. G. del Postigo, J. Torres, and J. M. Menéndez, "Vacant parking area estimation through background subtraction and transience map analysis," IET Intell. Transp. Syst., vol. 9, no. 9, pp. 835–841, 2015. https://doi.org/10.1049/iet-its.2014.0090

[102] M. Neuhausen, "Video-based parking space detection: Localisation of vehicles and development of an infrastructure for a routeing system," Inst. fr Neuroinformatik, 2015.

[103] J. Cho et al., "Automatic parking system using background subtraction with CCTV environment international conference on control, automation and systems (ICCAS 2016)," in 2016 16th International Conference on Control, Automation and Systems (ICCAS), 2016, pp. 1649–1652. https://doi.org/10.1109/iccas.2016.7832520

[104] R. Lin, X. Cao, Y. Xu, C. Wu, and H. Qiao, "Airborne moving vehicle detection for video surveillance of urban traffic," in 2009 IEEE Intelligent Vehicles Symposium, 2009, pp. 203–208. https://doi.org/10.1109/ivs.2009.5164278

[105] S. Muniruzzaman et al., "Deterministic algorithm for traffic detection in free-flow and congestion using video sensor," J. Built Environ. Technol. Eng., vol. 1, pp. 111–130, 2016.

[106] A. K. S. Kushwaha and R. Srivastava, "A framework of moving object segmentation in maritime surveillance inside a dynamic background," in Transactions on Computational Science XXV, Springer, 2015, pp. 35–54. https://doi.org/10.1007/978-3-662-47074-9_3

[107] M. A. M. Razif, M. Mokji, and M. Mun'im Ahmad Zabidi, "Low complexity maritime surveillance video using background subtraction on h. 264," in 2015 International Symposium on Technology Management and Emerging Technologies (ISTMET), 2015, pp. 364–368. https://doi.org/10.1109/istmet.2015.7359060

[108] J. Zhao, S. Pang, B. Hartill, and A. Sarrafzadeh, "Adaptive background modeling for land and water composition scenes," in International Conference on Image Analysis and Processing, 2015, pp. 97–107. https://doi.org/10.1007/978-3-319-23234-8_10

[109] V. F. Arguedas, G. Pallotta, and M. Vespe, "Maritime traffic networks: From historical positioning data to unsupervised maritime traffic monitoring," IEEE Trans. Intell. Transp. Syst., vol. 19, no. 3, pp. 722–732, 2017. https://doi.org/10.1109/tits.2017.2699635

[110] Z. Liu, F. Zhou, X. Chen, X. Bai, and C. Sun, "Iterative infrared ship target segmentation based on multiple features," Pattern Recognit., vol. 47, no. 9, pp. 2839–2852, 2014. https://doi.org/10.1016/j.patcog.2014.03.005

[111] A. Leykin and M. Tuceryan, "Detecting shopper groups in video sequences," in 2007 IEEE Conference on Advanced Video and Signal Based Surveillance, 2007, pp. 417–422. https://doi.org/10.1109/avss.2007.4425347

[112] A. Leykin and M. Tuceryan, "A vision system for automated customer tracking for marketing analysis: Low level feature extraction," in Human Activity Recognition and Modelling Workshop, 2005, vol. 3, p. 6.

[113] A. Leykin and M. Tuceryan, "Tracking and activity analysis in retail environments," 2005.

[114] N. Avinash, M. S. S. Kumar, and S. M. Sagar, "Automated video surveillance for retail store statistics generation," in Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012), 2013, pp. 585–596. https://doi.org/10.1007/978-81-322-0997-3_52

[115] V. Nogueira, H. Oliveira, J. A. Silva, T. Vieira, and K. Oliveira, "RetailNet: A deep learning approach for people counting and hot spots detection in retail stores," in 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2019, pp. 155–162. https://doi.org/10.1109/sibgrapi.2019.00029

[116] J. Zhong and others, "Segmenting foreground objects from a dynamic textured background via a robust kalman filter," in Proceedings Ninth IEEE International Conference on Computer Vision, 2003, pp. 44–50. https://doi.org/10.1109/iccv.2003.1238312

[117] X. Jin, P. Niu, and L. Liu, "A GMM-Based Segmentation Method for the Detection of Water Surface Floats," IEEE Access, vol. 7, pp. 119018–119025, 2019. https://doi.org/10.1109/access.2019.2937129

[118] T. Ko, S. Soatto, and D. Estrin, "Background subtraction on distributions," in European Conference on Computer Vision, 2008, pp. 276–289.

[119] T. Ko, S. Soatto, and D. Estrin, "Warping background subtraction," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010, pp. 1331–1338. https://doi.org/10.1109/cvpr.2010.5539813

[120] P. Dickinson, R. Freeman, S. Patrick, S. Lawson, and others, "Autonomous monitoring of cliff nesting seabirds using computer vision," 2008.

[121] P. Dickinson, C. Qing, S. Lawson, R. Freeman, and others, "Automated visual monitoring of nesting seabirds," in Workshop on visual observation and analysis of animal and insect behaviour, Istanbul, 2010.

[122] C. Spampinato, Y.-H. Chen-Burger, G. Nadarajan, and R. B. Fisher, "Detecting, Tracking and Counting Fish in Low Quality Unconstrained Underwater Videos.," VISAPP (2), vol. 2008, no. 514–519, p. 1, 2008. https://doi.org/10.5220/0001077705140519

[123] C. Spampinato, S. Palazzo, and I. Kavasidis, "A texton-based kernel density estimation approach for background modeling under extreme conditions," Comput. Vis. Image Underst., vol. 122, pp. 74–83, 2014. https://doi.org/10.1016/j.cviu.2013.12.003

[124] C. Spampinato, D. Giordano, R. Di Salvo, Y.-H. J. Chen-Burger, R. B. Fisher, and G. Nadarajan, "Automatic fish classification for underwater species behavior understanding," in Proceedings of the first ACM international workshop on Analysis and retrieval of tracked

events and motion in imagery streams, 2010, pp. 45–50. https://doi.org/10.1145/1877868.1877881

[125] C. Spampinato et al., "Understanding fish behavior during typhoon events in real-life underwater environments," Multimed. Tools Appl., vol. 70, no. 1, pp. 199–236, 2012. https://doi.org/10.1007/s11042-012-1101-5

[126] U. Knauer^1, M. Himmelsbach^1, F. Winkler^1, F. Zautke, K. Bienefeld, and B. Meffert^1, "Application of an adaptive background model for monitoring honeybees," 2005.

[127] B. Rezaei and S. Ostadabbas, "Moving object detection through robust matrix completion augmented with objectness," IEEE J. Sel. Top. Signal Process., vol. 12, no. 6, pp. 1313–1323, 2018. https://doi.org/10.1109/jstsp.2018.2869111

[128] M. Ashraphijuo, V. Aggarwal, and X. Wang, "On deterministic sampling patterns for robust low-rank matrix completion," IEEE Signal Process. Lett., vol. 25, no. 3, pp. 343–347, 2017. https://doi.org/10.1109/lsp.2017.2780983

[129] C. J. Cohen, D. Haanpaa, and J. P. Zott, "Machine vision algorithms for robust animal species identification," in 2015 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), 2015, pp. 1–7. https://doi.org/10.1109/aipr.2015.7444526

[130] C. J. Cohen, D. Haanpaa, S. Rowe, and J. P. Zott, "Vision algorithms for automated census of animals," in 2011 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), 2011, pp. 1–5. https://doi.org/10.1109/aipr.2011.6176371

[131] L. Maddalena and A. Petrosino, "Background subtraction for moving object detection in RGBD data: A survey," J. Imaging, vol. 4, no. 5, p. 71, 2018. https://doi.org/10.3390/jimaging4050071

[132] G. Levin, "Computer vision for artists and designers: pedagogic tools and techniques for novice programmers," AI Soc., vol. 20, no. 4, pp. 462–482, 2006. https://doi.org/10.1007/s00146-006-0049-2

[133] T. Ganokratanaa and S. Pumrin, "Hand gesture recognition algorithm for smart cities based on wireless sensor," Int. J. Online Biomed. Eng., vol. 13, no. 06, pp. 58–75, 2017. https://doi.org/10.3991/ijoe.v13i06.7022

[134] S. Paul and A. K. Pal, "An Efficient Surveillance Video Coding Scheme for Static Camera Based Captured Video Data," in 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), 2019, pp. 868–873. https://doi.org/10.1109/spin.2019.8711754

[135] H. Huang, X. Fang, Y. Ye, S. Zhang, and P. L. Rosin, "Practical automatic background substitution for live video," Comput. Vis. Media, vol. 3, no. 3, pp. 273–284, 2017. https://doi.org/10.1109/spin.2019.8711754

[136] G. Tzanidou, "Carried baggage detection and recognition in video surveillance with foreground segmentation," Loughborough University, 2014.

[137] B. U. Töreyin, Y. Dedeo\uglu, U. Güdükbay, and A. E. Cetin, "Computer vision based method for real-time fire and flame detection," Pattern Recognit. Lett., vol. 27, no. 1, pp. 49–58, 2006. https://doi.org/10.1016/j.patrec.2005.06.015

[138] W. Zhiliang, G. Jian, C. Xin, and Z. Weiming, "Organic light emitting diode pixel defect detection through patterned background removal," Sens. Lett., vol. 11, no. 2, pp. 356–361, 2013. https://doi.org/10.1166/sl.2013.2754

[139] A. Laurentini, "The visual hull concept for silhouette-based image understanding," IEEE Trans. Pattern Anal. Mach. Intell., vol. 16, no. 2, pp. 150–162, 1994. https://doi.org/10.1109/34.273735

[140] S. Molina-Giraldo et al., "People detection in video streams using background subtraction and spatial-based scene modeling," in 2015 20th Symposium on Signal Processing, Images

and Computer Vision (STSIVA), 2015, pp. 1–6. https://doi.org/10.1109/stsiva.2015.7330394

[141] L. Unzueta, M. Nieto, A. Cortés, J. Barandiaran, O. Otaegui, and P. Sánchez, "Adaptive multicue background subtraction for robust vehicle counting and classification," IEEE Trans. Intell. Transp. Syst., vol. 13, no. 2, pp. 527–540, 2011. https://doi.org/10.1109/tits.2011.2174358

[142] T.-H. Tran and T.-L. Le, "Vision based boat detection for maritime surveillance," in 2016 International Conference on Electronics, Information, and Communications (ICEIC), 2016, pp. 1–4. https://doi.org/10.1109/elinfocom.2016.7563033

[143] K. Higa and K. Iwamoto, "Robust estimation of product amount on store shelves from a surveillance camera for improving on-shelf availability," in 2018 IEEE International Conference on Imaging Systems and Techniques (IST), 2018, pp. 1–6. https://doi.org/10.1109/ist.2018.8577157

[144] Q. Ling, J. Yan, F. Li, and Y. Zhang, "A background modeling and foreground segmentation approach based on the feedback of moving objects in traffic surveillance systems," Neurocomputing, vol. 133, pp. 32–45, 2014. https://doi.org/10.1016/j.neucom.2013.11.034

[145] J. T. Lee, M. S. Ryoo, M. Riley, and J. K. Aggarwal, "Real-time detection of illegally parked vehicles using 1-D transformation," in 2007 IEEE Conference on Advanced Video and Signal Based Surveillance, 2007, pp. 254–259. https://doi.org/10.1109/avss.2007.4425319

[146] M. Sarker, K. Mostafa, C. Weihua, and M. K. Song, "Detection and recognition of illegally parked vehicles based on an adaptive gaussian mixture model and a seed fill algorithm," J. Inf. Commun. Converg. Eng., vol. 13, no. 3, pp. 197–204, 2015. https://doi.org/10.6109/jicce.2015.13.3.197

[147] M. Chunyang, M. Xing, and Z. Panpan, "Smart detection of vehicle in illegal parking area by fusing of multi-features," in 2015 9th International Conference on Next Generation Mobile Applications, Services and Technologies, 2015, pp. 388–392. https://doi.org/10.1109/ngmast.2015.28

[148] A. Filonenko, K.-H. Jo, and others, "Illegally parked vehicle detection using adaptive dual background model," in IECON 2015-41st Annual Conference of the IEEE Industrial Electronics Society, 2015, pp. 2225–2228. https://doi.org/10.1109/iecon.2015.7392432

[149] M. Teutsch, W. Krüger, and J. Beyerer, "Evaluation of object segmentation to improve moving vehicle detection in aerial videos," in 2014 11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2014, pp. 265–270. https://doi.org/10.1109/avss.2014.6918679

[150] V. Reilly and M. Shah, "Detecting, tracking, and recognizing activities in aerial video," 2012.

[151] G. Kopsiaftis and K. Karantzalos, "Vehicle detection and traffic density monitoring from very high-resolution satellite video data," in 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015, pp. 1881–1884. https://doi.org/10.1109/igarss.2015.7326160

[152] T. Yang et al., "Small moving vehicle detection in a satellite video of an urban area," Sensors, vol. 16, no. 9, p. 1528, 2016.

[153] J. Zhang, X. Jia, and J. Hu, "Error bounded foreground and background modeling for moving object detection in satellite videos," IEEE Trans. Geosci. Remote Sens., vol. 58, no. 4, pp. 2659–2669, 2019. https://doi.org/10.1109/tgrs.2019.2953181

[154] D. D. Bloisi, A. Pennisi, and L. Iocchi, "Background modeling in the maritime domain," Mach. Vis. Appl., vol. 25, no. 5, pp. 1257–1269, 2014. https://doi.org/10.1007/s00138-013-0554-5

[155] D. Bloisi and L. Iocchi, "Independent multimodal background subtraction.," in CompI-MAGE, 2012, pp. 39–44.

[156] D. Socek, D. Culibrk, O. Marques, H. Kalva, and B. Furht, "A hybrid color-based foreground object detection method for automated marine surveillance," in International Conference on Advanced Concepts for Intelligent Vision Systems, 2005, pp. 340–347. https://doi.org/10.1007/11558484_43

[157] D. Culibrk, O. Marques, D. Socek, H. Kalva, and B. Furht, "A neural network approach to bayesian background modeling for video object segmentation.," in VISAPP (1), 2006, pp. 474–479. https://doi.org/10.1109/tnn.2007.896861

[158] S. Zhang, Z. Qi, and D. Zhang, "Ship tracking using background subtraction and inter-frame correlation," in 2009 2nd International Congress on Image and Signal Processing, 2009, pp. 1–4. https://doi.org/10.1109/cisp.2009.5302115

[159] S. Lee, N. Kim, I. Paek, M. H. Hayes, and J. Paik, "Moving object detection using unstable camera for consumer surveillance systems," in 2013 IEEE International Conference on Consumer Electronics (ICCE), 2013, pp. 145–146. https://doi.org/10.1109/icce.2013.6486833

[160] Z. Zhou, L. Shangguan, X. Zheng, L. Yang, and Y. Liu, "Design and implementation of an RFID-based customer shopping behavior mining system," IEEE/ACM Trans. Netw., vol. 25, no. 4, pp. 2405–2418, 2017. https://doi.org/10.1109/tnet.2017.2689063

[161] F. El Baf, T. Bouwmans, and B. Vachon, "Fuzzy foreground detection for infrared videos," in 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008, pp. 1–6. https://doi.org/10.1109/cvprw.2008.4563057

[162] W. Shen, Y. Lin, L. Yu, F. Xue, and W. Hong, "Single channel circular SAR moving target detection based on logarithm background subtraction algorithm," Remote Sens., vol. 10, no. 5, p. 742, 2018. https://doi.org/10.3390/rs10050742

[163] I. Ali, J. Mille, and L. Tougne, "Space--time spectral model for object detection in dynamic textured background," Pattern Recognit. Lett., vol. 33, no. 13, pp. 1710–1716, 2012. https://doi.org/10.1016/j.patrec.2012.06.011

[164] I. Ali, J. Mille, and L. Tougne, "Adding a rigid motion model to foreground detection: application to moving object detection in rivers," Pattern Anal. Appl., vol. 17, no. 3, pp. 567–585, 2014. https://doi.org/10.1007/s10044-013-0346-6

[165] R. Bastos, "Monitoring of human sport activities at sea," Master Thesis, Portugal, 2015.

[166] B. Tamás, "Detecting and analyzing rowing motion in videos," in BME Scientific Student Conf, 2016, pp. 1–29.

[167] F. John, I. Hipiny, and H. Ujir, "Assessing performance of aerobic routines using background subtraction and intersected image region," in 2019 International Conference on Computer and Drone Applications (IConDA), 2019, pp. 38–41. https://doi.org/10.1109/iconda47345.2019.9034912

[168] A. Darwich, P.-A. Hébert, A. Bigand, and Y. Mohanna, "Background subtraction based on a new fuzzy mixture of Gaussians for moving object detection," J. Imaging, vol. 4, no. 7, p. 92, 2018. https://doi.org/10.3390/jimaging4070092

[169] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "Changedetection. net: A new change detection benchmark dataset," in 2012 IEEE computer society conference on computer vision and pattern recognition workshops, 2012, pp. 1–8. https://doi.org/10.1109/cvprw.2012.6238919

[170] Q. Li, E. A. Bernal, M. Shreve, and R. P. Loce, "Scene-independent feature-and classifier-based vehicle headlight and shadow removal in video sequences," in 2016 IEEE Winter Applications of Computer Vision Workshops (WACVW), 2016, pp. 1–8. https://doi.org/10.1109/wacvw.2016.7470115

[171] Y. Sheikh and M. Shah, "Bayesian object detection in dynamic scenes," in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, vol. 1, pp. 74–79. https://doi.org/10.1109/cvpr.2005.86

[172] Y. Sheikh and M. Shah, "Bayesian modeling of dynamic scenes for object detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 11, pp. 1778–1792, 2005. https://doi.org/10.1109/tpami.2005.213

[173] A. ElTantawy and M. S. Shehata, "A novel method for segmenting moving objects in aerial imagery using matrix recovery and physical spring model," in 2016 23rd International Conference on Pattern Recognition (ICPR), 2016, pp. 3898–3903. https://doi.org/10.1109/icpr.2016.7900243

[174] A. ElTantawy and M. S. Shehata, "Moving object detection from moving platforms using Lagrange multiplier," in 2015 IEEE International Conference on Image Processing (ICIP), 2015, pp. 2586–2590. https://doi.org/10.1109/icip.2015.7351270

[175] A. ElTantawy and M. S. Shehata, "UT-MARO: Unscented transformation and matrix rank optimization for moving objects detection in aerial imagery," in International Symposium on Visual Computing, 2015, pp. 275–284. https://doi.org/10.1007/978-3-319-27857-5_25

[176] A. ElTantawy and M. S. Shehata, "MARO: Matrix rank optimization for the detection of small-size moving objects from aerial camera platforms," Signal, Image Video Process., vol. 12, no. 4, pp. 641–649, 2018. https://doi.org/10.1007/s11760-017-1203-7

[177] A. ElTantawy and M. S. Shehata, "KRMARO: Aerial detection of small-size ground moving objects using kinematic regularization and matrix rank optimization," IEEE Trans. Circuits Syst. Video Technol., vol. 29, no. 6, pp. 1672–1686, 2018. https://doi.org/10.1109/tcsvt.2018.2843761

[178] E. Toropov, L. Gui, S. Zhang, S. Kottur, and J. M. F. Moura, "Traffic flow from a low frame rate city camera," in 2015 IEEE International Conference on Image Processing (ICIP), 2015, pp. 3802–3806. https://doi.org/10.1109/icip.2015.7351516

[179] L. Li, W. Huang, I. Y.-H. Gu, and Q. Tian, "Statistical modeling of complex backgrounds for foreground object detection," IEEE Trans. Image Process., vol. 13, no. 11, pp. 1459–1472, 2004. https://doi.org/10.1109/tip.2004.836169

[180] S. Singh, S. A. Velastin, and H. Ragheb, "Muhavi: A multicamera human action video dataset for the evaluation of action recognition methods," in 2010 7th IEEE International Conference on Advanced Video and Signal Based Surveillance, 2010, pp. 48–55. https://doi.org/10.1109/avss.2010.63

[181] L. Maddalena and A. Petrosino, "Towards benchmarking scene background initialization," in International conference on image analysis and processing, 2015, pp. 469–476.

[182] C. Cuevas, E. M. Yáñez, and N. Garc\'\ia, "Labeled dataset for integral evaluation of moving object detection algorithms: LASIESTA," Comput. Vis. Image Underst., vol. 152, pp. 103–117, 2016. https://doi.org/10.1016/j.cviu.2016.08.005

[183] M. I. Chacon-Murguia, J. A. Ramirez-Quintana, and G. Ramirez-Alonso, "Evaluation of the background modeling method auto-adaptive parallel neural network architecture in the sbm-net dataset," in 2016 23rd International Conference on Pattern Recognition (ICPR), 2016, pp. 137–142. https://doi.org/10.1109/icpr.2016.7899622

[184] M. Camplani, L. Maddalena, G. M. Alcover, A. Petrosino, and L. Salgado, "A benchmarking framework for background subtraction in RGBD videos," in International Conference on Image Analysis and Processing, 2017, pp. 219–229. https://doi.org/10.1007/978-3-319-70742-6_21

[185] F. El Baf, T. Bouwmans, and B. Vachon, "Comparison of background subtraction methods for a multimedia application," in 2007 14th International Workshop on Systems, Signals and Image Processing and 6th EURASIP Conference focused on Speech and Image Processing,

Multimedia Communications and Services, 2007, pp. 385–388. https://doi.org/10.1109/iws-sip.2007.4381122

[186] X. P. Burgos-Artizzu, P. Dollár, D. Lin, D. J. Anderson, and P. Perona, "Social behavior recognition in continuous video," in 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 1322–1329. https://doi.org/10.1109/cvpr.2012.6247817

[187] T. Forrester, W. J. McShea, R. W. Keys, R. Costello, M. Baker, and A. Parsons, "eMammal--citizen science camera trapping as a solution for broad-scale, long-term monitoring of wildlife populations," Sustain. Pathways Learn. from Past Shap. Futur., 2013.

[188] I. Kavasidis, S. Palazzo, R. Di Salvo, D. Giordano, and C. Spampinato, "An innovative web-based collaborative platform for video annotation," Multimed. Tools Appl., vol. 70, no. 1, pp. 413–432, 2014. https://doi.org/10.1007/s11042-013-1419-7

[189] S. Beery, G. Van Horn, and P. Perona, "Recognition in terra incognita," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 456–473.

[190] J. Wei, J. Zhao, Y. Zhao, and Z. Zhao, "Unsupervised anomaly detection for traffic surveillance based on background modeling," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 129–136. https://doi.org/10.1109/cvprw.2018.00025

[191] L. A. Marcomini and A. L. Cunha, "A comparison between background modelling methods for vehicle segmentation in highway traffic videos," arXiv Prepr. arXiv1810.02835, 2018.

[192] L. O. Andrews Sobral, L. Schnitman, and F. De Souza, "Highway traffic congestion classification using holistic properties," in 10th IASTED International Conference on Signal Processing, Pattern Recognition and Applications, 2013. https://doi.org/10.2316/p.2013.798-105

[193] T. Bouwmans, F. Porikli, B. Höferlin, and A. Vacavant, Background modeling and foreground detection for video surveillance. CRC press, 2014. https://doi.org/10.1201/b17223

[194] A. Sobral, T. Bouwmans, and E. Zahzah, "Lrslibrary: Low-rank and sparse tools for background modeling and subtraction in videos," Robust Low-Rank Sparse Matrix Decompos. Appl. Image Video Process., 2016. https://doi.org/10.1201/b20190-24

[195] D. Bloisi, "Background modeling and foreground detection for maritime video surveillance," Chapter Handb. Backgr. Model. Foregr. Detect. Video Surveill. Tradit. Recent Approaches, Implementations, Benchmarking Eval. Chapman Hall/CRC, pp. 11–14, 2014. https://doi.org/10.1201/b17223-18

[196] S. Zeevi, "Background Subtractor CNT Project." 2019.

[197] J. Han, J. Pei, and M. Kamber, Data mining: concepts and techniques. Elsevier, 2011.

11 Authors

Maryam A. Yasir received her bachelor's degree in computer science from the University of Baghdad (UOB) -Iraq 2004. Since 2008, she is working as a lecturer at the computer science department, college of science, University of Baghdad up till now. In 2011 she received a certificate for a seven months course in IT Administration from the Technische Universität Berlin (TU)Berlin-Germany. In 2014 she received her master's degree in computer science from University Putra Malaysia (UPM)-Malaysia. Meanwhile, she is a PhD candidate at the University of Technology (UOT) -Iraq. Maryam has participated in many scholar courses and activities, local and abroad including Fulbright visiting scholar at the University of Central Oklahoma (UCO)-US 2015.

Dr. Yossra Hussain Ali is an Assistant Professor. She received her B.Sc , M.Sc and PhD degrees in 1996, 2002 and 2006 respectively from Iraq, University of technology, department of Computer Sciences. She Joined the University of Technology, Iraq in 1997. During her postgraduate studies, she worked on Computer Network, Information systems, Agent Programming and Image Processing as well as some experience in Artificial Intelligent and Computer Data Security. She is a reviewer at many conferences and journals and she supervised a number of undergraduates and postgraduates (PhD. and MSc.) dissertations in Computer sciences. Yossra has many professional certificates and she has published in well regarded journals . (e-mail: 110017@uotechnology.edu.iq)