Computer Vision and Image Understanding 152 (2016) 41–57

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Detection of stationary foreground objects: A survey

Carlos Cuevas*, Raquel Martínez, Narciso García

Grupo de Tratamiento de Imágenes, Universidad Politécnica de Madrid (UPM), E-28040 Madrid, Spain

ARTICLE INFO

Article history:
Received 23 October 2015
Revised 1 July 2016
Accepted 1 July 2016
Available online 2 July 2016

Keywords:
Stationary foreground
Abandoned object
Removed object
Background subtraction
Survey
State of the art
Overview

ABSTRACT

Detection of stationary foreground objects (i.e., moving objects that remain static throughout several frames) has attracted the attention of many researchers over the last decades and, consequently, many new ideas have been recently proposed, trying to achieve high-quality detections in complex scenarios with the lowest misdetections, while keeping real-time constraints. Most of these strategies are focused on detecting abandoned objects. However, there are some approaches that also allow detecting partially-static foreground objects (e.g. people remaining temporarily static) or stolen objects (i.e., objects removed from the background of the scene).

This paper provides a complete survey of the most relevant approaches for detecting all kind of stationary foreground objects. The aim of this survey is not to compare the existing methods, but to provide the information needed to get an idea of the state of the art in this field: kinds of stationary foreground objects, main challenges in the field, main datasets for testing the detection of stationary foreground, main stages in the existing approaches and algorithms typically used in such stages.

1. Introduction

Applications for video analysis and understanding (e.g. video surveillance (Liu et al., 2015), augmented reality (Pauly et al., 2015), or analysis of people behavior (Morozov, 2015)) typically include strategies for separating the moving objects (MOs) in the scene, called foreground (FG), from the static information, called background (BG). These strategies, commonly known as background subtraction strategies, have been widely studied since the 90s and currently there are thousands of algorithms to carry them out. Moreover, several surveys of background subtraction have also been published up to date (Bouwmans, 2014; Cristani et al., 2010; Elhabian et al., 2008; Sobral and Vacavant, 2014).

Some of these applications (Evangelio and Sikora, 2011; Kim and Kang, 2014; Zeng et al., 2015) have shown a growing interest in detecting a particular type of FG objects: those that stop and remain static throughout several consecutive frames, called stationary foreground objects (SFOs). The graph in Fig. 1 illustrates that the cumulative number of works proposed over the last 20 years to detect SFOs has rapidly grown. Moreover, it can also be observed that the rate of emergence of new proposals has increased since 2006 (see the knee in the graph), due to the detection of SFOs has been the key theme in some relevant computer vision conferences.

Most of these applications are focused exclusively on detecting abandoned objects (Lim and Davis, 2006) or people with suspicious behavior (Foresti et al., 2002) to prevent terrorist incidents and reduce crime in public areas (e.g. airports, railways, subway stations or sport event venues). On the other hand, there is a small portion of these applications (Ferrando et al., 2006) that is not only interested in the detection of abandoned objects, but in detecting stolen objects in public places (e.g. museums or stores). Moreover, there are also some recent strategies (Martínez et al., 2015) that detect SFOs with the aim of improving the quality of classical moving object detection algorithms in scenarios featuring moving objects that stop frequently (e.g. people in offices or vehicles on urban roads). In these scenarios, the classical algorithms frequently lead to misdetections that can be avoided by identifying the moving objects remaining temporarily static (Bouwmans et al., 2008; Sobral and Vacavant, 2014).

Despite the increasing demand for applications where the detection of SFOs is critical and the consequent emergence of many approaches focused on such detection task, up to our knowledge, the only existing work that analyzes the existing algorithms is a survey of 2009 (Bayona et al., 2009). Since many relevant SFO detection approaches have been published since then, this survey is outdated. Besides, it only analyzes a very specific subset of the existing methods.

This paper provides a new and complete survey of the most relevant SFO detection strategies published throughout the last two decades. This survey neither attempts to compare the SFO detection methods among them, nor aims to analyze the advantages

* Corresponding author.
E-mail addresses: ccr@gti.ssr.upm.es (C. Cuevas), rms@gti.ssr.upm.es (R. Martínez), narciso@gti.ssr.upm.es (N. García).

http://dx.doi.org/10.1016/j.cviu.2016.07.001
1077-3142/

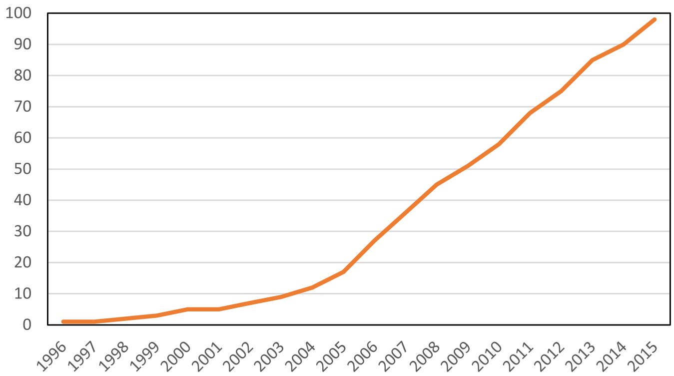

Fig. 1. Cumulative number of relevant publications related to the detection of SFOs over the last 20 years.

and disadvantages of each method. The purpose of this survey is to offer an updated and complete overview of the most relevant aspects in the field of SFO detection: the main stages in the detection process, the most typical algorithms applied to each stage, the databases that can be used to assess the quality of this kind of detection methods, etc. In this way, the researchers beginners in this topic will be able to have an overview of the state of the art, the current trends and the challenges of interest. Moreover, although the descriptions of the methods listed in this survey are generally short, they provide enough information to understand if each strategy is appropriate or not for the purposes of the reader. For example, it will be seen that some detection algorithms can only be applied if a pixel-based BG subtraction is first used, whereas other detection algorithms take as starting point the object-level results provided by tracking methods. Moreover, this survey also provides information regarding the typical challenges addressed by each of the reviewed strategies. In this way, readers will be able to easily locate the algorithms that best suit their needs.

All the reviewed strategies and their main characteristics appear summarized in Tables 1 and 2, sorted according to their publication year. A more detailed analysis of such characteristics will be performed along the rest of the paper. First, the main challenges in the detection of SFOs and the public datasets typically used to assess the quality of the detections are described, respectively, in Sections 2 and 3. In Section 4, the kinds of SFOs defined in the literature and the amount of works focusing in each definition are discussed. Then, the algorithms used to carry out the main stages in the detection of SFOs are described in Sections 5, 6 and 7. These stages are the following:

- FG detection (Section 5): A motion detection strategy, commonly based on BG subtraction methods, is applied to locate the objects not belonging to the BG of the scene.
- SFO detection (Section 6): The detected FG objects are analyzed to try to determine if they have stopped moving or not.
- Removed object (RO) detection (Section 7): Some strategies include this third stage to try to discriminate between abandoned objects and removed (e.g. stolen) objects.

Finally, Section 8 presents some brief conclusions of the analysis that is carried out in this survey.

2. Challenges

Most SFO detection strategies include algorithms for the detection of FG. Therefore, some of the challenges in the detection of FG are also challenges for the detection of SFOs. In addition, there are some challenges directly associated with the detection of SFOs, which are related to the speed and persistence of the detections and the capability of the algorithms to deal with some specific situations (e.g. SFOs occluded by FG objects).

2.1. Challenges in MO detection

- Image noise: It appears either in sequences recorded with poor quality cameras or after applying a compression process on the video.
- Illumination changes: They can be gradual (e.g., light variations along the day) or sudden (e.g., turning on the lights in a room) and cause many false detections in large areas of the images.
- Low contrast: The FG detection methods must be able to detect the moving objects in sequences with low contrast. This situation is typical in sequences recorded at night.
- Camera automatic adjustments: Some automatic adjustments of the modern cameras (e.g. auto focus, gain control, white balance and brightness control) make difficult to achieve successful detections, since these adjustments modify the dynamic of the color level of the pixels.
- Dynamic BG: Many sequences contain BG elements that are not completely static but move periodically or irregularly (e.g., waving flags, trees and shrubs shaken by the wind, escalators, or water waves). Despite being in motion, these moving elements must be considered as part of the BG.
- Camera jitter: The sequences may have been recorded with non-stabilized cameras (e.g., a camera endowed in a mobile phone or fixed cameras affected by the wind). This camera motion typically results in many false detections.
- FG aperture: If a FG object has regions with uniform colors, the changes inside these regions may not be detected.
- Camouflage: A FG object and the BG behind it can have similar appearance, which complicates distinguishing between them. Occasionally, when the camouflage is very intense, the FG object can be detected only if their shape is previously known.
- Shadows (SHs) and highlights (HLs): The shadows and highlights cast by moving objects are commonly detected as part of such the FG, which significantly decreases the quality of the detections. This problem appears in outdoor sequences, where hard shadows typically appear, and also in most indoor scenarios, where the moving objects produce medium and soft shadows and highlights.
- Bootstraping: In some cases, a training period (images free from FG) to obtain an initial representative BG model is not available.

2.2. Challenges in SFO detection

- Occluded SFOs: An abandoned object can be temporarily occluded by a second object. This second object can move in front of the first one, or it can even stop just when it is placed in front of the first object (becoming a new SFO). In these cases, the correct detection of the initial abandoned object can fail both during the occlusion and after it. Moreover, this case is further aggravated if the first object starts moving when it is occluded by the second object. Note that this example can be extended to other cases with multiple objects overlapping simultaneously.
- Long-term SFOs: The FG objects that remain static very long periods of time typically end up not being detected (they are incorporated to the BG).
- Partially-stationary foreground objects (PSFOs): In many video-surveillance scenarios (e.g. airports, malls, offices, etc.) many people become SFOs for a while. However, it is not realistic to assume that these people remain completely static when they stop walking, since their upper body (torso, arms and head) is not usually completely static. Nevertheless, if the static area of

Table 1
Main strategies proposed over the last years for detecting SMOs (part 1).

Strategy	Type	BG subtraction	Num. models	SFO detection	Detection level	Databases
1996-Gibbins (Gibbins et al., 1996)	A, B	None	–	Classifier	Object	Other
1998-Murino (Murino, 1998)	A	None	–	Classifier	Object	Other
1999-Wang (Wang and Ooi, 1999)	A, B, C	Basic	1	Persistence	Object	Other
2000-Sacchi (Sacchi and Regazzoni, 2000)	A	Basic	1	Persistence	Pixel	Other
2000-Stringa (Stringa and Regazzoni, 2000)	A	Basic	1	Persistence	Pixel	Other
2002-Amer (Amer et al., 2002)	A	–	–	Tracking	Object	Other
2002-Foresti (Foresti et al., 2002)	A	Basic	1	Tracking	Object	Other
2003-Beynon (Beynon et al., 2003)	A	SGM	1	Tracking	Object	Other
2003-Spengler (Spengler and Schiele, 2003)	A	GMM	1	Tracking	Object	Other
2004-Aubert (Aubert et al., 2004)	A, B, C	None	–	DFC	Pixel	Other
2004-Fujiyoshi (Fujiyoshi and Kanade, 2004)	A, B	Basic	1	Persistence	Pixel	Other
2004-Yang (Yang et al., 2004)	A, B, C	Basic	2	DFC	Pixel	Other
2005-Black (Black et al., 2005)	A, B, C	MM	1	Tracking	Object	Other
2005-Mathew (Mathew et al., 2005)	A, B, C	GMM	1	GS	Pixel	PETS('01, '04)
2005-Melli (Melli et al., 2005)	B	MM	1	Tracking	Object	Other
2005-Tian (Tian et al., 2005)	A, B, C	GMM	1	GS	Pixel	PETS'01
2005-Velastin (Velastin et al., 2005)	A, B, C	MM	1	Tracking	Object	Other
2006-Auvinet (Auvinet et al., 2006)	A	MM	1	Tracking	Object	PETS'06
2006-Bird (Bird et al., 2006)	A	GMM	1	Tracking	Object	Other
2006-Ferrando (Ferrando et al., 2006)	A	Basic	1	Tracking	Object	Other
2006-Guler (Guler and Farrow, 2006)	A	GMM	1	Persistence	Object	PETS'06
2006-Li (Li et al., 2006)	A	None	–	Tracking	Object	PETS'06
2006-Lim (Lim and Davis, 2006)	A	OF	1	Other	Pixel	Other
2006-Lv (Lv et al., 2006)	A	–	1	Tracking	Object	PETS'06
2006-Martínez (Martínez-del Rincón et al., 2006)	A	MM	2	Tracking	Object	PETS'06
2006-Smith (Smith et al., 2006)	A	GMM	1	Tracking	Object	PETS'06
2006-Spagnolo (Spagnolo et al., 2006)	A	SGM	1	Tracking	Object	PETS'06
2007-Arsic (Arsic et al., 2007)	A, C	GMM	1	Tracking	Object	PETS'07
2007-Bevilacqua (Bevilacqua and Vaccari, 2007)	B	Other	1	Tracking	Object	i-LIDS
2007-Bhargava (Bhargava et al., 2007)	A	OF	1	Classifier - Tracking	Object	i-LIDS
2007-Cheng (Cheng et al., 2007)	A, B	CM	2	DFC	Pixel	PETS'06
2007-Chuang (Chuang et al., 2007)	A	GMM	1	Tracking	Object	Other
2007-Dalley (Dalley et al., 2007)	A	SGM	1	Tracking	Object	PETS'07
2007-Denman (Denman et al., 2007)	A	CM	1	Persistence	Pixel	PETS'06
2007-Guler (Guler et al., 2007)	A, B	Basic	1	Tracking	Object	i-LIDS
2007-Porikli (Porikli, 2007)	A, B	GMM	2	DFC	Pixel	PETS'06, i-LIDS
2008-Cho (Cho et al., 2008)	A	GMM	1	Other	Object	Other
2008-Gallego (Gallego et al., 2008)	A, B, C	SGM	1	Persistence	Pixel	PETS'01
2008-Liao (Liao et al., 2008)	A	SGM	1	Persistence	Object	PETS'06, i-LIDS
2008-Lin (Lin and Wang, 2008)	A	GMM	01-Feb	DFC - GS	Pixel	Other
2008-Miezianko (Miezianko and Pokrajac, 2008)	A	SLM	1	Classifier	Object	PETS('06,'07), i-LIDS, CAVIAR
2008-Porikli (Porikli et al., 2008)	A, B	GMM	2	DFC	Pixel	PETS'06, i-LIDS
2008-Rodriguez (Rodriguez-Fernandez et al., 2008)	A, B	Basic	1	Persistence	Pixel	i-LIDS
2008-SanMiguel (San Miguel and Martínez, 2008)	A	SGM	1	Tracking	Object	PETS'06, i-LIDS
2008-Tian (Tian et al., 2008)	A, B, C	GMM	1	GS	Pixel	PETS'06, i-LIDS
2009-Chuang (Chuang et al., 2009)	A	GMM	1	Tracking	Object	Other
2009-Lee (Lee et al., 2009)	B	GMM	1	Tracking	Object	i-LIDS
2009-Li (Li et al., 2009)	A, B	SGM	2	DFC	Pixel	i-LIDS, CAVIAR
2009-Magno (Magno et al., 2009)	A	–	1	Other	Object	Other

these objects is large enough, they can lead to erroneous detection of abandoned objects.

- Removed objects (ROs): The identification of situations in which a BG object is removed by someone is of great interest in many surveillance applications. However, this situations are easily mistaken with object abandonments.
- Ghost regions (GRs): When a moving object stops moving it will eventually be incorporated into the BG model. If the object now begins to move, the area it previously occupied will be incorrectly detected as a FG blob, commonly referred to as a ghost. This ghost will remain until the BG model adapts to the newly exposed BG. The GRs are typical in scenes with parked vehicles that start moving.

3. Databases

The last column in both Tables 1 and 2 shows the databases used to assess the quality of the reviewed strategies. Note that for many of these strategies, the label in such column is “Other”, since they use databases that are not public and that, typically, have been recorded by the own authors. The remaining strategies use one or multiple of the following public datasets:

- PETS: The Performance Evaluation of Tracking and Surveillance (PETS) program (PET, 2000-2007) collects videos for evaluating surveillance algorithms from 2000 to 2007. Its content varies from simple and uncrowded sequences to crowdy complex scenarios recorded from multiple viewpoints. Those collections including sequences with abandoned and subtracted objects (PETS'01, PETS'04, PETS'06, PETS'07) have been used by almost 50% of the reviewed works.
- i-LIDS: The Imagery Library for Intelligent Detection Systems (i-LIDS) (i LIDS Team, 2006) contains several sequences to test 4 situations: abandoned objects, parked vehicles, doorway and restricted areas. All the sequences are supplied with XML files describing temporal events and the size and spatial position of the moving objects. This database has been the second most used in the literature (it has been used by the 34% of the reviewed strategies).

Table 2
Main strategies proposed over the last years for detecting SMOs (part 2).

Strategy	Type	BG subtraction	Num. models	SFO detection	Detection level	Databases
2009-Singh (Singh et al., 2009)	A	MM	2	DFC - Tracking	Object	PETS'06, i-LIDS
2009-Wen (Wen et al., 2009)	A	Combination	1	Classifier	Object	Other
2010-Bayona (Bayona et al., 2010)	A, B	SGM	1	Persistence	Object	PETS('06,'07), i-LIDS
2010-Chang (Chang et al., 2010)	A	Basic	1	Persistence	Object	PETS'06, i-LIDS
2010-Evangelio (Evangelio and Sikora, 2010)	A, B	GMM	2	DFC - FSM	Pixel	PETS'06, i-LIDS, CAVIAR
2010-Kwak (Kwak et al., 2010)	A	SGM	1	Tracking - FSM	Object	PETS'06
2010-Li (Li et al., 2010)	A	GMM	2	DFC - Classifiers	Pixel	PETS('06,'07)
2010-Singh (Singh et al., 2010)	A	GMM	1	Tracking	Object	Other
2010-Wang (Wang and Liu, 2010)	A, B	GMM	2	DFC	Pixel	i-LIDS
2011-Albiol (Albiol et al., 2011)	B	None	–	Classifier	Object	i-LIDS
2011-Evangelio (Evangelio and Sikora, 2011)	A, B	GMM	2	DFC - FSM	Pixel	PETS'06, i-LIDS, CAVIAR
2011-Fan (Fan and Pankanti, 2011)	A, B, C	GMM	1	GS - FSM	Object	i-LIDS
2011-Fu (Fu et al., 2011)	A, B	GMM	1	Persistence - Tracking	Object	Other
2011-Geng (Geng and Xiao, 2011)	A	CM	1	Other	Object	PETS'01
2011-Pan (Pan et al., 2011)	A, B, C	Other	1	Persistence	Object	i-LIDS
2011-Raheja (Raheja et al., 2011)	A	GMM	1	GS	Pixel	Other
2011-Tian (Tian et al., 2011)	A, B, C	GMM	1	GS	Pixel	PETS'06, i-LIDS
2011-Yang (Yang and Rothkrantz, 2011)	A	CM	1	Classifier - Tracking	Object	Other
2011-Zin (Zin et al., 2011)	A, B, C	Basic	2	DFC	Object	PETS'06
2012-Bangare (Bangare et al., 2012)	A	GMM	1	Tracking	Object	Other
2012-Fan (Fan and Pankanti, 2012)	A	GMM	1	GS - FSM	Object	i-LIDS
2012-Kim (Kim et al., 2012)	A, B	MM	1	Classifier	Object	PETS('06,'07), i-LIDS
2012-Lai (Lai et al., 2012)	A	Other	1	DFC	Object	Other
2012-Morde (Morde et al., 2012)	A	Other	1	Persistence	Pixel	PETS'06, Change detection
2012-Xiya (Xiya et al., 2012)	A	Basic	2	DFC - Tracking	Object	PETS'06
2012-Zin (Zin et al., 2012)	A, B, C	Basic	2	DFC	Pixel	PETS'06
2013-Chang (Chang et al., 2013)	A	SLM	1	Classifier - Tracking	Object	PETS'06, i-LIDS
2013-Collazos (Collazos et al., 2013)	A, B, C	Basic	1	Persistence - FSM	Object	Other
2013-Fan (Fan et al., 2013)	A	GMM	1	GS - FSM	Object	PETS'06, i-LIDS
2013-Ferryman (Ferryman et al., 2013)	A, C	GMM	2	DFC - Classifiers	Object	PETS'06
2013-Hassan (Hassan et al., 2013)	A, B, C	GMM	1	Persistence	Pixel	PETS'06, i-LIDS
2013-Maddalena (Maddalena and Petrosino, 2013)	A, B, C	NN	1	Persistence	Pixel	i-LIDS, VISOR
2013-Muchtar (Muchtar et al., 2013)	A	GMM	2	DFC - Persistence	Pixel	CAVIAR
2013-Ortego (Ortego and SanMiguel, 2013)	A, B, C	SGM	1	Persistence	Pixel	PETS('06,'07), i-LIDS
2013-Sajith (Sajith and Nair, 2013)	A	CM	2	DFC - Tracking	Object	PETS('06,'07)
2013-Tripathi (Tripathi et al., 2013)	A	SGM	1	Persistence	Object	PETS('06,'07)
2014-Joglekar (Joglekar et al., 2014)	A	GMM	2	DFC - Tracking	Object	Other
2014-Kim (Kim and Kang, 2014)	A	None	–	Classifiers - FSM	Object	PETS'06, i-LIDS, CAVIAR
2014-Lopez (Lopez-Mendez et al., 2014)	A	Other	1	GS	Object	PETS'06
2014-Ortego (Ortego and SanMiguel, 2014)	A, B, C	SGM	1	Classifier	Pixel	PETS('06,'07), i-LIDS
2014-Wang(Wang et al., 2014)	A, B, C	Combination	2	DFC	Object	Change detection
2015-Baxter (Baxter et al., 2015)	A	Basic	1	Persistence	Object	PETS'06
2015-Ingersoll (Ingersoll et al., 2015)	A, B	GMM	1	Tracking	Object	PETS'06
2015-Jardim (Jardim et al., 2015)	A	SLM	1	Classifier	Object	VDAO
2015-Lin (Lin et al., 2015)	A	GMM	2	DFC - FSM - Tracking	Pixel	i-LIDS, PETS'06
2015-Mahin (Mahin et al., 2015)	A	Basic	1	Tracking	Object	Other
2015-Martínez (Martínez et al., 2015)	A, B, C	KDE	3	DFC - FSM	Pixel	LASIESTA
2015-Wahyono (Filonenko et al., 2015)	A, B, C	Basic	2	DFC - Tracking	Object	PETS'06, i-LIDS
2015-Zeng (Zeng et al., 2015)	A, B	Combination	2	DFC	Pixel	Other

- CAVIAR: The Content Aware Vision using Image-based Active Recognition (CAVIAR) dataset (Fisher et al., 2005) contains surveillance sequences recorded in different scenarios of interest. These include people walking alone, meeting with others, fighting or abandoning packages.
- Change detection (Goyette et al., 2012): It is composed by 31 video sequences with more than 80,000 images recorded in real indoor and outdoor scenarios. The sequences are classified in 6 categories that address different challenges, such as dynamic BG, camera jitter, or intermittent object motion.
- VISOR: The Video Surveillance Online Repository (VISOR) (Vezzani and Cucchiara, 2010) contains a large set of multimedia data distributed among 14 categories. The sequences in some of these categories contain abandoned objects, vehicles that are parked and people remaining partially static.
- VDAO: The Video Database for Abandoned-Object Detection in a Cluttered Environment (VDAO) (da Silva et al., 2014) comprises more than 8 hours of video sequences recorded in scenarios of different complexity, with different amount of cameras and different light conditions. Along with i-LIDS, this database is the only one specifically created to test the quality of strategies for detecting abandoned objects.
- LASIESTA: The Labeled and Annotated Sequences for Integral Evaluation of SegmenTation Algorithms (LASIESTA) database (LAS, 2016) is composed by 48 video sequences to evaluate several challenges in FG detection algorithms, such as camouflage, occlusions, illumination changes, subtractions, abandonments or camera motion. This database is the only one that is fully annotated and both pixel and object levels.

4. Types of SFO

The SFOs can be classified into three types:

- Type A: Inert MO that is abandoned by a human. This definition comprises very typical cases such as, for example, abandoning a backpack or a suitcase.
- Type B: Inert MO that becomes static without any apparent interaction with humans. This is the case of a vehicle that has

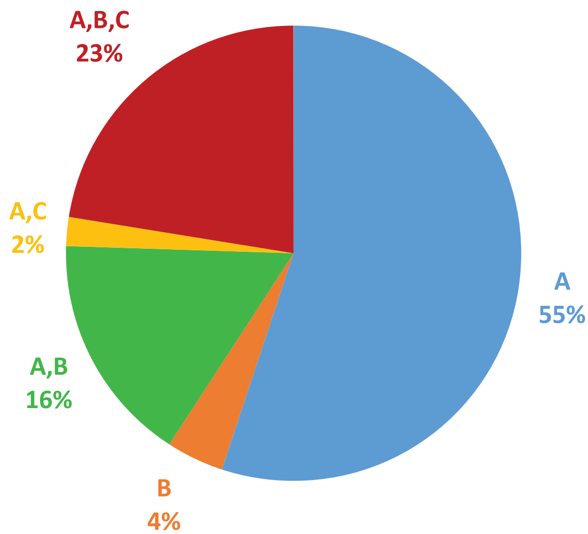

Fig. 2. Proportions of strategies that detect the different types of SFO.

been parked, or an object thrown into the scene by a person who is not within the field of view of the camera.
- Type C: A moving person that becomes totally or partially static.

The second column in both Tables 1 and 2 shows the type of SFO considered by the reviewed works and Fig. 2 illustrates the proportions of works that focus on each type.

It can be observed that most methods (55%) only consider the type A, since detecting abandoned objects is a key task commonly required by security applications (Spengler and Schiele, 2003), where the abandoned objects are those generating most interest. Despite the large amount of works proposed since 1996, many algorithms for detecting SFOs of type A have been recently proposed, trying to achieve high-quality detections in complex scenarios or improve the computational efficiency and the usability of previous approaches. For example, the strategy in Sajith and Nair (2013) tries to detect SFOs in complex surveillance videos with illumination changes or occlusions, whereas the algorithms proposed in Kim and Kang (2014) try to reduce the amount of threshold values required by previous approaches.

Some of these algorithms are able to detect more than one type of SFO. However, they discard the detections corresponding to types B and/or C in post-processing stages. For example, in Sajith and Nair (2013) the objects of other types are identified and discarded by a matching method based on edge detection and user defined parameters such as size and position. Other strategies include mechanisms to detect not only the SFOs but also their owners. For example, the algorithm in Chang et al. (2013) performs this task using online classifiers, whereas the strategies in Foresti et al. (2002) and Ferryman et al. (2013) use trackers to relate abandoned objects and the people who left them. By the identification of the owners of the objects, these strategies are able of selecting only those objects abandoned by people and discard the rest of SFOs (i.e. types B and C).

In the case of SFOs of type B, the algorithms designed to detect only this kind of objects are focused on the detection of vehicles that have stopped. Some of these approaches analyze the behavior of the vehicles in roads, such as the one proposed in Melli et al. (2005), which describes a probabilistic and predictive tracking strategy to detect stopped vehicles in tunnels and highways from cluttered views. In other cases, the developed algorithms analyze urban scenarios (Albiol et al., 2011) or mixed scenarios (roads and cities) (Bevilacqua and Vaccari, 2007). So, in Bevilacqua and Vaccari (2007) a method to detect stopped vehicles and measure their stopping time is described, which is able to deal with tracking errors due to camera vibrations or partial occlusions. In Albiol et al. (2011) an algorithm to detect the presence of vehicles parked in street lanes is proposed, which is able to provide robust results in complex scenarios (e.g. high traffic density or illumination changes) by the detection and analysis of features in cars and asphalt.

Finally, as can be seen in Fig. 2, there are no strategies focused only in the detection of SFOs of type C. However, many strategies to detect multiple types of SFO have been proposed. Some algorithms try to detect types A and B (Bayona et al., 2010; Kim et al., 2012; Rodriguez-Fernandez et al., 2008), whereas other approaches are focused on the detection of the 3 types of SFOs (Maddalena and Petrosino, 2013; Mathew et al., 2005; Wang et al., 2014).

5. BG subtraction

The first stage in most strategies focused on the detection of SFOs consists in separating the FG from the rest of elements in the scene using a BG subtraction algorithm. As stated in the introduction, BG subtraction is a crucial stage not only in the detection of SFOs but in many computer vision applications such as video surveillance, multimedia or augmented reality. The typical scheme used to detect FG objects by subtracting the BG comprises the following three steps:

- BG initialization: An initial BG model, which must not contain FG objects, is constructed from data of one or more frames at the beginning of the sequence.
- FG detection: By comparing the current frame with the BG model, each pixel is classified as BG or FG.
- BG maintenance: The BG model is updated along time to adapt the changes in the BG.

To be robust against illumination changes or permanent BG changes (e.g. a door that is opened or an object moved by someone), objects that stop moving must be integrated in the BG model. However, by doing this the SFOs are also absorbed by the BG. Consequently, the third step in these scheme is crucial for SFO detection strategies, since they must be able of selectively updating the BG model.

The BG subtraction methods used to detect SFOs vary widely. Some of them (Collazos et al., 2013; Zin et al., 2012) use non-statistical models that, in the simplest cases, are never updated (Wang and Ooi, 1999). On the other hand, other strategies use popular statistical BG modeling approaches that are able to deal with very complex scenarios (e.g. dynamic BG and illumination changes). Moreover, some authors modify these typical approaches to improve the results in some situations (e.g. long-term SFOs or object removal).

As it is shown in the fourth column of both Tables 1 and 2, most authors propose strategies that are based on using a single BG model. However, it is possible to find some of them using two (Cheng et al., 2007) or three (Martínez et al., 2015) models.

The BG subtraction methods used by the reviewed SFO detection strategies have been classified into 12 categories (third column of Tables 1 and 2), which are described in the following subsections. The percentages of SFO detection strategies encompassed within each of these categories appears in Fig. 3.

The left half of the Tables 3 and 4 show a summary of the MO detection challenges that are addressed by each of the BG subtraction algorithms in the reviewed strategies. This summary allows a quick comparison between the analyzed methods and, therefore, it helps to identify the advantages and disadvantages of the methods that are described in the following subsections.

Table 3
Summary of challenges addressed by the reviewed strategies (part 1).

Strategy	BG subtraction method	MO detection										SFO detection method	SFO detection				
		Image noise	Illum. changes	Low contrast	Camera adjust.	Dynamic BG	Camera jitter	FG aperture	Camouflage	SHs & HLs	Bootstraping		Occlusions	Long-term	PSFO	RO	GR
1996-Gibbins (Gibbins et al., 1996)	None											Classifier		x			x
1998-Murino (Murino, 1998)	None											Classifier		x			x
1999-Wang (Wang and Ooi, 1999)	Basic			x				x	x			Persistence		x			
2000-Sacchi (Sacchi and Regazzoni, 2000)	Basic		x	x				x	x			Persistence		x			
2000-Stringa (Stringa and Regazzoni, 2000)	Basic			x				x	x			Persistence		x			
2002-Amer (Amer et al., 2002)	–											Tracking	x	x			
2002-Foresti (Foresti et al., 2002)	Basic			x			x	x	x			Tracking	x	x			
2003-Beynon (Beynon et al., 2003)	SGM	x	x	x	x			x	x	x		Tracking	x	x			
2003-Spengler (Spengler and Schiele, 2003)	GMM	x	x	x	x	x		x	x			Tracking	x	x			
2004-Aubert (Aubert et al., 2004)	None											DFC			x		
2004-Fujiyoshi (Fujiyoshi and Kanade, 2004)	Basic			x				x	x			Persistence		x			
2004-Yang (Yang et al., 2004)	Basic			x				x	x			DFC			x		x
2005-Black (Black et al., 2005)	MM	x		x				x	x		x	Tracking	x	x			
2005-Mathew (Mathew et al., 2005)	GMM	x	x	x	x	x		x	x			GS			x		
2005-Melli (Melli et al., 2005)	MM	x		x				x	x	x	x	Tracking	x	x			
2005-Tian (Tian et al., 2005)	GMM	x	x	x	x			x	x			GS			x	x	
2005-Velastin (Velastin et al., 2005)	MM	x		x				x	x		x	Tracking	x	x			
2006-Auvinet (Auvinet et al., 2006)	MM	x		x				x	x		x	Tracking	x	x			
2006-Bird (Bird et al., 2006)	GMM	x	x	x	x	x		x	x			Tracking	x	x			
2006-Ferrando (Ferrando et al., 2006)	Basic			x				x	x			Tracking	x	x		x	
2006-Guler (Guler and Farrow, 2006)	GMM	x	x	x	x	x		x	x	x		Persistence		x			
2006-Li (Li et al., 2006)	None											Tracking	x	x			
2006-Lim (Lim and Davis, 2006)	OF	x	x	x	x		x	x	x			Other	x	x			x
2006-Lv (Lv et al., 2006)	–											Tracking	x	x			
2006-Martínez (Martínez-del Rincón et al., 2006)	MM	x		x				x	x		x	Tracking	x	x			
2006-Smith (Smith et al., 2006)	GMM	x	x	x	x	x		x	x			Tracking	x	x			
2006-Spagnolo (Spagnolo et al., 2006)	SGM	x	x	x	x			x	x			Tracking	x	x		x	
2007-Arsic (Arsic et al., 2007)	GMM	x	x	x	x	x		x	x			Tracking	x	x			
2007-Bevilacqua (Bevilacqua and Vaccari, 2007)	Other	x	x	x	x			x	x		x	Tracking	x	x			
2007-Bhargava (Bhargava et al., 2007)	OF	x	x	x	x		x	x	x			Classifier - Tracking	x	x			x
2007-Cheng (Cheng et al., 2007)	CM	x		x	x	x		x	x			DFC					
2007-Chuang (Chuang et al., 2007)	GMM	x	x	x	x	x		x	x			Tracking	x	x		x	
2007-Dalley (Dalley et al., 2007)	SGM	x	x	x	x		x	x	x			Tracking	x	x			
2007-Denman (Denman et al., 2007)	CM	x	x	x	x	x	x	x	x			Persistence		x	x		
2007-Guler (Guler et al., 2007)	Basic			x				x	x			Tracking	x	x			
2007-Porikli (Porikli, 2007)	GMM	x	x	x	x	x		x	x			DFC			x		
2008-Cho (Cho et al., 2008)	GMM	x	x	x	x	x		x	x			Other					
2008-Gallego (Gallego et al., 2008)	SGM	x	x	x	x			x	x			Persistence		x			
2008-Liao (Liao et al., 2008)	SGM	x	x	x	x			x	x			Persistence		x			
2008-Lin (Lin and Wang, 2008)	GMM	x	x	x	x	x		x	x			DFC - GS			x	x	
2008-Miezianko (Miezianko and Pokrajac, 2008)	SLM	x	x	x			x	x	x			Classifier		x			x
2008-Porikli (Porikli et al., 2008)	GMM	x	x	x	x	x		x	x			DFC			x		
2008-Rodriguez (Rodriguez-Fernandez et al., 2008)	Basic			x				x	x			Persistence		x			
2008-SanMiguel (San Miguel and Martínez, 2008)	SGM	x	x	x	x		x	x	x			Tracking				x	
2008-Tian (Tian et al., 2008)	GMM	x	x	x	x			x	x			GS			x	x	
2009-Chuang (Chuang et al., 2009)	GMM	x	x	x	x	x		x	x			Tracking	x	x			
2009-Lee (Lee et al., 2009)	GMM	x	x	x	x	x		x	x			Tracking	x	x			
2009-Li (Li et al., 2009)	SGM	x	x	x	x			x	x			DFC			x	x	
2009-Magno (Magno et al., 2009)	–											Other				x	

Table 4
Summary of challenges addressed by the reviewed strategies (part 2).

Strategy	BG subtraction method	MO detection										SFO detection method	SFO detection				
		Image noise	Illum. changes	Low contrast	Camera adjust.	Dynamic BG	Camera jitter	FG aperture	Camouflage	SHs & HLs	Bootstraping		Occlusions	Long-term	PSFO	RO	GR
2009-Singh (Singh et al., 2009)	MM	x		x				x	x		x	DFC - Tracking	x				
2009-Wen (Wen et al., 2009)	Combination	x	x	x	x	x		x	x	x	x	Classifier		x		x	x
2010-Bayona (Bayona et al., 2010)	SGM	x	x	x	x			x	x			Persistence	x	x		x	x
2010-Chang (Chang et al., 2010)	Basic			x				x	x			Persistence		x			
2010-Evangelio (Evangelio and Sikora, 2010)	GMM	x	x	x	x	x		x	x			DFC - FSM		x	x		
2010-Kwak (Kwak et al., 2010)	SGM	x	x	x	x		x	x	x	x		Tracking - FSM	x	x			
2010-Li (Li et al., 2010)	GMM	x	x	x	x	x		x	x			DFC - Classifiers		x	x		x
2010-Singh (Singh et al., 2010)	GMM	x	x	x	x	x		x	x	x		Tracking	x	x			
2010-Wang (Wang and Liu, 2010)	GMM	x	x	x	x	x		x	x			DFC			x	x	
2011-Albiol (Albiol et al., 2011)	None											Classifier		x			x
2011-Evangelio (Evangelio and Sikora, 2011)	GMM	x	x	x	x	x		x	x			DFC - FSM		x	x	x	
2011-Fan (Fan and Pankanti, 2011)	GMM	x	x	x	x		x	x	x			GS - FSM		x		x	
2011-Fu (Fu et al., 2011)	GMM	x	x	x	x	x		x	x			Persistence - Tracking	x	x			
2011-Geng (Geng and Xiao, 2011)	CM	x		x	x	x		x	x			Other				x	
2011-Pan (Pan et al., 2011)	Other	x	x	x	x		x	x	x	x		Persistence		x	x		
2011-Raheja (Raheja et al., 2011)	GMM	x	x	x	x	x		x	x			GS			x	x	
2011-Tian (Tian et al., 2011)	GMM	x	x	x	x			x	x			GS			x	x	
2011-Yang (Yang and Rothkrantz, 2011)	CM	x		x	x	x	x	x	x			Classifier - Tracking	x	x			x
2011-Zin (Zin et al., 2011)	Basic		x	x				x	x	x		DFC			x		
2012-Bangare (Bangare et al., 2012)	GMM	x	x	x	x	x		x	x			Tracking	x	x			
2012-Fan (Fan and Pankanti, 2012)	GMM	x	x	x	x	x		x	x			GS - FSM	x	x			x
2012-Kim (Kim et al., 2012)	MM	x		x				x	x	x	x	Classifier		x		x	x
2012-Lai (Lai et al., 2012)	Other	x		x				x	x			DFC			x	x	
2012-Morde (Morde et al., 2012)	Other	x	x	x	x	x	x	x	x	x		Persistence		x			
2012-Xiya (Xiya et al., 2012)	Basic			x				x	x			DFC - Tracking	x				
2012-Zin (Zin et al., 2012)	Basic			x				x	x	x		DFC			x	x	
2013-Chang (Chang et al., 2013)	SLM	x	x	x			x	x	x			Classifier - Tracking	x	x			x
2013-Collazos (Collazos et al., 2013)	Basic		x	x	x			x	x	x		Persistence - FSM		x			x
2013-Fan (Fan et al., 2013)	GMM	x	x	x	x		x	x	x			GS - FSM	x	x			x
2013-Ferryman (Ferryman et al., 2013)	GMM	x	x	x	x	x		x	x			DFC - Classifiers		x			x
2013-Hassan (Hassan et al., 2013)	GMM	x	x	x	x	x		x	x			Persistence		x		x	x
2013-Maddalena (Maddalena and Petrosino, 2013)	NN	x		x	x	x	x	x	x	x		Persistence		x	x	x	x
2013-Muchtar (Muchtar et al., 2013)	GMM	x	x	x	x	x		x	x			DFC - Persistence		x		x	
2013-Ortego (Ortego and SanMiguel, 2013)	SGM	x	x	x	x		x	x	x			Persistence		x	x		
2013-Sajith (Sajith and Nair, 2013)	CM	x		x	x	x		x	x			DFC - Tracking	x			x	
2013-Tripathi (Tripathi et al., 2013)	SGM	x	x	x	x		x	x				Persistence		x			
2014-Joglekar (Joglekar et al., 2014)	GMM	x	x	x	x	x		x	x			DFC - Tracking	x				
2014-Kim (Kim and Kang, 2014)	None											Classifiers - FSM		x		x	x
2014-Lopez (Lopez-Mendez et al., 2014)	Other	x	x	x	x	x	x	x	x	x		GS					
2014-Ortego (Ortego and SanMiguel, 2014)	SGM	x	x	x	x		x	x	x			Classifier	x	x	x		x
2014-Wang(Wang et al., 2014)	Combination	x	x	x	x	x	x	x	x			DFC				x	
2015-Baxter (Baxter et al., 2015)	Basic			x				x	x			Persistence		x		x	
2015-Ingersoll (Ingersoll et al., 2015)	GMM	x	x	x	x	x		x	x			Tracking	x	x			
2015-Jardim (Jardim et al., 2015)	SLM	x		x				x	x			Classifier		x			x
2015-Lin (Lin et al., 2015)	GMM	x	x	x	x	x		x	x			DFC - FSM - Tracking	x	x	x	x	
2015-Mahin (Mahin et al., 2015)	Basic			x				x	x			Tracking	x	x			
2015-Martínez (Martínez et al., 2015)	KDE	x	x	x	x	x		x	x		x	DFC - FSM	x	x	x		
2015-Wahyono (Filonenko et al., 2015)	Basic			x				x	x			DFC - Tracking	x			x	
2015-Zeng (Zeng et al., 2015)	Combination	x	x	x	x	x		x	x			DFC			x		

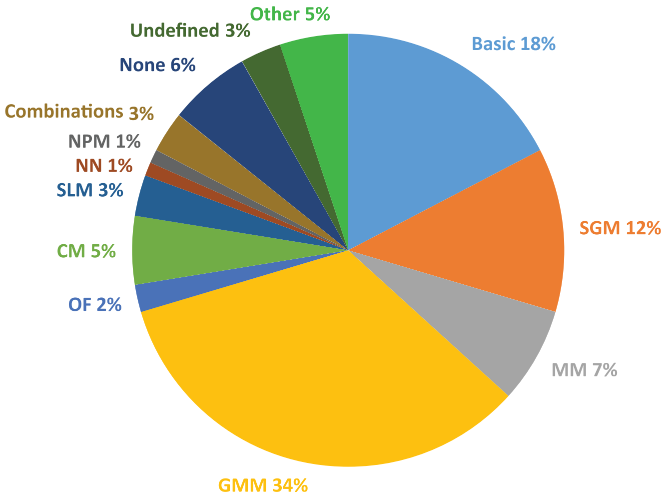

Fig. 3. Percentages of SFO detection strategies according to the BG subtraction method they use.

5.1. Basic

This category comprises those subtraction methods not using any statistical model: the initial BG model is obtained from one or multiple frames without FG, typically at the beginning of the sequence, that are manually selected. These methods can be considered the most simple ones, since they assume that, at each instant, the BG is formed by fixed values (neither the noise of the pixels or the BGs with dynamic areas are taken into account).

Often, the reason for using so inflexible models is motivated by the simplicity of the strategies and their computational efficiency (Fujiyoshi and Kanade, 2004). It is also common to find this type of decisions in works involving stages that reach beyond the detection of SFOs and, therefore, they are not interested in putting effort in the FG detection. For example, Stringa and Regazzoni (2000) describes a system that, first, detects the presence of abandoned objects and then performs a temporal video segmentation to facilitate to human operators the task of retrieving the cause of alarms. In Foresti et al. (2002), a strategy for video shot detection and layered indexing was proposed, which is able of detecting potentially dangerous situations (abandoned objects and people with suspicious behavior) that must be notified to human operators in guarding environments. In Baxter et al. (2015) several methods to deal with many complex human activities are proposed.

The selection of the reference frames used to construct the BG model can be done attending to different criteria. In Wang and Ooi (1999), a frame without moving objects is manually selected as BG, whereas in Chang et al. (2010) the BG image is obtained as the average of some manually selected frames in which the presence of FG objects is minimum.

In the simplest cases (Mahin et al., 2015; Wang and Ooi, 1999) the initial BG is never updated and it is used to be compared with all the frames along the sequence. In other cases (Ferrando et al., 2006), to deal with permanent changes in the scene, it is periodically refreshed. This update can be complete (all the pixels in the BG model are refreshed simultaneously) (Sacchi and Regazzoni, 2000) or partial (only some pixels are updated) (Guler et al., 2007; Rodriguez-Fernandez et al., 2008). In Sacchi and Regazzoni (2000) the full BG image is updated when a significant change in the scene occurs (e.g. a noticeable variation in light). In Guler et al. (2007) the BG model is initialized with the first frame of the sequence and it is periodically updated with the values of those pixels that are not part of MOs or candidate SFOs. The BG model used in the strategy in Rodriguez-Fernandez et al. (2008) updates the pixels remaining unchanged for a given number of consecutive frames. Other works (Yang et al., 2004) only explain their update mechanism, but they not specify how the model is initialized.

Whereas most approaches use a unique BG model, some strategies propose constructing two models (Zin et al., 2011), which are statistically updated at different speeds (Xiya et al., 2012) or apply two different selective BG updates (Filonenko et al., 2015). These models are commonly called short-term model and long-term model. The way in which these two models are used to detect SFOs is detailed in Section 6.

All the strategies in this category are able to address the most basic challenges: low contrast, FG aperture and camouflage. Additionally, some of them are able to deal with other specific challenges. For example, the strategy in Sacchi and Regazzoni (2000) includes mechanisms to be robust to illumination changes, the strategy in Foresti et al. (2002) avoids false detections due to camera jitter, and the strategy in Zin et al. (2011) includes a post-processing stage to remove shadows from the detections and provide robustness against lighting changes.

5.2. Single Gaussian models (SGM)

To be robust against the typical noise of the camera sensor, gradual illumination changes and camera automatic adjustments, some strategies try to statistically model the variations of each pixel with a Gaussian distribution, which is updated at each instant to try to adapt to the BG changes. The FG pixels will be those whose Mahalanobis distance to the Gaussians is greater that a predefined threshold (Gallego et al., 2013).

Some SFO detection strategies take as starting point the popular Running Gaussian Average (RGA) method proposed in Wren et al. (1997). Some of these approaches (Gallego et al., 2008) directly adopt this method, whereas other strategies complement it with additional stages to improve the FG detection in some situations. In Dalley et al. (2007) the RGA modeling is combined with Markov Random Fields (MRFs), which spatially smooths the BG modeling results, thus avoiding the need of applying an independent threshold to each pixel and improving the quality of the results in sequences with camera jitter. In Li et al. (2009) the FG masks are constructed from two BG models: the first one (short-term model) is obtained with the RGA method, whereas the second one (long-term model) is an approximation of the average of the previous.

Other strategies (Beynon et al., 2003; Kwak et al., 2010), instead of using the RGA method, take as starting point the algorithm proposed in Horprasert et al. (1999). This algorithm models the brightness and the chromaticity of the pixels using Gaussian distributions. In this way, it successfully avoids many false detections due to shadows and highlights cast by the MOs.

Another single Gaussian-based modeling strategy that has been used by several authors (Bayona et al., 2010; Ortego and San-Miguel, 2014; San Miguel and Martínez, 2008) is the one proposed in Cavallaro and Ebrahimi (2000), which is able to deal with complex camera noise and a large variety of slow environmental light changes. Additionally, by modeling not at pixel level but in a spatial neighborhood around each pixel, it provides robust detections in a wider variety of situations (e.g. camera jitter).

5.3. Median models (MM)

It is also possible to find a significant amount of SFO detection algorithms using median filtering to model the BG (Black et al., 2005; Martínez-del Rincón et al., 2006). These methods define the BG to be the median at each pixel location of all the frames in a buffer (Elhabian et al., 2008), assuming that the pixels stay in the BG for more than half of such frames. The major strengths of the median-based modeling are its computational efficiency, robustness to noise and efficiency. Additionally, these methods are also able to provide successful detections in bootstrapping sequences. However, their most noticeably limitation is that they do not model the variance of the pixels. Consequently, they are not able of dealing with illumination changes or camera automatic adjustments.

In Velastin et al. (2005) a median-based BG model at block-level is performed, which is selectively updated from motion information of the scene.

The median filtering proposed in Cucchiara et al. (2003), which is able to classify among FG, shadows, and subtracted objects, has been adopted by some authors (Kim et al., 2012; Melli et al., 2005) to detect SFOs. In the case of Melli et al. (2005), the original method was modified to selectively exclude stopped vehicles from the BG updating.

In Singh et al. (2009), two models are obtained and combined to detect SFOs, both of them constructed with the Approximate Median Model described in McFarlane and Schofield (1995). One of the models is updated frequently and it detects only the MOs. The second model has a slower update rate and detects also the SFOs.

5.4. Gaussian mixture models (GMM)

The BG of the scene usually contains non-static regions (e.g. tree branches, bushes, water surface or flags). These BG changes can not be modeled using one Gaussian distribution per pixel. Instead, a generalization based on a mixture of Gaussians can be used to model such variations. In Stauffer and Grimson (1999), Stauffer and Grimson proposed a Gaussian Mixture Model (GMM) to deal with scenarios with dynamic BGs. This method allows the BG model to be a mixture of several Gaussians (typically between 3 and 5). Each current pixel is compared against the existing Gaussians. The parameters of the matched model are updated based on a learning factor. If there is not a match, the least-likely model is discarded and replaced by a new Gaussian distribution initialized with the value of the current pixel. The current pixels that are correctly represented by a predefined portion of the GMM are classified as BG and the rest of pixels are classified as FG.

As it can be observed in Fig. 3, most of the strategies for detecting SFOs use a GMM to subtract the BG (34%), since these models typically allow dealing with a large amount of challenges in MO detection: image noise, illumination changes, low contrast, camera automatic adjustments, dynamic background, FG aperture and camouflage.

Many of these approaches (Hassan et al., 2013; Raheja et al., 2011; Smith et al., 2006) directly use the original method proposed by Stauffer and Grimson, not applying any substantial modification to the original method and not specifying any parameter configuration. Other approaches propose the use of a specific number of Gaussian distributions (Spengler and Schiele, 2003) or alternative initialization processes (Arsic et al., 2007). Moreover, some approaches proposing significant changes to the original GMM method can be found, which try to improve the quality of the results or reduce the computational efficiency of the algorithms. The authors of Singh et al. (2010) propose an alternative measure to order the Gaussians. In this way, they are able of reducing the computational cost in the modeling. In Fu et al. (2011) an intermittent updating scheme is proposed, which only updates the model when the surrounding environment of the pixel is severely changing. Doing this, the inclusion of the SFOs in the BG is delayed. The strategies in Guler and Farrow (2006) and (Singh et al., 2010) include mechanisms to avoid shadows from the detections.

Other strategies (Cho et al., 2008), instead of directly using or modifying the original GMM method, use modifications developed by other authors. For example, the BG subtraction described in Bird et al. (2006) uses the GMM-based method proposed in Atev et al. (2004), which instead of using Gaussians with diagonal covariance matrices uses complete covariance matrices. Additionally, in the interest of real-time operation, the authors of Bird et al. (2006) propose processing only every tenth frame.

Similarly to the cases of the basic and the median models discussed in the previous subsections, there are also some GMM-based strategies that construct two models with different learning factors (i.e. different BG refreshing rates) (Evangelio and Sikora, 2010; Ferryman et al., 2013; Joglekar et al., 2014; Lin et al., 2015). The way in which the SFOs are detected from these two models is discussed in detail in Section 6.

Among GMM-based strategies, the strategy in Lin and Wang (2008) is the only one proposing two alternatives to detect the FG objects. The first one takes as starting point the strategy in Tian et al. (2005), which uses a single GMM. The second one is based on the comparison of two GMM as proposed in Porikli (2007). According to the authors, the first option must be used to achieve high efficiency, whereas the second one must be selected to obtain the highest quality.

Finally, there are some works that do not use the GMMs to adapt the complex variations of the pixel values (Fan and Pankanti, 2011; Tian et al., 2011), but use the GMMs to detect SFOs by analyzing the stability of the different Gaussians in the mixture model (see Section 6.5). Therefore, these works are not able to deal with dynamic BG.

5.5. Other BG subtraction methods

The BG modeling methods described in the previous subsections have been used by the 71% of the SFO detection strategies. However, the remaining 29% uses other types of modeling, which are briefly described along this subsection.

5.5.1. Optical flow (OF)

Motion segmentation using optical flow is based on the analysis of coherent motion of points or features between frames to detect regions or pixels that have changed (Wang et al., 2003). In this way, these methods are usually able of dealing with illumination changes and camera jitter. The authors of Lim and Davis (2006) propose a novel strong discriminative measure to identify BG pixels, which is based on optical flow information acquired at pixel level. In Bhargava et al. (2007), the initialization algorithm in Gutchess et al. (2001) is adapted to build the BG model, which is based on a local optical flow analysis that helps to determine which temporal interval is most likely to display the BG.

5.5.2. Cluster models (CM)

Cluster models suppose that each pixel in the frame can be temporally represented by clusters and they are able to deal with dynamic BGs and noise from video compression (Bouwmans, 2014). In Denman et al. (2007), a clustering modeling using color and motion information is proposed, which is robust to lighting changes. In Cheng et al. (2007), the CM described in Luo and Bhandarkar (2005) is used. The SFO detection strategy in Geng and Xiao (2011) proposes a BG modeling using a modified codebook (the clusters are formed by codewords including appearance information of the pixels) with two layers for detecting, respectively, MOs and SFOs. The BG model proposed in Yang and Rothkrantz (2011) is initialized along the first 100 frames and the subsequent frames are segmented using the codebook method in Kim et al. (2005). The authors of Sajith and Nair (2013) also take as starting point the codebook algorithm proposed in Kim et al. (2005), but in this case they use it to obtain two BG models with different learning rates that will allow discriminating between MOs and SFOs.

5.5.3. Subspace learning models (SLM)

Some of the strategies proposed to detect SFOs include a BG modeling stage based on learning models. These methods are used to model the BG in the idea to represent inline data content while reducing dimension significantly (Bouwmans, 2009). The most common SLMs used in BG subtraction are the following: Principal Component Analysis (PCA) (Jolliffe, 2002), Independent Component Analysis (ICA) (Hyvärinen et al., 2004) and Non-negative Matrix Factorization (NMF) (Lee and Seung, 2001). The strategy in Miezianko and Pokrajac (2008) uses the BG modeling in Latecki and Miezianko (2006), which is based on an incremental PCA decomposition over non-overlapping texture blocks that is to robust image noise, illumination changes and camera jitter. In Jardim et al. (2015), a Robust PCA (RPCA) is used to explore the low-rank similarities between reference and target videos. Finally, the BG subtraction in Chang et al. (2013) is able to deal with image noise and illumination changes by using a robust block-based boosting learning method and a classifier that makes use of Haar-like features (Viola and Jones, 2001).

5.5.4. Neural networks (NN)

The BG model is represented by means of the weights of a neural network suitably trained on several clean frames. The network learns how to classify each pixel as BG or FG. These methods have shown to be able of dealing with most typical challenges in FG detection (Maddalena and Petrosino, 2012). However, as they depend on a training period, they fail when abrupt changes occur in the scene (e.g. abrupt light changes) and in bootstrapping sequences. In Maddalena and Petrosino (2013), a 3-D neural model that automatically adapts to scene changes in a self-organizing manner is proposed, which is used to model both the BG and the FG of the scene.

5.5.5. Kernel density estimation (KDE)

Kernel density estimation methods, also known as nonparametric modeling methods, estimate the density function directly from the data without any assumptions about the underlying distribution, avoiding having to choose a model and estimating its distribution parameters (Elhabian et al., 2008). Consequently, they are able to provide high-quality detections in a large variety of complex situations, such as dynamic BG, illumination changes and bootstrapping sequences. In Martínez et al. (2015) a nonparametric-based BG modeling is proposed, which simultaneously obtains three BG models that are selectively updated using different criteria.

5.5.6. Combinations

Some approaches use more than one of the previously described BG modeling strategies. The authors of Wen et al. (2009) propose the use of one of three possible BG models (a model based on a texture analysis (Heikkilä and Pietikäinen, 2006), a model based on an incremental PCA algorithm (Zhao et al., 2008) and a KDE-based modeling (Elgammal et al., 2002)), together with different shadow removal techniques, according to the characteristics of the analyzed sequence. In Wang et al. (2014), an optical flow-based analysis is combined with a GMM-based modeling that model both the BG and the FG and automatically adapts the amount of Gaussians required by each pixel. Finally, in Zeng et al. (2015) two BG models are constructed simultaneously: the first one is based on a three frame differencing, whereas the second one is an improved version of the classical GMM.

5.5.7. None

Some approaches (Gibbins et al., 1996; Murino, 1998) do not model the BG but instead they extract sets of discriminative features of the objects of interest (learned beforehand through machine learning algorithms), that are used to detect such objects in subsequent stages. The strategy in Aubert et al. (2004) proposes the use of two algorithms to detect BG changes. The first one is based on level sets and allows detecting general changes. The second one makes use of the first, at different time scales, to discriminate between BG, MOs and SFOs. In Li et al. (2006), the BG appearance is characterized using three types of features: spectral, spatial and temporal. The strategy in Albiol et al. (2011) is based on the detection of corner points using the Harris algorithm (Harris and Stephens, 1988). In Kim and Kang (2014), several features are introduced in a Finite State Machine (FSM) to classify each pixel among BG, MO and SFO.

5.5.8. Unspecified

Some of the reviewed strategies (Amer et al., 2002; Lv et al., 2006; Magno et al., 2009) stay that they use a BG subtraction algorithm as a first stage. However, details on the used subtraction method are not specified (neither the initialization nor the update). These works have been labeled with a hyphen in the third column of Tables 1 and 2.

5.5.9. Other

This category comprises those works using BG subtraction algorithms that can not be associated to any typical category of methods. The strategy in Bevilacqua and Vaccari (2007) uses the coarse-to-fine strategy (a block-based stage followed by a pixel-level stage) proposed in Bevilacqua et al. (2005). A hybrid differencing-based strategy at region level is proposed in Pan et al. (2011). The authors of Lai et al. (2012) propose using a timeliness BG that uses real world time instead of the frames. The BG modeling strategy in Morde et al. (2012), avoids making assumptions of normality pixel distributions and uses the Chebyschev probability inequality. The strategy in Lopez-Mendez et al. (2014) includes a multilayer BG subtraction algorithm that takes as starting point the texture-based BG modeling proposed in Yao and Odobez (2007).

6. SFO detection

Once the FG has been detected (through the application of any of the methods described in Section 5), it is necessary to discriminate between MOs and SFOs. To do this, many detection algorithms of different complexities have been proposed in the literature, which are reviewed along this section. Some authors have opted by algorithms that directly analyze the results provided by the BG subtraction stage. On the other hand, some strategies include image analysis stages specifically oriented to the detection of SFOs. There are also some approaches that deal with other typical challenges, such as the detection of long-term SFOs, PSFOs or GRs. Besides, a significant amount of authors include an additional stage to detect ROs (the algorithms that are typically used in this additional stage are presented in Section 7).

As discussed in Section 5, the BG subtraction is carried out at pixel level by most strategies. However, in the process of detecting SFOs many more strategies (64%) work at region or object level, since in this way they can discard small noisy detections resulting from the BG subtraction and, moreover, they are able to group FG regions that actually belong to the same object (this typically occurs when a portion of a FG object has similar appearance to the BG region that it is covering). On the other hand, only those methods working at pixel level are able to deal with the detection of SFOs of type C (i.e., PSFOs), since people who stop walking are neither static or in motion (i.e. they have both moving pixels and static pixels). It must be noted that part of the strategies working at pixel level are not focused on the detection of static people and, although they are able of detecting PSFOs, they discard such objects in post-processing stages.

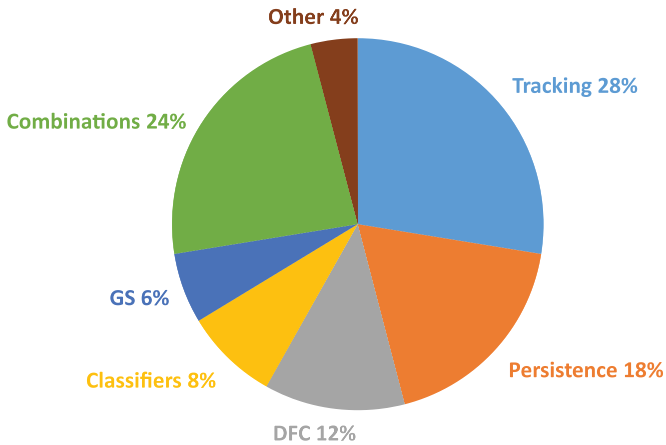

Fig. 4. Percentages of strategies according to the method they use to detect SFOs.

The SFO detection approaches can be classified in the seven categories that are presented in the following subsections. The fifth column in Tables 1 and 2 summarizes the categories corresponding to each of the reviewed strategies, whereas the sixth column in such tables shows the detection level (pixel or object) that is considered by each strategy. Fig. 4 illustrates the percentages of strategies encompassed within each category.

To facilitate comparing the detection methods among them, and also to easily observe the strengths and weaknesses of the algorithms that are introduced along the following subsections, a summary of the challenges addressed by all these algorithms has been included in the right half of the Tables 3 and 4. As stated in Section 5, the data in these tables refer to the challenges that each method is supposedly capable of coping with.

6.1. Tracking

A very large amount of the strategies for detecting SFOs (28%) include a tracking algorithm to determine what FG regions become static. These algorithms allow detecting both short-term and long-term SFOs and, additionally, they deal with occluded SFOs. However, since they work at object level, they are not able of dealing with sequences containing PSFOs.

As it can be observed in Tables 1 and 2, most of these strategies were proposed along the first reviewed years. Nevertheless, in the recent years most authors have opted for other kinds of detection processes, since they are typically able to provide more consistent results in a larger amount of situations.

On the one hand, some tracking algorithms are just focused on determining when a FG object stops moving. The simplest ones try to relate objects between pairs of consecutive images by analyzing colors, distances, velocities, or object sizes (Lee et al., 2009; Mahin et al., 2015). Other algorithms include more complex strategies that are based on higher-level information, such as the peripheral tracker proposed in Guler and Farrow (2006), which is inspired by human's visual cognition processes. Additionally, many approaches use standard tracking algorithms (e.g. a Kalman-based tracking in Melli et al. (2005) and Ingersoll et al. (2015), a pyramidal Kanade-Lucas-Tomasi algorithm in Bevilacqua and Vaccari (2007) or particle filters in Spengler and Schiele (2003) and (Smith et al., 2006)). To assure that FG object is a SFO, many of these algorithms establish that the object should remain static for a minimum time (e.g. 25 s in Arsic et al. (2007) or 3 s in Auvinet et al. (2006)) or along some consecutive frames (e.g. 50 frames in San Miguel and Martínez (2008)).

On the other hand, other tracking strategies not only try to identify when a FG object stops moving, but analyze the link between abandoned objects and the people who left them (Dalley et al., 2007; Ferrando et al., 2006; Foresti et al., 2002; Lv et al., 2006). In this way, they are able to determine if a SFO can result in a dangerous situation (someone leaves an object and moves away from it) or not (someone leaves an object but stays near to it).

In addition to the above mentioned strategies, some multi-camera tracking strategies can also be found (Beynon et al., 2003; Martínez-del Rincón et al., 2006), which are able to provide consistent results in a larger amount of situations (e.g. in presence of complex occlusions).

It is also possible to find some strategies analyzing the motion at block level (Black et al., 2005; Velastin et al., 2005).

Finally, it should be noted that some of these tracking-based strategies include additional stages to detect ROs. Some of them are based in the analysis of edges (Spagnolo et al., 2006) or histograms (Chuang et al., 2007), whereas other works analyze multiple types of information (e.g., shape, contours and color in San Miguel and Martínez (2008)). A more detailed discussion on the typical mechanisms to detect ROs and separate them from the SFOs is provided in Section 7.

6.2. Persistence

The detection process in some strategies consists in an analysis of the persistence of the pixels that have been previously classified as FG. That is, one pixel is considered part of a SFO if it is classified as FG for a predetermined time or along several frames (consecutive or not). As it is shown in Tables 3 and 4, these methods are able of dealing with most challenges in SFO detection. The only challenge that they are not able to cope with is that of occluded SFOs.

In (Denman et al., 2007) a counter is associated to each pixel. When such counter reaches a threshold the pixel is considered to be part of a SFO. This strategy allows detecting short-term and long-term SFOs. Moreover, since it works at pixel level, it also allows detecting PSFOs. However, it does not include any stage to detect ROs or GRs.

In Liao et al. (2008) the SFOs are detected by analyzing the persistence in 6 FG masks obtained from 6 frames equispaced along the 30 previous seconds. This method, despite its simplicity, is sufficient to detect the most typical cases of abandoned objects. In Bayona et al. (2010), this idea was used as starting point, but some additional stages were added to deal with the detection of stolen objects, some occlusions and false detections (e.g., GRs).

The authors of Chang et al. (2010) assume that a FG object is a SFO when it remains static T consecutive seconds. After, they apply a color-based analysis to discard stationary people from the results.

Other authors, instead of choosing a fixed number of frames, just mention that this number must be set according to the desired responsiveness of the system. This is the case of the strategy in Maddalena and Petrosino (2013). This strategy includes several post-processing stages for determining when a SFO resumes motion and to deal with occluded SFOs.

Some strategies trying to reduce the dependency of the results with fixed counters that must be set by the users have also been proposed. The authors of Pan et al. (2011) use a score function that is updated, at each new frame, from the BG subtraction results. A similar update process is applied in Guler and Farrow (2006), where the final decisions depend on a stationary object confidence image. The strategy in Pan et al. (2011) is able of detecting PSFOs. However, since it is focused in the detection of abandoned objects, it includes a stage to discard these objects from the results. On the other hand, this strategy does not consider the detection of neither ROs nor GRs. The strategy in Guler and Farrow (2006) is also focused in the detection of abandoned objects. To carry out this task, it tries to detect when a FG object splits in two parts:

a MO and a SFO. This strategy does not address the detection of PSFOs, ROs, or GRs.

To improve the quality of the results in some complex situations, some authors add other analysis stages after applying the persistence analysis. In Gallego et al. (2008), the SFOs are modeled using a Gaussian distribution, which improves the quality of the results when the color features of the SFOs change slightly. In Hassan et al. (2013), an edge-based tracking is applied to remove false FG detections (mainly due to illumination changes), ROs and GRs.

Instead of analyzing the persistence of the pixels that are part of the FG, some algorithms focus on analyzing the persistence of the appearance (color, gradients, etc.) of those pixels. That is, if a pixel is classified as FG, its appearance is analyzed along the subsequent frames. If such appearance persists for a minimum time, it is established that the pixel is part of a SFO. In some cases (Sacchi and Regazzoni, 2000; Stringa and Regazzoni, 2000; Tripathi et al., 2013), this persistence is analyzed only between two consecutive frames. Whereas in other cases (Fujiyoshi and Kanade, 2004; Wang and Ooi, 1999), the appearance must persist along more frames. These strategies are able of detecting SFOs in more scenarios than the algorithms not using appearance data. However, none of them include mechanisms to detect PSFOs, ROs or GRs.

Similarly to the tracking-based detection strategies, there are some strategies using persistence to cope with the detection of ROs (Hassan et al., 2013; Maddalena and Petrosino, 2013) (see Section 7).

6.3. Dual FG comparison (DFC)

Taking as starting point the strategy proposed by Porikli in Porikli (2007), a large amount of strategies try to identify the SFOs by comparing two binary FG masks at pixel level. These masks are obtained from two BG models constructed with different learning rates. In the strategy proposed by Porikli, and some other later works (Joglekar et al., 2014; Li et al., 2010; Wang and Liu, 2010), the models are constructed using multiple Gaussians. However, other modeling choices can also be found in the literature: non-statistical models (basic) in Zin et al. (2011), Xiya et al. (2012), Zin et al. (2012) and Filonenko et al. (2015), SGMs in Li et al. (2009), MMs in Singh et al. (2009), or CMs in Cheng et al. (2007).

The BG model with the highest learning rate (commonly named short-term model) must be configured to adapt rapidly to the changes in the scene. So, it will only detect the short duration changes (i.e. the MOs). In contrast, the model with the lowest learning rate (long-term model) must be configured to be more resistant against the changes. So, it will also detect the large duration changes (i.e. the SFOs). Let $F_S(h, w)$ and $F_L(h, w)$ be the binary FG masks obtained from, respectively, the short-term and the long-term models, where (h, w) represents the coordinates of the pixels (rows and columns). Depending on the values of these masks, the pixels are classified in one of the following 4 classes:

- BG: If $(F_S(h,w), F_L(h,w)) = (0,0)$.
- MO: If $(F_S(h,w), F_L(h,w)) = (1,1)$.
- SFO: If $(F_S(h,w), F_L(h,w)) = (0,1)$.
- Uncovered BG: If $(F_S(h,w), F_L(h,w)) = (1,0)$.

These methods are commonly able of providing successful detections in scenarios with complex backgrounds (Joglekar et al., 2014; Zeng et al., 2015). Additionally, many of them allow detecting PSFOs (Aubert et al., 2004; Lai et al., 2012; Wang and Liu, 2010; Zin et al., 2012). Moreover, since ROs give rise to uncovered BG, these strategies have the capability of correctly detecting these kind of objects. However, only some of the methods using DFCs focus on separating abandoned objects from the removed ones (Lai et al., 2012; Wang et al., 2014; Wang and Liu, 2010).

Regarding the GRs, some authors (Porikli, 2007) claim that DFC-based methods are suitable to avoid GRs form the detections. However, this is true only if the learning rate of the long-term model is high enough compared to the time that an object remains stationary before returning to move.

On the other hand, they are not able of maintaining the detections when the FG objects remain static for a long time (long-term SFO challenge), since, sooner or later, the objects remaining stopped long periods of time are absorbed by both long-term and short-term BG models. Additionally, they lose the detected SFOs when other objects pass in fort of them (occlusions challenge). In addition, to provide successful results, the configurations of the long-term and short-term models must be adapted to the characteristics of each analyzed sequence. Thus, the usability of these methods is low, which is an important drawback.

Although most strategies using DFC takes as starting point the idea proposed by Porikli, some alternative analyses have also been proposed. For example, the strategy in Yang et al. (2004) constructs an additional BG model that is selectively updated, with the values of an initial BG model, only for those pixels classified as FG at each instant. In this way, the additional model allows detecting if a SFO moves again. This strategy is able of dealing with PSFOs and GRs. In Aubert et al. (2004), instead of using FG masks resulting from BG models, two feature-based change detectors (short term and long term) are compared to detect SFOs. This strategy also deals with the detection of PSFOs. Finally, the strategy in Cheng et al. (2007) combines two GMM-based detectors to detect SFOs in simple scenarios. However, it is not able to cope with any of the challenges described in Section 2.2.

As it can be seen in Tables 3 and 4, many DFC-based methods are able to cope with many challenges in MO detection (Section 2.1). However, in the case of the challenges related to SFO detection (Section 2.2), these methods usually deal with less challenges than tracking-based and persistence-based methods.

6.4. Classifiers

A significant portion of the reviewed strategies use classifiers to separate the SFOs from the rest of detected FG objects. These strategies are able of detecting SFOs in a large variety of scenarios and regardless the time that such objects remain static (i.e., they cope with the long-term SFO challenge). Additionally, they also deal with GRs. An other important advantage of the strategies using classifiers is that they are probably those that provide the best results in crowd sequences.

On the other hand, their main drawbacks are that they require a training period and that, typically, they do not cope with occluded SFOs.

Some of the proposed classifiers are based on the analysis of the physical properties of the objects (e.g. shape or size) (Bhargava et al., 2007). Other classifiers analyze the spatio-temporal variations of their features (e.g. color, edges, etc.) (Ortego and SanMiguel, 2014; Wen et al., 2009).

In some cases the classifiers are used to discriminate between only two classes. For example, the authors of Yang and Rothkrantz (2011) use the skin color detector proposed in Hsu et al. (2002) to classify the FG objects into human or luggage. However, there are other works that consider more classes. For example, the classifier in Kim et al. (2012), which is based on area, saliency and motion information, is able to classify among MOs, ROs, shadows cast by MOs or shadows cast by ROs.

Whereas most classifiers try to identify any kind of SFO (Murino, 1998), some authors have proposed classifiers to detect specific SFOs, such as the one proposed in Miezianko and Pokrajac (2008).

There are also some of these strategies that allow detecting ROs. For example, in Kim et al. (2012) the ROs are detected thanks to a post-processing stage based on the analysis of the edges of the detected SFOs (see Section 7).

6.5. Gaussian stability (GS)

Some of the reviewed works determine that a pixel is part of a SFO by analyzing the stability of the Gaussians in a GMM associated to such pixel. When a MO appears in a pixel, a new Gaussian is created in its GMM, which represents the new value of the pixel. If the object stops moving, that new Gaussian will begin to gain importance in the mixture model. So, if one is able to identify this situation, it will be possible to determine when the MOs become SFOs.

This idea was first proposed in Mathew et al. (2005) and, virtually simultaneously, also in Tian et al. (2005) (with small differences between them). Later, it has been incorporated in other strategies (Lopez-Mendez et al., 2014; Raheja et al., 2011; Tian et al., 2011; 2008).

GS-based detection methods are computationally efficient and easy of implement. Additionally, they depend on few parameters, which increases their usability. Most of these methods perform the detections at pixel level (the only exception in the strategy proposed in Lopez-Mendez et al. (2014), which works at blob level). Therefore, they are suitable for the detection of PSFOs. However, since they use only one BG model, they do not allow dealing with occluded SFOs. Additionally, if a SFO remains static for too long, the Gaussian used to model the SFO will become more important than the Gaussian modeling the BG. Consequently, the long-term SFOs are misdetected.

6.6. Combinations

To improve the quality of the SFO detection in a wide variety of situations, many authors have proposed combining more than one of the previously described strategies.

In many cases, the proposed approaches include post-processing tracking strategies. For example, in Fu et al. (2011), an edge-based persistence analysis is complemented with a tracking strategy. The DFC-based methods in Singh et al. (2009), Xiya et al. (2012), Sajith and Nair (2013), Joglekar et al. (2014) and Filonenko et al. (2015) are also complemented with tracking algorithms. Finally, the strategies in Bhargava et al. (2007), Yang and Rothkrantz (2011) and Chang et al. (2013) propose combining classifiers with tracking modules. In all these works, the main purpose of the tracking step is to improve the quality of the results in sequences with occluded MOs and occluded SFOs. However, since they perform the tracking at object level, none of them is able of detecting PSFOs.

To improve the robustness against the detection of long-term SFOs, other approaches combine DFC-based algorithms with classifiers (Ferryman et al., 2013; Li et al., 2010) or persistence (Muchtar et al., 2013).

The strategy in Lin and Wang (2008), looking for efficiency and robustness, try to detect SFO by the application of two detection strategies in parallel. The first one is a DFC-based method, whereas the second one in based on GS. This strategy is also able of detecting ROs.

Most of the detection strategies described in the previous subsections are based on FG masks obtained from a BG modeling stage. Consequently they are not able of detecting long-term SFOs (objects remaining stopped very large periods of time) since, sooner or later, these objects are always absorbed by the BG models. To avoid this drawback, a very significant amount of authors use traditional SFO detection methods in conjunction with a FSM. In addition, by using FSMs, some works are also able of coping with occluded SFOs. In Evangelio and Sikora (2010), a strategy that uses DFC and a FSM was proposed. An extended description of this strategy was later proposed in Evangelio and Sikora (2011). The approach in Martínez et al. (2015) also uses FSM, but in this case using 3 FG masks. In this way, it is able of dealing with occluded SFOs. In Kim and Kang (2014), a FSM is controlled by several classifiers. In Kwak et al. (2010), the FSM is supplied with the results provided by a tracking module. In Fan and Pankanti (2011), the FSM is used in conjunction to the results of a GS analysis. Improved versions of this work were later published in Fan and Pankanti (2012) and Fan et al. (2013), which are also able of dealing with occluded SFOs and false detections (e.g. GRs). The persistence analysis in Collazos et al. (2013) is also supported by a FSM. Finally, in Lin et al. (2015) a FSM is combined with a DFC to detect candidate SFOs and a tracker is then used to verify whether such candidates are really abandoned or not.

6.7. Other

There are some strategies for detecting SFOs that can not be associated to any of the categories previously described.

The detection strategy in Lim and Davis (2006) is based on a statistical evaluation of the shape and color properties using the Hausdorff distance and a histogram similarity measure. This strategy is able of detecting abandoned objects under severe occlusions and, moreover, its usability is very high.

In Cho et al. (2008) a vertical line scanner is proposed, which is based on the idea that moving people have varied vertical distances to the top of the scene, whereas the SFOs have static vertical distances. As is can be seen in Table 3, this work is not able of dealing with any of the challenges described in Section 2.2.

In Magno et al. (2009) it is assumed that all the detections correspond to SFOs or subtracted objects. So, any specific strategy to discriminate them from the MOs is applied. To discriminate among SFOs and ROs, it applies an edge-based analysis (see Section 7).

Finally, in Geng and Xiao (2011), a modified codebook in proposed, which is able of detecting MOs, SFOs and also ROs. Additionally, the authors state that one of the main advantages of this strategy over previous approaches is its low computational and memory cost.

7. RO detection

The detection of stolen objects (i.e. objects removed from the BG of a scene) is of great interest in some public places, such as museums or stores. When an object is stolen, similarly to the case of the SFOs, a permanent change in the BG of the scene occurs. Consequently, most of the reviewed strategies are able of detecting ROs.

However, a 65% of these strategies do not consider the possible appearance of ROs in the scene. Therefore, they do not include any algorithm to discriminate them from the SFOs. There are also some authors (Geng and Xiao, 2011; Zin et al., 2012) that state that their algorithms are able of detecting both SFOs and ROs, but these works do not include any specific stage to differentiate between them. On the other hand, as it can be seen in Tables 3 and 4, about the 35% of the reviewed strategies apply a last stage with specific mechanisms to cope with the detection of ROs.

Many of these specific mechanisms (Evangelio and Sikora, 2011; Kim et al., 2012; Lai et al., 2012; Magno et al., 2009; Spagnolo et al., 2006; Tian et al., 2005; Wang et al., 2014) rely on the analysis of the edges of both the current image and a BG image (point estimate of the BG that is typically obtained from the BG model). The intuition is that, in many cases, if the BG is covered with a SFO the number of edges in the image will grow. Based on this assumption, a static FG region will be classified as a SFO if the BG image contains less edges than the current frame along such FG region. On the contrary, if the current frame has less edges than the BG image, the FG region will be classified as a RO. Moreover, these strategies are also able of locating false detections (e.g. GRs) under the assumption that the amount of the edges in a false detection must be similar to the amount of edges in the BG image. The strategies including these edge-based algorithms provide successful results in many challenge situations (e.g. camouflage and illumination changes). However, their main drawback is that they require to have available a point estimate of the BG of the scene at every frame and this can become very difficult (or even impossible (Martínez et al., 2015)) for some BG modeling strategies. In addition, obtaining these models is usually computationally expensive.

Other edge-based algorithms (Hassan et al., 2013), instead of comparing the edges in the current and BG images, analyze the edges along successive frames. In this way, they avoid the costly estimation of the BG at each frame. However, these methods are not able of discriminating between ROs and GRs.

Other works try to verify how a static region is compatible with its surroundings. For this purpose, they use region growing processes to grow from the boundaries of the static region to the exterior (Tian et al., 2011; 2008), or apply area analyses on the segmented regions (Sajith and Nair, 2013). Although these strategies achieve high-quality results in many situations, they typically fail in scenarios with SFOs that have similar appearance to the BG (camouflage challenge).

Other approaches (Kim and Kang, 2014; Li et al., 2009; Muchtar et al., 2013; Wang and Liu, 2010) are based on the fact that the color of the static FG regions is usually inconsistent with the one of the BG around it, while the color of the region occupied by the object is often similar with the one of the surrounding region. In this way, they can discriminate between SFOs and ROs by analyzing the color richness of the detected static regions.

Finally, it must be mentioned that there is a portion of strategies that consider the possible appearance of ROs but, nevertheless, they are not interested in their detection. Therefore, they apply some of the previously described algorithms to discard ROs from their final results. For example, in Fan and Pankanti (2011), the region growing analysis proposed in Tian et al. (2008) is used for this purpose, whereas in Raheja et al. (2011) the ROs are discarded by applying an edge-based analysis.

8. Conclusions

This paper provides a complete survey of the most relevant strategies published over the last two decades for detecting SFOs.

First, some general aspects of the reviewed strategies have been analyzed, such as the definitions of SFO they use, the challenges they typically face, and the datasets used to assess the quality of their results.

Second, the most popular algorithms for detecting candidate SFOs and to finally separate the SFOs from the MOs have also been analyzed. Regarding the detection of candidate SFOs, it has been found that most of the strategies make use of BG subtraction methods to perform such detection. Among these subtraction methods, those based on GMM are the most widely used, since these mixture models are able to provide high-quality detections in a wide variety of scenarios (e.g. dynamic BG, illumination changes, camera noise, etc.), while preserving the computational requirements of the applications looking for real-time performance. Regarding the strategies to identify SFOs and separate them from the MOs, the tracking-based ones are the most commonly used. However, the most recent approaches have opted for other alternatives, such as methods for analyzing the temporal persistence of the detections, or methods based on comparing multiple BG models constructed with different learning rates.

It has also been shown that the detection of ROs is of great interest in many applications (e.g. applications to avoid theft in museums or stores) and that, consequently, some strategies include an additional stage to detect ROs. The mechanisms typically applied in this last stage have also been analyzed.

Acknowledgments

This work was supported in part by the Ministerio de Economía y Competitividad of the Spanish Government under project TEC2013-48453 (MR-UHDTV).

References

Albiol, A., Sanchis, L., Mossi, J.M., 2011. Detection of parked vehicles using spatiotemporal maps. Intell. Transp. Syst. IEEE Trans. 12, 1277–1291.
Amer, A., Dubois, E., Mitiche, A., 2002. Context-independent real-time event recognition: application to key-image extraction. In: Pattern Recognition, 2002. Proceedings. 16th International Conference on, vol. 2. IEEE, pp. 945–948.
Arsic, D., Hofmann, M., Schuller, B., Rigoll, G., 2007. Multi-camera person tracking and left luggage detection applying homographic transformation. Proceeding Tenth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, PETS.
Atev, S., Masoud, O., Papanikolopoulos, N., 2004. Practical mixtures of gaussians with brightness monitoring. In: Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on. IEEE, pp. 423–428.
Aubert, D., Guichard, F., Bouchafa, S., 2004. Time-scale change detection applied to real-time abnormal stationarity monitoring. Real-Time Imaging 10, 9–22.
Auvinet, E., Grossmann, E., Rougier, C., Dahmane, M., Meunier, J., 2006. Left-luggage detection using homographies and simple heuristics. In: Proc. 9th IEEE International Workshop on Performance Evaluation in Tracking and Surveillance (PETS'06). Citeseer, pp. 51–58.
Bangare, P.S., Uke, N.J., Bangare, S.L., 2012. An approach for detecting abandoned object from real time video. Int. J. Eng. Res. Appl. 2, 2646–2649.
Baxter, R.H., Robertson, N.M., Lane, D.M., 2015. Human behaviour recognition in data-scarce domains. Pattern Recognit. 48, 2377–2393.
Bayona, Á., SanMiguel, J.C., Martínez, J.M., 2009. Comparative evaluation of stationary foreground object detection algorithms based on background subtraction techniques. In: Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference on. IEEE, pp. 25–30.
Bayona, Á., SanMiguel, J.C., Martínez, J.M., 2010. Stationary foreground detection using background subtraction and temporal difference in video surveillance. In: Image Processing (ICIP), 2010 17th IEEE International Conference on. IEEE, pp. 4657–4660.
Bevilacqua, A., Stefano, L.D., Lanza, A., 2005. Coarse-to-fine strategy for robust and efficient change detectors. In: Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on. IEEE, pp. 87–92.
Bevilacqua, A., Vaccari, S., 2007. Real time detection of stopped vehicles in traffic scenes. In: Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on. IEEE, pp. 266–270.
Beynon, M.D., Van Hook, D.J., Seibert, M., Peacock, A., Dudgeon, D., 2003. Detecting abandoned packages in a multi-camera video surveillance system. In: Advanced Video and Signal Based Surveillance, 2003. Proceedings. IEEE Conference on. IEEE, pp. 221–228.
Bhargava, M., Chen, C.-C., Ryoo, M.S., Aggarwal, J.K., 2007. Detection of abandoned objects in crowded environments. In: Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on. IEEE, pp. 271–276.
Bird, N., Atev, S., Caramelli, N., Martin, R., Masoud, O., Papanikolopoulos, N., 2006. Real time, online detection of abandoned objects in public areas. In: Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on. IEEE, pp. 3775–3780.
Black, J., Velastin, S., Boghossian, B., 2005. A real time surveillance system for metropolitan railways. In: Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on. IEEE, pp. 189–194.
Bouwmans, T., 2009. Subspace learning for background modeling: a survey. Recent Pat. Comput. Sci. 2, 223–234.
Bouwmans, T., 2014. Traditional and recent approaches in background modeling for foreground detection: an overview. Comput. Sci. Rev. 11, 31–66.
Bouwmans, T., El Baf, F., Vachon, B., 2008. Background modeling using mixture of Gaussians for foreground detection-a survey. Recent Pat. Comput. Sci. 1, 219–237.
Cavallaro, A., Ebrahimi, T., 2000. Video object extraction based on adaptive background and statistical change detection. In: Photonics West 2001-Electronic Imaging. International Society for Optics and Photonics, pp. 465–475.
Chang, J.-Y., Liao, H.-H., Chen, L.-G., 2010. Localized detection of abandoned luggage. EURASIP J. Adv. Signal Process. 2010, 675784.
Chang, L., Zhao, H., Zhai, S., Ma, Y., Liu, H., 2013. Robust abandoned object detection and analysis based on online learning. In: Robotics and Biomimetics (ROBIO), 2013 IEEE International Conference on. IEEE, pp. 940–945.

Cheng, S., Luo, X., Bhandarkar, S.M., 2007. A multiscale parametric background model for stationary foreground object detection. In: Motion and Video Computing, 2007. WMVC'07. IEEE Workshop on. IEEE, p. 18.
Cho, C.-Y., Tung, W.-H., Wang, J.-S., 2008. A crowd-filter for detection of abandoned objects in crowded area. In: Sensing Technology, 2008. ICST 2008. 3rd International Conference on. IEEE, pp. 581–584.
Chuang, C.-H., Hsieh, J.-W., Fan, K.-C., 2007. Suspicious object detection and robbery event analysis. In: Computer Communications and Networks, 2007. ICCCN 2007. Proceedings of 16th International Conference on. IEEE, pp. 1189–1192.
Chuang, C.-H., Hsieh, J.-W., Tsai, L.-W., Chen, S.-Y., Fan, K.-C., 2009. Carried object detection using ratio histogram and its application to suspicious event analysis. Circuits Syst. Video Technol. IEEE Trans. 19, 911–916.
Collazos, A., Fernández-López, D., Montemayor, A.S., Pantrigo, J.J., Delgado, M.L., 2013. Abandoned object detection on controlled scenes using kinect. In: Natural and Artificial Computation in Engineering and Medical Applications. Springer, pp. 169–178.
Cristani, M., Farenzena, M., Bloisi, D., Murino, V., 2010. Background subtraction for automated multisensor surveillance: a comprehensive review. EURASIP J. Adv. Signal Process. 2010, 43.
Cucchiara, R., Grana, C., Piccardi, M., Prati, A., 2003. Detecting moving objects, ghosts, and shadows in video streams. Pattern Anal. Mach. Intell. IEEE Trans. 25, 1337–1342.
Dalley, G., Wang, X., Grimson, W.E.L., 2007. Event detection using an attention-based tracker. In: 10th International Workshop on Performance Evaluation for Tracking and Surveillance. Citeseer, pp. 71–79.
Denman, S., Chandran, V., Sridharan, S., 2007. Abandoned object detection using multi-layer motion detection. In: Proceedings of International Conference on Signal Processing and Communication Systems 2007. DSP for Communication Systems, pp. 439–448.
Elgammal, A., Duraiswami, R., Harwood, D., Davis, L.S., 2002. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. Proc. IEEE 90, 1151–1163.
Elhabian, S.Y., El-Sayed, K.M., Ahmed, S.H., 2008. Moving object detection in spatial domain using background removal techniques-state-of-art. Recent Pat. Comput. Sci. 1, 32–54.
Evangelio, R.H., Sikora, T., 2010. Static object detection based on a dual background model and a finite-state machine. EURASIP J. Image Video Process. 2011, 858502.
Evangelio, R.H., Sikora, T., 2011. Complementary background models for the detection of static and moving objects in crowded environments. In: Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on. IEEE, pp. 71–76.
Fan, Q., Gabbur, P., Pankanti, S., 2013. Relative attributes for large-scale abandoned object detection. In: Computer Vision (ICCV), 2013 IEEE International Conference on. IEEE, pp. 2736–2743.
Fan, Q., Pankanti, S., 2011. Modeling of temporarily static objects for robust abandoned object detection in urban surveillance. In: Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on. IEEE, pp. 36–41.
Fan, Q., Pankanti, S., 2012. Robust foreground and abandonment analysis for large-scale abandoned object detection in complex surveillance videos. In: Advanced Video and Signal-Based Surveillance (AVSS), 2012 IEEE Ninth International Conference on. IEEE, pp. 58–63.
Ferrando, S., Gera, G., Massa, M., Regazzoni, C.S., 2006. A new method for real time abandoned object detection and owner tracking. In: Image Processing (ICIP), 2006. 15th IEEE International Conference on, pp. 3329–3332.
Ferryman, J., Hogg, D., Sochman, J., Behera, A., Rodriguez-Serrano, J.A., Worgan, S., Li, L., Leung, V., Evans, M., Cornic, P., et al., 2013. Robust abandoned object detection integrating wide area visual surveillance and social context. Pattern Recognit. Lett. 34, 789–798.
Filonenko, A., Jo, K.-H., et al., 2015. Detecting abandoned objects in crowded scenes of surveillance videos using adaptive dual background model. In: Human System Interactions (HSI), 2015 8th International Conference on. IEEE, pp. 224–227.
Fisher, R., Santos-Victor, J., Crowley, J., 2005. Caviar: Context aware vision using image-based active recognition. http://homepages.inf.ed.ac.uk/rbf/CAVIAR.
Foresti, G.L., Marcenaro, L., Regazzoni, C.S., 2002. Automatic detection and indexing of video-event shots for surveillance applications. Multimedia IEEE Trans. 4, 459–471.
Fu, H., Xiang, M., Ma, H., Ming, A., Liu, L., 2011. Abandoned object detection in highway scene. In: Pervasive Computing and Applications (ICPCA), 2011 6th International Conference on. IEEE, pp. 117–121.
Fujiyoshi, H., Kanade, T., 2004. Layered detection for multiple overlapping objects. IEICE Trans. Inf. Syst. 87, 2821–2827.
Gallego, G., Cuevas, C., Mohedano, R., Garcia, N., 2013. On the mahalanobis distance classification criterion for multidimensional normal distributions. Signal Process. IEEE Trans. 61, 4387–4396.
Gallego, J., Pardas, M., Landabaso, J.-L., 2008. Segmentation and tracking of static and moving objects in video surveillance scenarios. In: Image Processing (ICIP), 2008. 15th IEEE International Conference on. IEEE, pp. 2716–2719.
Geng, L., Xiao, Z., 2011. Real time foreground-background segmentation using two-layer codebook model. In: Control, Automation and Systems Engineering (CASE), 2011 International Conference on. IEEE, pp. 1–5.
Gibbins, D., Newsam, G.N., Brooks, M.J., 1996. Detecting suspicious background changes in video surveillance of busy scenes. In: Applications of Computer Vision, 1996. WACV'96., Proceedings 3rd IEEE Workshop on. IEEE, pp. 22–26.
Goyette, N., Jodoin, P.-M., Porikli, F., Konrad, J., Ishwar, P., 2012. Changedetection.net: A new change detection benchmark dataset. In: Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on. IEEE, pp. 1–8.
Guler, S., Farrow, M.K., 2006. Abandoned object detection in crowded places. In: Proc. of PETS. Citeseer, pp. 18–23.
Guler, S., Silverstein, J.A., Pushee, I.H., 2007. Stationary objects in multiple object tracking. In: Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on. IEEE, pp. 248–253.
Gutchess, D., Trajković, M., Cohen-Solal, E., Lyons, D., Jain, A.K., 2001. A background model initialization algorithm for video surveillance. In: Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on, vol. 1. IEEE, pp. 733–740.
Harris, C., Stephens, M., 1988. A combined corner and edge detector. In: Alvey Vision Conference, vol. 15. Citeseer, p. 50.
Hassan, W., Birch, P., Mitra, B., Bangalore, N., Young, R., Chatwin, C., 2013. Illumination invariant stationary object detection. IET Comput. Vis. 7, 1–8.
Heikkilä, M., Pietikäinen, M., 2006. A texture-based method for modeling the background and detecting moving objects. Pattern Anal. Mach. Intell. IEEE Trans. 28, 657–662.
Horprasert, T., Harwood, D., Davis, L.S., 1999. A statistical approach for real-time robust background subtraction and shadow detection. In: Computer Vision, 1999. ICCV 1999. Proceedings. IEEE International Conference on, vol. 99, pp. 1–19.
Hsu, R.-L., Abdel-Mottaleb, M., Jain, A.K., 2002. Face detection in color images. Pattern Anal. Mach. Intell. IEEE Trans. 24, 696–706.
Hyvärinen, A., Karhunen, J., Oja, E., 2004. Independent component analysis, vol. 46. John Wiley & Sons.
Ingersoll, K., Niedfeldt, P.C., Beard, R.W., 2015. Multiple target tracking and stationary object detection in video with recursive-ransac and tracker-sensor feedback. In: Unmanned Aircraft Systems (ICUAS), 2015 International Conference on. IEEE, pp. 1320–1329.
Jardim, E., Bian, X., da Silva, E.A., Netto, S.L., Krim, H., 2015. On the detection of abandoned objects with a moving camera using robust subspace recovery and sparse representation. In: Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on. IEEE, pp. 1295–1299.
Joglekar, U.A., Awari, S.B., Deshmukh, S.B., Kadam, D.M., Awari, R.B., 2014. An abandoned object detection system using background segmentation. International Journal of Engineering Research and Technology, vol. 3. ESRSA Publications.
Jolliffe, I., 2002. Principal Component Analysis. Wiley Online Library.
Kim, J., Kang, B., 2014. Nonparametric state machine with multiple features for abnormal object classification. In: Advanced Video and Signal Based Surveillance (AVSS), 2014 11th IEEE International Conference on. IEEE, pp. 199–203.
Kim, J., Kang, B., Wang, H., Kim, D., 2012. Abnormal object detection using feedforward model and sequential filters. In: Advanced Video and Signal-Based Surveillance (AVSS), 2012 IEEE Ninth International Conference on. IEEE, pp. 70–75.
Kim, K., Chalidabhongse, T.H., Harwood, D., Davis, L., 2005. Real-time foreground–background segmentation using codebook model. Real Time Imag. 11, 172–185.
Kwak, S., Bae, G., Byun, H., 2010. Abandoned luggage detection using a finite state automaton in surveillance video. Opt. Eng. 49, 027007.
Lai, T.Y., Kuo, J.Y., Liu, C.-H., Wu, Y.W., Fanjiang, Y.-Y., Ma, S.-P., 2012. Intelligent detection of missing and unattended objects in complex scene of surveillance videos. In: Computer, Consumer and Control (IS3C), 2012 International Symposium on. IEEE, pp. 662–665.
Lasiesta: Labeled and annotated sequences for integral evaluation of segmentation and tracking algorithms, 2016. http://www.gti.ssr.upm.es/data.
Latecki, L.J., Miezianko, R., 2006. Object tracking with dynamic template update and occlusion detec. In: Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, Vol. 1. IEEE, pp. 556–560.
Lee, D.D., Seung, H.S., 2001. Algorithms for non-negative matrix factorization. In: Advances in neural information processing systems, pp. 556–562.
Lee, J.T., Ryoo, M.S., Riley, M., Aggarwal, J., 2009. Real-time illegal parking detection in outdoor environments using 1-d transformation. Circ. Syst. Video Technol. IEEE Trans. 19, 1014–1024.
Li, L., Luo, R., Ma, R., Huang, W., Leman, K., 2006. Evaluation of an ivs system for abandoned object detection on pets 2006 datasets. In: Proceedings of the 9th IEEE International Workshop on Performance Evaluation in Tracking and Surveillance (PETS'06). Citeseer, pp. 91–98.
Li, Q., Mao, Y., Wang, Z., Xiang, W., 2009. Robust real-time detection of abandoned and removed objects. In: Image and Graphics, 2009. ICIG'09. Fifth International Conference on. IEEE, pp. 156–161.
Li, X., Zhang, C., Zhang, D., 2010. Abandoned objects detection using double illumination invariant foreground masks. In: Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, pp. 436–439.
Liao, H.-H., Chang, J.-Y., Chen, L.-G., 2008. A localized approach to abandoned luggage detection with foreground-mask sampling. In: Advanced Video and Signal Based Surveillance, 2008. AVSS'08. IEEE Fifth International Conference on. IEEE, pp. 132–139.
i LIDS Team, 2006. Imagery library for intelligent detection systems (i-lids); a standard for testing video based detection systems. In: IEEE Carnahan Conf. Security Technology. IEEE, pp. 75–80.
Lim, S.N., Davis, L.S., 2006. A One-Threshold Algorithm for Detecting Abandoned Packages Under Severe Occlusions Using a Single Camera. University of Maryland, College Park.

Lin, C.-Y., Wang, W.-H., 2008. An abandoned objects management system based on the gaussian mixture model. In: Convergence and Hybrid Information Technology, 2008. ICHIT'08. International Conference on. IEEE, pp. 169–175.
Lin, K., Chen, S.-C., Chen, C.-S., Lin, D.-T.D., Hung, Y.-P., 2015. Abandoned object detection via temporal consistency modeling and back-tracing verification for visual surveillance. Inf. Forensics Secur. IEEE Trans. 10, 1359–1370.
Liu, N., Wu, H., Lin, L., 2015. Hierarchical ensemble of background models for ptz-based video surveillance. Cybern. IEEE Trans. 45, 89–102.
Lopez-Mendez, A., Monay, F., Odobez, J.-M., 2014. Exploiting scene cues for dropped object detection. In: 9th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications..
Luo, X., Bhandarkar, S.M., 2005. Real-time and robust background updating for video surveillance and monitoring. In: Image Analysis and Recognition. Springer, pp. 1226–1233.
Lv, F., Song, X., Wu, B., Singh, V.K., Nevatia, R., 2006. Left luggage detection using bayesian inference. In: Proc. of IEEE Int. Workshop on Performance Evaluation of Tracking and Surveillance. Citeseer, pp. 83–90.
Maddalena, L., Petrosino, A., 2012. The sobs algorithm: what are the limits? In: Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on. IEEE, pp. 21–26.
Maddalena, L., Petrosino, A., 2013. Stopped object detection by learning foreground model in videos.. IEEE Trans. Neural Netw. Learn. Syst. 24, 723–735.
Magno, M., Tombari, F., Brunelli, D., Di Stefano, L., Benini, L., 2009. Multimodal abandoned/removed object detection for low power video surveillance systems. In: Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference on. IEEE, pp. 188–193.
Mahin, F.S., Islam, M.N., Schaefer, G., Ahad, M.A.R., 2015. A simple approach for abandoned object detection. In: Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), p. 427.
Martínez, R., Cuevas, C., Berjón, D., García, N., 2015. Detection of static moving objects using multiple nonparametric background models. In: Consumer Electronics (ISCE), 2015 IEEE International Symposium on. IEEE, pp. 1–2.
Martínez-del Rincón, J., Herrero-Jaraba, J.E., Gómez, J.R., Orrite-Urunuela, C., 2006. Automatic left luggage detection and tracking using multi-camera ukf. In: Proceedings of the 9th IEEE International Workshop on Performance Evaluation in Tracking and Surveillance (PETS'06). Citeseer, pp. 59–66.
Mathew, R., Yu, Z., Zhang, J., 2005. Detecting new stable objects in surveillance video. In: Multimedia Signal Processing, 2005 IEEE 7th Workshop on. IEEE, pp. 1–4.
McFarlane, N.J., Schofield, C.P., 1995. Segmentation and tracking of piglets in images. Mach. Vis. Appl. 8, 187–193.
Melli, R., Prati, A., Cucchiara, R., de Cock, L., Traficon, N., 2005. Predictive and probabilistic tracking to detect stopped vehicles.. In: WACV/MOTION, pp. 388–393.
Miezianko, R., Pokrajac, D., 2008. Detecting and recognizing abandoned objects in crowded environments. In: Computer Vision Systems. Springer, pp. 241–250.
Morde, A., Ma, X., Guler, S., 2012. Learning a background model for change detection. In: Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on. IEEE, pp. 15–20.
Morozov, A., 2015. Development of a method for intelligent video monitoring of abnormal behavior of people based on parallel object-oriented logic programming. Pattern Recognit. Image Anal. 25, 481–492.
Muchtar, K., Lin, C.-Y., Kang, L.-W., Yeh, C.-H., 2013. Abandoned object detection in complicated environments. In: Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific. IEEE, pp. 1–6.
Murino, V., 1998. Structured neural networks for pattern recognition. Syst. Man Cybern. Part B Cybern. IEEE Trans. 28, 553–561.
Ortego, D., SanMiguel, J.C., 2013. Stationary foreground detection for video-surveillance based on foreground and motion history images. In: Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on. IEEE, pp. 75–80.
Ortego, D., SanMiguel, J.C., 2014. Multi-feature stationary foreground detection for crowded video-surveillance. In: Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, pp. 2403–2407.
Pan, J., Fan, Q., Pankanti, S., 2011. Robust abandoned object detection using region-level analysis. In: Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE, pp. 3597–3600.
Pauly, O., Diotte, B., Fallavollita, P., Weidert, S., Euler, E., Navab, N., 2015. Machine learning-based augmented reality for improved surgical scene understanding. Comput. Med. Imag. Graph. 41, 55–60.
Pets: The performance evaluation of tracking and surveillance, 2000–2007. http://www.cvg.rdg.ac.uk/.
Porikli, F., 2007. Detection of temporarily static regions by processing video at different frame rates. In: Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on. IEEE, pp. 236–241.
Porikli, F., Ivanov, Y., Haga, T., 2008. Robust abandoned object detection using dual foregrounds. EURASIP J. Adv. Signal Process. 2008, 30.
Raheja, J.L., Malireddy, C., Singh, A., Solanki, L., 2011. Detection of abandoned objects in real time. In: Electronics Computer Technology (ICECT), 2011 3rd International Conference on, vol. 2. IEEE, pp. 199–203.
Rodriguez-Fernandez, D., Vilarino, D., Pardo, X., 2008. Cnn implementation of a moving object segmentation approach for real-time video surveillance. In: Cellular Neural Networks and Their Applications, 2008. CNNA 2008. 11th International Workshop on. IEEE, pp. 129–134.
Sacchi, C., Regazzoni, C.S., 2000. A distributed surveillance system for detection of abandoned objects in unmanned railway environments. Veh. Technol. IEEE Trans. 49, 2013–2026.
Sajith, K., Nair, K.R., 2013. Abandoned or removed objects detection from surveillence video using codebook. International Journal of Engineering Research and Technology, vol. 2. ESRSA Publications.
San Miguel, J.C., Martínez, J.M., 2008. Robust unattended and stolen object detection by fusing simple algorithms. In: Advanced Video and Signal Based Surveillance, 2008. AVSS'08. IEEE Fifth International Conference on. IEEE, pp. 18–25.
da Silva, A.F., Thomaz, L., Carvalho, G., Nakahata, M.T., Jardim, E., de Oliveira, J.F., da Silva, E.A., Netto, S.L., Freitas, G., Costa, R.R., et al., 2014. An annotated video database for abandoned-object detection in a cluttered environment. In: Telecommunications Symposium (ITS), 2014 International. IEEE, pp. 1–5.
Singh, A., Sawan, S., Hanmandlu, M., Madasu, V.K., Lovell, B.C., 2009. An abandoned object detection system based on dual background segmentation. In: Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference on. IEEE, pp. 352–357.
Singh, R., Vishwakarma, S., Agrawal, A., Tiwari, M., 2010. Unusual activity detection for video surveillance. In: Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia. ACM, pp. 297–305.
Smith, K.C., Quelhas, P., Gatica-Perez, D., 2006. Detecting Abandoned Luggage Items in a Public Space. Technical Report. IDIAP.
Sobral, A., Vacavant, A., 2014. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. Comput. Vis. Image Understand. 122, 4–21.
Spagnolo, P., Caroppo, A., Leo, M., Martiriggiano, T., D'Orazio, T., 2006. An abandoned/removed objects detection algorithm and its evaluation on pets datasets. In: Video and Signal Based Surveillance, 2006. AVSS'06. IEEE International Conference on. IEEE, p. 17.
Spengler, M., Schiele, B., 2003. Automatic detection and tracking of abandoned objects. In: Proceedings of the Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance. Citeseer.
Stauffer, C., Grimson, W.E.L., 1999. Adaptive background mixture models for real–time tracking. In: Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., vol. 2. IEEE.
Stringa, E., Regazzoni, C.S., 2000. Real-time video-shot detection for scene surveillance applications. Image Process. IEEE Trans. 9, 69–79.
Tian, Y., Feris, R.S., Liu, H., Hampapur, A., Sun, M.-T., 2011. Robust detection of abandoned and removed objects in complex surveillance videos. Syst. Man Cybern. Part C: Appl. Rev. IEEE Trans. 41, 565–576.
Tian, Y.-l., Feris, R., Hampapur, A., 2008. Real-time detection of abandoned and removed objects in complex environments. The Eighth International Workshop on Visual Surveillance-VS2008.
Tian, Y.-L., Lu, M., Hampapur, A., 2005. Robust and efficient foreground analysis for real-time video surveillance. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1. IEEE, pp. 1182–1187.
Tripathi, R.K., Jalal, A.S., Bhatnagar, C., 2013. A framework for abandoned object detection from video surveillance. In: Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2013 Fourth National Conference on. IEEE, pp. 1–4.
Velastin, S.A., Boghossian, B.A., Lo, B.P., Sun, J., Vicencio-Silva, M.A., 2005. Prismatica: toward ambient intelligence in public transport environments. Syst. Man Cybern., Part A: Syst. Hum. IEEE Trans. 35, 164–182.
Vezzani, R., Cucchiara, R., 2010. Video surveillance online repository (visor): an integrated framework. Multimedia Tools Appl. 50, 359–380.
Viola, P., Jones, M., 2001. Rapid object detection using a boosted cascade of simple features. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1. IEEE, pp. I–511.
Wang, J., Ooi, W.-T., 1999. Detecting Static Objects in Busy Scenes. Technical Report. Cornell University.
Wang, L., Hu, W., Tan, T., 2003. Recent developments in human motion analysis. Pattern Recognit. 36, 585–601.
Wang, R., Bunyak, F., Seetharaman, G., Palaniappan, K., 2014. Static and moving object detection using flux tensor with split gaussian models. In: Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on. IEEE, pp. 420–424.
Wang, W., Liu, Z., 2010. A new approach for real-time detection of abandoned and stolen objects. In: Electrical and Control Engineering (ICECE), 2010 International Conference on. IEEE, pp. 128–131.
Wen, J., Gong, H., Zhang, X., Hu, W., 2009. Generative model for abandoned object detection. In: Image Processing (ICIP), 2009 16th IEEE International Conference on. IEEE, pp. 853–856.
Wren, C.R., Azarbayejani, A., Darrell, T., Pentland, A.P., 1997. Pfinder: Real-time tracking of the human body. Pattern Anal. Mach. Intell. IEEE Trans. 19, 780–785.
Xiya, L., Jingling, W., Qin, Z., 2012. An abandoned object detection system based on dual background and motion analysis. In: Computer Science & Service System (CSSS), 2012 International Conference on. IEEE, pp. 2293–2296.
Yang, T., Pan, Q., Li, S.Z., Li, J., 2004. Multiple layer based background maintenance in complex environment. In: Multi-Agent Security and Survivability, 2004 IEEE First Symposium on. IEEE, pp. 112–115.
Yang, Z., Rothkrantz, L., 2011. Surveillance system using abandoned object detection. In: Proceedings of the 12th International Conference on Computer Systems and Technologies. ACM, pp. 380–386.
Yao, J., Odobez, J.-M., 2007. Multi-layer background subtraction based on color and texture. In: Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on. IEEE, pp. 1–8.

Zeng, Y., Lan, J., Ran, B., Gao, J., Zou, J., 2015. A novel abandoned object detection system based on three-dimensional image information. Sensors 15, 6885–6904.
Zhao, Y., Gong, H., Lin, L., Jia, Y., 2008. Spatio-temporal patches for night background modeling by subspace learning. In: Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE, pp. 1–4.
Zin, T.T., Tin, P., Hama, H., Toriu, T., 2011. Unattended object intelligent analyzer for consumer video surveillance. Consum. Electron. IEEE Trans. 57, 549–557.
Zin, T.T., Tin, P., Toriu, T., Hama, H., 2012. A probability-based model for detecting abandoned objects in video surveillance systems. In: Proceedings of the World Congress on Engineering, vol. 2.