Existing Structures, Problems and Future Directions in Background Subtraction of Practical Application

T. Logeswari

Abstract: PC video-based vision frameworks additionally include the ID of moving articles in their first stage. Subtraction of the foundation is then applied to recognize the foundation from the frontal area. Setting subtraction in writing is unquestionably one of the most contemplated zones of PC vision and incorporates an immense number of productions. The majority of them are worried about adjusting scientific and AI strategies to be progressively effective in the game difficulties. A definitive objective, however, is that the setting subtraction approaches worked in research can be utilized in genuine applications for example, traffic checking. However taking a gander at the writing, we should see that there is frequently a distinction in basic research between the ebb and flow strategies utilized in genuine applications and the new techniques. Be that as it may, the recordings broke down in enormous scale archives are not far reaching in the sense they tended to only 50% of the full range of issues in genuine applications. In any case, taking a gander at the writing, we will take note of that there is frequently a distinction between the ebb and flow strategies utilized in genuine applications and the new techniques utilized in crucial research. In any case, the recordings broke down in enormous scale archives are not extensive in the sense they tended to only 50% of the full range of issues in genuine applications. We additionally recognize the foundation models that are utilized adequately in these applications to distinguish potential usable new foundation models regarding heartiness, time and memory necessity.

Keywords: Background subtraction Background initialization Foreground detection Visual surveillance.

I. INTRODUCTION

Ground initialization and ground subtraction was commonly used in various computer vision applications based on video captured by fixed cameras with the growth of the different sensors. Such projects include a wide range of situations with different challenges and different types of shifting objects of interest in the foreground [1]. Intelligent video surveillance systems of human activities such as road traffic monitoring, airport and marine surveillance are undoubtedly the most well-known and oldest technologies. But in order to study the actions of the observed animals in their environment, identification of moving objects is also needed for intelligent visual observation systems of animals and insects. Nonetheless, it involves perceptual perception with specific challenges in natural environments such as trees, ponds, streams, oceans and submarines. Certain specific systems such as visual motion capture, human-machine interface framework, hand gesture identification dependent on vision, Content-based video coding and background substitution both involve initialization and context subtraction in their first phase[2]. Even if the identification of moving objects is commonly used in these real applications, there is no full sample in the literature defining them. Current models and challenges in videos taken with fixed cameras for these applications were met by reviews and groups in one paper. In addition, most of the research is carried out on large-scale data sets that often consist of videos taken for the purpose of researchers’ evaluation, and thus are not covered by several challenging situations that appear in real cases[5]. Such background creates a conflict between the complexity of research-grade approaches and actual practical issues, both in terms of used models and solved challenges[3]. In particular, research focuses on future paths of advanced mathematical models, machine learning models, signal processing models and classification models, whereas practical systems still use models such as mean, Median and MOG models evolved twenty years ago. Background subtraction methods can, however, be classified as statistical models, fuzzy models and Dempster–Shafer models for mathematical concepts; subspace learning models that are either reconstructive, discriminatory and mixed; robust subspace learning through matrix decomposition or tensor decomposition; robust subspace tracking; support vector machines, neural networks and deep learning; In practice, however, most authors are in real applications. In order to address the above concerns, we first attempt, in this analysis, to examine most of the individual applications that used context initialization and subtraction in their method by classifying them in terms of targets, conditions and items of interest[4][6].

II. LITERATURE REVIEW: SUBTRACTION OF BACKGROUND

In this part, the goal remains for us momentarily to remove from a video stream the backdrop to segment static and shifting foreground objects. In many video surveillance systems, this function is the key step, and context subtraction provides an appropriate solution that offers a good quality detection and estimation time compromise[9]. The foregoing are the following measures of context subtraction: The first background image is computed in the initialization of backgrounds[10] (also called the background generation, background extraction and background reconstruction). Context modeling defines the construct used for the context[7] (also known as Background Representation). Maintenance of backgrounds is about the updating process to adjust the layout to time-specific changes.
I. In the Ground / Moveable Target Identification of the Pixels (also known as the Foreground Detection) is labeled as the Context or the Moveable Subject type[8].

Such multiple measures using approaches that have different objectives and restrictions. Therefore, you need different characteristics of algorithms. The initialization of context includes algorithms "offline" which are "batch" by simultaneously using all the info. On the other side, context support needs "at-line" algorithms that are "incremental" by taking input data one by one. Initialization, modeling and maintaining backgrounds require rebuilding algorithms, while primary detection requires discriminatory algorithms.

II. BACKGROUND SUBTRACTION BASED APPLICATIONS

The segmentation of static and shifting context artifacts from a picture source is a major step in many computer vision systems, in which background subtraction provides an appropriate solution that offers a good degree of identification and measurement time efficiency balance[11].

1. The aim is to identify and track objects of interest in various environments. Visual surveillance of human activity.

Traffic scenes (also called highway or path) are the most popular conditions for examining accidents such as road stopped vehicles or for measuring traffic density on highways, then graded into flat, smooth, heavy and Jam. Therefore, the number of vehicles must be registered, tracked, or counted. The use of background detection can also be made for congestion detection, for illegal detection of parking and for recognition of free parking spaces in urban traffic surveillance. For safety at train Stations and airports, where unattended baggage can be a key objective, it is also essential.

The number of ships operating in maritime or port areas and the identification and monitoring of ships in river canals may also take place at marine scenes to observe human activities. Shop scenes to identify and monitor customers are other worlds[14].

2. Visual hull computation: For the image reconstruction, visual hull is used to achieve an estimated geometric representation of a static entity. In the first case a practical model is required through the use of an image-based object model. In the second case, a human image model can be used for optical motion capture.

Virtual hull is a geometrical structure obtained with the 3D reconstruction technique from Laurentini's form-from-silhouette. Next, the schematic subtraction separates the front and rear entity to achieve the first mask called the profile, which is then assumed to be the 2 D projection of the foreground object.

3. Human–Machine Interaction (HMI): This requires human–machine interaction in many areas such as the sciences, sports and ludo applications. For games, the player can display his own picture or outline, as in the PlayStation Eye-Toy, designed into a virtual scene. Many implementations for ludo-multimedia include the identification of a chosen moving object.

4. Framework for hand gesture detection dependent on vision: This framework requires to recognize, register and classify hand gesture for several purposes, including the human-computer interface, behavioral analysis, perception and understanding of sign language, teleconferencing, distance learning, robots, gaming collection or virtual object handling.

5. Video content coded: only the main structures are conveyed via moving objects for example, MPEG-4 Multimedia Communications Standard improves content-based flexibility by using a Video Objective Plant (VOP) as a fundamental coding feature in the video content coding scheme for transmitting such as events, multimedia films and video phones.

III. SMART VISUAL MONITORING HUMAN ACTIVITIES

Visual Monitoring is the primary method of context modeling and foreground detection and we only checked papers published after 1997, because variations in two or three frames due to computer constraints have been developed before the techniques used to monitor static or moving objects[12]. Practically, the purpose of visual analysis is to track static or shifting foreground artifacts automatically as follows:

a. Traffic monitoring: Examples of traffic monitoring have their own property, with respect to camera locations, environments and moving body forms, as follows:

1.1 Location of the cameras

1.2 Quality of the cameras: CCTV cameras are mostly used, but cameras' output can range from low to high quality (HD cameras). In order to decrease the data and the size, low quality cameras produce one to two frames per second and 100k pixels / frame[13].

In reality, it is estimated that a video produced by a low-quality city camera is about 1 GB to 10 GB per day.
1.3 Environments: traffic scenes present their various challenges to roads, roads and urban traffic environments. Shadows and luminous shifts also arise in road scenes. Route scenes often have backgrounds.

Airport visual surveillance concerns primarily those areas where specialized ground vehicles such as fueling vehicles and cars and tracking persons, such as employees, are parked and maintained. Visual supervision is needed because:
(1) the plane of an airport is a security-related area; (2) the need for transit is helped, i.e. the time when the aircraft park on the apron and (3) the need for airport operators to minimize costs because the personnel are able to be deployed more efficiently; and the need for passenger latencies is minimized as necessary[16].

2. Challenges because of foreground objects:
– During the monitoring time, ground cars that modify their form extensively, e.g. luggage cars. Strictly static action and model structures cannot be used to control such cars.
On the tarmac there may be large aircraft with clear details and little aircraft with boring outlines.
– Land airport security camouflage also exists as equal patterns of gray and white color are exchanged between aircraft and ground.

C. Maritime Marine Surveillance
Visible radiation or IR frequencies can be used for marine monitoring. The goal is to list, monitor and classify boats in water and in the seas, for instance. There is a general tool called Movement Meerkat for the analysis of movement in aquatic settings. Motion Meerkat eliminates the task of processing video streams through the retrieval of camera movement objects. The Meerkat trend could either be used as the history pattern by Running Gaussian Average or MOG. Motion Meerkat has been very popular in a number of environmentally friendly habitats but remains prone to obstacles such as rapid changes in lighting and camouflage. Virtually any natural environment is extremely challenging, for example the woodland canopy, the river and the shore, because key items can be blended with the context undoubtedly.
1. The goal is to identify humans and animals, but the shifting of the leaves causes quick transitions from light to dark.
In fact, divisions may partly occult human beings or creatures. The foreground blocks are then attached to represent the foreground areas which are applicants for the region to be confirmed to be human, object or history in order to mark the region with labelled bounding boxes[17].
2. Riverbeds The aim is to locate foreign objects in the rivers (bottles, forests) for (1) river conservation or (2) civil infrastructure security including bridges and river gates. River ecosystem River restoration is the concept. In the first example, the atmosphere is polluted with foreign objects and livestock are harmed.

Bloisi used the Independent Multimodal Bottom Subtraction (IMBS) to cope with highly dynamic situations that are marked by non-regular and high-frequency noise, such as a water backdrop in another job. IMBS is a non-recursive perpixel. Certain techniques, such as morphology and saliency identification, are often used for treating false positive detections caused by water flow.

The goal is to track foreign objects in natural environments including land, sea and river to preserve biodiversity.

Active visual detection in nature ecosystems like Fauna and vegetation environments.
In the second example, the probability of trees’ destruction is directly proportional to the scale of foreign objects such as falling trees and bushes, branches of fallen forests and other small pieces of wood. Larger falling trees than those in smaller parts of fallen trees are more vulnerable.

3. Ocean environments The aim is (1) to monitor traffic improving ships, for the most part, (2) foreign items, to avoid collision with alien artifacts, and (3) foreign persons (intruders), as ships are at risk for attacks by the pirates in open water or in a port area. Boat wakes and environmental issues often contribute to the development of a highly dynamic context, such as dawn, noon, evening, cloud, rain and fog.

VI. THE BACKGROUND SUBTRACTION IS USED FOR THE VISUAL EVALUATION OF HUMAN ACTIVITIES

like in sport, as a means of fast taking important decisions (1), (2) for accurate athletic performance analysis, and (3) for surveillance of dangerous activities. The analysis of human activities is automatically visualized. John, for starters... Provide a system to ensure that a coach gets input in real time to make sure the process is carried out properly[16]. The stationary camera records the background image without the consumer during the initialization step, and then the outline is removed from each current image.

Figure 4 : Human Activities for Background Subtraction

The goal is to create a 3D image of a person shot by multi-cameras that can involve a marker or not. Digital motion capture Due to the difficulty of a detailed 3D reconstruction of the human model, the outline (also called the visual hull) measures the 3D voxel estimate provided by each frame[18]. The gestures are then tracked and replicated in the human body model known as avatar. The gesture capturing is used in computer games, the technology for synthetic clothes and augmented reality.8. Optical motion capture The goal is to get a 3D image of a person who can or cannot be labelled, shot by a machine of multi-cameras.

IV. HUMAN-MACHINE INTERACTIONS

Most applications require human-machine interactions by real-time video captured by fixed-time devices, such as gaming (Microsoft’s Kinect) and ludo-applications.–Arts and games: The body pixels are first placed with context subraction, and this detail is then used as a foundation for graphical answers in Levin's immersive structures. In 2003, Warren proposed various key strategies for contact that could use this type of body-pixel info. The system is effective in "close to mirrors," as is the videogame by Myron Krueger[19].-Ludo-multimedia apps: The consumer may pick a moving object that is of interest on the monitor in this type of application and then give information. The visitor to an aquarium can choose fish from an interactive system which are captured on-line using a remote video camera as a representative example. The game is a two-part touch screen. The first one mentions the fish in the tank and is always helpful.

Figure 5 : Human Machine Interaction

III. RESULT ANALYSIS

We addressed in this analysis all methods for the identification of static or shifting important artifacts utilizing context subtraction. We also discussed the challenges associated with the multiple scenarios and moving objects. Thus the following results can be reached:–Such existing tests also show the value of moving object detection in video as monitoring, recognition or behavioral analysis is a first step. So, for real-time applications the foreground mask must be as accurate as possible and quickly achieved.In these various applications there are several requirements and they must address specific critical situations including (1) the position of cameras that can generate small or large moving objects for terms of the picture scale to be observed, (2) the nature of atmosphere and (3) the form of context subtractions.

As the conditions are very diverse, after implementation the context model needs to address various challenges. In fact, the moving objects of concern have very distinct presentation characteristics. Thus, different background models must be built for a particular application or a generalized background model can be used in all applications. For a general model of history, a different context model can be better created with specific challenges and the appropriate background model can be used when the relevant difficulties are defined.

–Basic models are stable enough for indoor applications in managed environments, such as visual motion capture. Statistical models are a good system for traffic monitoring, but problems such as increases in lighting and in-depth sleeping / beginning items will incorporate new changes. More rigorous history methods than the top
ChangeDetection.net methods are needed for coastal and underwater environments for natural environments and in particular, for maritime and aquatic ecosystems, as developed in 2017. The main conclusion here is that BGS approaches commonly used in real implementations are usually slightly enhanced over older methods. That means that new RPCA and deep learning

| Applications | Scenes                  | Challenges | Solved–unsolved |
|--------------|-------------------------|------------|------------------|
| (1) Intelligent visual surveillance of human activities | | | |
| (1.1) Traffic surveillance | Outdoor scenes | Multi-modal backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| | | Camera jitter | Partially solved |
| (1.2) Airport surveillance | Outdoor scenes | Illumination changes | Partially solved |
| | | Camera jitter | Partially solved |
| | | Illumination changes | Partially solved |
| (1.3) Maritime surveillance | Outdoor scenes | Multimodal backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| | | Camera jitter | Partially solved |
| (1.4) Store surveillance | Indoor scenes | Multimodal backgrounds | Solved |
| (1.5) Coastal Surveillance | Outdoor scenes | Multimodal backgrounds | Partially solved |
| (1.6) Swimming Pools Surveillance | Outdoor scenes | Multimodal backgrounds | Solved |
| (2) Intelligent visual observation of animal and insect behaviors | | | |
| (2.1) Birds surveillance | Outdoor scenes | Multimodal backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| | | Camera jitter | Partially solved |
| (2.2) Fish surveillance | Aquatic scenes | Multimodal backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| (2.3) Dolphins surveillance | Aquatic scenes | Multimodal backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| (2.4) Honeybees surveillance | Outdoor scenes | Small objects | Partially solved |
| (2.5) Spiders surveillance | Outdoor scenes | | |
| (2.6) Lizards surveillance | Outdoor scenes | Multimodal backgrounds | Partially solved |
| (2.7) Pigs surveillance | Indoor scenes | Illumination changes | Partially solved |
| (2.7) Hinds surveillance | Outdoor scenes | Multimodal backgrounds | Partially solved |
| | | Low-frame rate | Partially solved |
| (3) Intelligent visual observation of natural environments | | | |
| (3.1) Forest | Outdoor scenes | Multimodal backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| | | Low-frame rate | Partially solved |
| (3.2) River | Aquatic scenes | Multimodal backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| (3.3) Ocean | Aquatic scenes | Multimodal backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| (3.4) Submarine | Aquatic scenes | Multimodal backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| (4) Intelligent analysis of human activities | | | |
| (4.1) Soccer | Outdoor scenes | Small objects | Solved |
| | | Illumination changes | Solved |
| (4.2) Rowing | Indoor scenes | | |
| (4.3) Surf | Aquatic scenes | Dynamic backgrounds | Partially solved |
| | | Illumination changes | Partially solved |
| (5) Visual hull computing | | | |
| Image-based modeling | Indoor scenes | Shadows/highlights | Solved |
| Optical motion capture | Indoor scenes | Shadows/highlights | Solved (SG) |
| (6) Human–Machine interaction (HMI) | | | |
| Arts | Indoor scenes | | |
| Games | Indoor scenes | | |
| Ludo-Multimedia | Indoor scenes | | |
| (7) Vision-based hand gesture recognition | | | |
| Human–computer interface (HCI) | Indoor scenes | | |
| Behavior analysis | Indoor scenes | | |
| Sign language interpretation and learning | Indoor scenes | | |
| Robotics | Indoor scenes | | |
| (7) Content based Video coding | | | |
| | | | |

Figure 6: Application of Scenes
models have to be looked at with this type of environment. Two processes are the result. Second, many engineers use public implementations to restrict their work to old proven methods because they are constrained by the period of growth. Third, owing to their memory and computation criteria, modern approaches can sometimes not be used in real applications. Nevertheless, some of them have real-time capabilities for comprehensive subspace computing, such as PCP and Re ProCS.

The correct studies are usually based on a large dataset that is comprehensive enough to guarantee a fair assessment. The scientists used either available databases or their own datasets most of the time. In this segment, we rapidly survey available public data sets throughout order to evaluate algorithms in conditions similar to those actual. A number of data sets are required for visual analysis of human activities.

IV. CONCLUSION

We discussed first of all the software used to identify static or shifting objects of interest with a context subtraction. Next, we discussed the problems associated with various situations and traveling objects. So the following results can be concluded:–All these examples highlight the value of moving object detection in images, as tracking’s, recognition or behavioral analysis are the first step. Therefore, for real-time implementations, the front mask has to be as accurately and as quickly as possible. As the conditions are very diverse, after the submission, the context paradigm will contend with various challenges. In fact, the moving objects of concern have very distinct design characteristics. Thus different background models must be built for a particular application or a common background model can be identified that can be used in all applications. For a universal background model, a specific background model can be created with particular challenges, and when the related difficulties are found, the appropriate background model can be used. Basic models are versatile enough for applications such as optical motion capture in indoor scenes in controlled environments. Statistical models provide a good framework for traffic monitoring, but special developments must be added to challenges such as lighting and sleeping beginning foreground objects. The main conclusion is that the BGS approaches used in modern software structures are usually slightly better by old methods. In natural environments, especially marine and aquatic circumstances. Next, engineers are often restricted to ancient well-known methods with available public implementations since they are pressured.

REFERENCES

1. N. Jain, R. Saini, P. Mittal, A review on traffic monitoring system techniques, Soft Comput.: Theories Appl. (2018) 569–577.
2. C. Stauffer, E. Grimson, Adaptive background mixture models for realtime tracking, in: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 1999, 1999, pp. 246–252.
3. J. Pulgarin-Giraldo, A. Alvarez-Meza, D. Insuasti-Ceballos, T. Bouwmans,G. Castellanos-Dominguez, GMM Background modeling using divergence based weight updating, in: Conference Ibero American Congress on Pattern Recognition, CIARP 2016, 2016.
4. L. Martins, P. Carvalho, L. Corte-Real, J. Alba-Castro, BMOG: Boosted Gaussian mixture model with controlled complexity, in: BPRIA 2017,2017.
5. X. Lu, C. Xu, L. Wang, L. Teng, Improved background subtraction method for detecting moving objects based on GMM, IEEJ Trans. Electr. Electron.Eng. (2018).
6. S. Ali, K. Goyal, J. Singhai, Moving object detection using self adaptive Gaussian mixture model for real time applications, in: International Conference on Recent Innovations in Signal processing and Embedded
7. Systems, RISE 2017, 2017, pp. 153–156.
8. C. Postigo, J. Torres, J. Menendez, Vacant parking area estimation through background subtraction and transience map analysis, IET Intell. Transp. Syst. 9 (9) (2015) 835–841.
9. M. Neuhausen, Video-based Parking Space Detection: Localisation of Vehicles and Development of an Infrastructure for a Routing System, Institut fur Neuro informatik, 2015.
10. J. Cho, J. Park, U. Baek, D. Hwang, S. Choi, S. Kim, K. Kim, Automatic parking system using background subtraction with CCTV environment in: International Conference on Control, Automation and Systems, IKCAS 2016, 2016, pp. 1649–1652.
11. A. Kushwaha, R. Srivastava, A framework of moving object segmentation in maritime surveillance inside a dynamic backgrounds, Trans. Comput. Sci. (2015) 35–54.
12. M. Razif, M. Mokji, M. Zahidi, Low complexity maritime surveillance video using background subtraction on h.26, in: International Symposium on Technology Management and Emerging Technologies, ISTMET 2015, 2015,pp. 361–368.
13. J. Zhao, S. Pang, B. Hartill, A. Sarrafa-zadeh, Adaptive background modeling for land and water composition scenes, in: International Conference on Image Analysis and Processing, ICIAP 2015, 2015.
14. V. Fernandez-Aguedas, G. Pallotta, M. Vespe, Maritime traffic networks: From historical positioning data to unsupervised maritime traffic monitoring, IEEE Trans. Intell. Transp. Syst. 19 (3) (2018) 722–732.
15. A. Leykin, M. Tuceryan, Detecting shopper groups in video sequences, in: International Conference on Advanced Video and Signal-Based Surveillance, AVSBS 2007, 2007.
16. A. Leykin, M. Tuceryan, A vision system for automated customer tracking for marketing analysis: low level feature extraction, in: HAREM 2005,2005.
17. A. Leykin, M. Tuceryan, Tracking and Activity Analysis in Retail Environments, TR 620 Computer Science Department, Indiana University, 2005.
18. N. Aminash, M.S. Kumar, S. Sagar, Automated video surveillance for retail store statistics generation, in: International Conference on Signal and Image Processing, ICSIP, 2012, pp. 585–596.
19. V. Nogueira, H. Oliveira, J. Silva, T. Vieira, K. Oliveira, RetailNet: A deep learning approach for people counting and hot spots detection in retail stores, in: SIBGRAPI 2019, 2019.
20. U. Knauer, M. Himmelsbach, F. Winkler, F. Zautke, K. Bienefeld, B. Meffert, Application of an adaptive background model for monitoring honeybees: in: VIIIP 2005, 2005.
21. Belmar Garcia-Garcia a, Thierry Boumans b,*, Alberto Jorge Rosales Silva: Background subtraction in real applications: Challenges, current models and future directions, 2019 elsevier

AUTHOR PROFILE

Dr. T.Logeswari is a Associate Professor, Department of Computer Applications, New Horizon College, Bangalore. She obtained her Bachelor’s degree in Applied science with Computer Technology from Bharathiar University during 2002, Masters Degree in Computer Applications from Alagappa University during 2006 and M.Phil in Computer Applications from Alagappa University, during 2007. She has completed her PhD in Computer Science at Mother Teresa Women’s University during November 2012. Her specialized areas are Image Processing, Pattern Recognition and Neural Network. Her Ph.D work is focused on Investigations on Intelligent system to detect Brain tumor in the earlier Stage. She has presented many research papers in National, International conferences and Journals.