Moving Object Detection in Night Time: A Survey

Anu Singha, Mrinal Kanti Bhowmik
Department of Computer Science and Engineering
Tripura University (A Central University)
Agartala-799022, India
anusingh5012@gmail.com, mrinalkantibhowmik@tripurauniv.in

***Abstract*—The object monitoring performance is greatly depends on the enhanced quality of images. The quality of these images is affected at night video systems due to low contrast of salient objects as well as several atmospheric conditions such as fog, rain, dust etc. As a consequence, the intensity, color, polarization, coherence of scenes alters. The awful appearance of night images under subjective lighting and atmospheric conditions is a general problem for analysis in computer vision. In this paper, we have aimed to provide a survey work to researchers under four categories: (i) review on change detection datasets either thermal or visual-thermal, (ii) review on feature extraction based object detection techniques, (iii) review on background modeling based object detection techniques, (iv) review on convolutional neural network (CNN) based object detection techniques. This paper will benefits for those who require a details review over change detection datasets and detection techniques for moving object detection related applications.**

Keywords—night vision, object detection, feature extraction, background model, deep learning

I. Introduction

Night object detection is important in computer vision due to its still challenging issues specially for military border intruder monitoring, and advance driver assistance systems (ADAS) [1][2][3]. Most automatic night vision systems for monitoring intelligently moving objects presume that the input images have clear visibility under lane light but unfortunately this does not ensue all the time [4]. The quality of these images is affected by several atmospheric conditions that alter the key characteristics (i.e. intensity, color, polarization, coherence) of light source due to scattering by medium aerosols [5][6]. Although computer vision systems have benefit from success in indoor or outdoor environments at day time but facing issues in outdoor night environments. There a good solution is very necessary because the major safety problem from vehicle-objects collision that darkness causes [5][6]. In the last few decades, large datasets are designed to meet the increasing demands in developing new models for night object detection under bad atmospheric conditions [7][8]. However, there is still a lack of video datasets for moving object detection tasks that can provide a balanced coverage in atmosphere degraded outdoor scenes, especially at night.

Furthermore, detecting moving objects using a visual digital camera or a normal charge-coupled device (CCD) camera generally have advantage of high resolution which is more suitable at day time or night time with proper lighting setup. However, they are ineffective in environments with poor illumination or visibility such as total darkness as well as due to bad atmosphere because appearance of objects in the captured images is not as clear as in images captured during the normal atmosphere [9][4]. To address the limitations of visual and CCD camera at night time, many studies have been conducted on methods that detect objects with infrared based cameras [10][11][12][13][14] – including near-infrared (NIR) and far-infrared (FIR) camera. The NIR cameras are robust against darkness, and they are less costly than FIR cameras. But the NIR have similar drawback faced by CCD camera when the interferences produced by vehicle headlights. As well, the attenuation of visual, CCD, and NIR radiation produced through atmospheric aerosols is more due to their short wavelengths. In contrast, FIR cameras allow robust object detection regardless of the atmospheric conditions because as the spectrum wavelength increases the affect of bad atmospheric will be decreased [15]. Far less research has been carried out on moving object detection at night time under different atmospheric conditions using thermal images because of the high price of a FIR camera.

In this paper, we aim to provide a survey work to researchers who require a details review over change detection datasets and detection techniques for moving object detection related applications. The whole paper is organized as; the object detection dataset literature work is introduced in Section II. In Section III, the methodological review on the moving object detection techniques is described. Section IV provides a review on different feature, background model, and deep learning based object detection techniques. And finally, section V concludes the paper.

II. Review on CHANGE DETECTION DATASETS

Frame based object detection is interrelated with video based object detection, background subtraction, and moving object segmentation. Several datasets have been designed in the past to evaluate moving object detection methods. The Table I provides a summary of existing object detection database either thermal or visual-thermal in day or night time. Among these datasets, four have been recorded with thermal sensors to detect and tracking objects (i.e. OSU-T, BU-TIV, LTIR, ASL-TID), where BU-TIV dataset is primarily designed for visual analysis tasks. These datasets only contain day-time video sequences with challenges of cluttered background, occlusion, static and moving camera, object size variation; where OSU-T dataset includes weather conditions with low resolution thermal camera to detect only pedestrian.

Next numerous datasets (LITIV, AIC-TV, OSU-CT, CVC-14, KAIST, CDNet 2012, CDNet 2014) contains both colour and thermal video sequences, few of them (LITIV, OSU-CT, KAIST) works on fusion between two modalities to robust detection. The night video sequences contains in AIC-TV, CV-14, KAIST, and CDNet 2014. These datasets consist of various challenges such as scale variations, lighting conditions, dynamic backgrounds, shadows, camera jitter, low frame rate, turbulence, but very rare datasets consider weather conditions except CDNet 2014 although day time. As a consequence, it is difficult to evaluate robustness of object detection methods in atmospheric

978-1-7281-0744-8/19/$31.00 ©2019 IEEE

TABLE I. REVIEW ON CHANGE DETECTION DATASETS FOR MOVING OBJECT DETECTION

Modality	Dataset	Purpose	Object Category	Number of Video Sequences	Number of Frames	Frame Resolution	Ground Truth Type	Environment Condition	Key Challenges
Thermal	OSU-T [16], 2005	Person detection in thermal imagery	Pedestrian	10	284	360x240	Bounding Boxes	Outdoor	Light Rain, Cloudy, Haze
	BU-TIV [17], 2014	Address several visual analysis tasks in thermal infrared videos	Pedestrian, Runner Car, Bicycle, Motorcycle, Bat	16	63,782	512x512 to 1024x1024	Bounding Boxes	Outdoor	Single/Multi-view Single/Multi Object Tracking, Counting, Group Motion
	ASL-TID [18], 2014	Designed for Object Detect, not tracking	Pedestrian, Cat, Horse	4,381	8	324x240	Bounding Box	Outdoor	Moving Camera, cluttered background, occlusion
	LTIR [19], 2015	Object tracking	Rhinoceros, Human, Horse, Car, Dog, Quadrocopter	20	11,269	320x240 to 1920x480	Bounding Box	Indoor and Outdoor	static, hand-held, moving camera, cluttered background, occlusions, size-change
Visual-Thermal	LITIV [20][21], 2012,2015	People tracking	Pedestrian	9	6,236	320X240	Polygons	Indoor	thermal–visible video registration, sensor fusion
	AIC-TV [22], 2006	Object tracking	People, bicycles, faces, motorbike	6	2,013	-	-	Indoor and Outdoor	Scale variation, dark night-time, occlusion
	OSU-CT [23], 2007	Fusion-based object detection	Pedestrian	6	17,089	320x240	-	Outdoor	Color-Thermal image registration
	CVC-14 [24], 2016	Object detection	Pedestrian	2	17,036	Upto 1280x1024	Bounding Box	Outdoor	Varying lighting conditions at day and night time
	KAIST [25], 2015	Object detection	Pedestrian	-	95,000	Upto 640x480	Bounding Box	Outdoor	Align Color-Thermal Image pairs, Varying lighting conditions at day and night time
	CDNet 2012 [26]	Object detection	Boats, cars, trucks, and pedestrians	31	Nearly 90,000	320x240 To 720x576	Pixel-wise labeling	Indoor and Outdoor	Dynamic background, Camera Jitter, Shadows, Intermittent Object Motion, Thermal, Baseline
	CDNet 2014 [27]	Object detection	Boats, cars, trucks, and pedestrians	53	1,60,000	320x240 to 720x4 86	Pixel-wise labeling	Indoor and Outdoor	Dynamic background, Camera Jitter, Shadows, Intermittent Object Motion, Thermal, Bad Weather, Low Framerate, Night videos, PTZ, Baseline, Turbulence

conditions especially in night vision because more than half of object related accidents occur in the night time.

Therefore, there is need of designing a several atmospheric weather degraded conditions based standard video dataset in night vision that cover many real-world scenarios. The considerable atmospheric conditions may be rainy, foggy, dust and so on. The preferable sensor will be far-infrared thermal sensor to take the advantages of the higher wavelength, which will be less affected in such adverse weather conditions along with being able to see in total darkness in night vision.

III. REVIEW ON OBJECT DETECTION TECHNIQUES

Object detection is one of mature and challenging research in computer vision. Over the decades, the object detection methods those are using for visual or day-time frames, are

TABLE II. REVIEW ON FEATURE BASED OBJECT DETECTION TECHNIQUES

Name of the Publisher	Author	Proposed Method	Used Databases
IEEE Trans. PAMI, 2002	T. Ojala, et al [28]	LBP	Video sequences name not mentioned
IET Intell. Transp. Syst., 2015	P. Hurney, et al [29]	HOG-LBP	Their own database (2000 FIR images, 15000 frames of captured FIR video)
Optics Letters, OSA, 2012	A.Ko, et al [30]	CS-LBP	Human and Non-Human Thermal Images
Proc. CGIV, 2014	M. R. Jeong, et al [31]	OCS-LBP	Pedestrian and Non- Pedestrian Thermal Images
Comput. Vis. Image Underst. Elsevier, 2007	A.Dai, et al. [32]	Generalized Expectation-Maximization (EM)	OSU Infrared Image Database and WVU Infrared Video Database
Patt. Recog. Ltt., Elsevier, 2012	J. Wang, et al [33]	Shape Describer	OSU Color-Thermal Database
Patt. Recog., Elsevier, 2015	X.Y. Zhao, et al [34]	SDH	OSU thermal pedestrian database
Proc. CVPR, 2005	N. Dalal, et al [35]	HOG	MIT Pedestrian Database
IEEE Trans. PAMI, 2012	P. Dollar, et al [36]	16 representative state-of the-art pedestrian detectors	Caltech Pedestrian Database
Proc. PSIVT, 2009	T. Watanabe, et al [37]	CoHOG	DaimlerChrysler pedestrian classification benchmark dataset INRIA person data set
Proc. TISC, 2014	B. Qi, et al [38]	Scattered Difference of Directional Gradients (SDDG)	Thermal Corridor Dataset
IEEE Trans. Intell. Transp. Syst., 2015	J. Kim, et al [39]	PI-HOG	Caltech, IR, Pittsburgh, and Kitti datasets
Sensors, MDPI, 2017	J. Baek, et al [40]	TPI-HOG	KAIST pedestrian dataset

also using for thermal or night-time frames. A number of holistic approaches have been anticipated those are using specific features with combination of popular image processing techniques [28-40]. They are very successful but their performances degrade when the shape is deformed. To address these problems, various background models and systems have been developed [41-52]. Recently, few popular deep learning approaches to detecting moving objects work have been carried out in the literature over last half decades [53-57]. Therefore, the reviews on object detection techniques are arranged in three subsections in Table II, III, IV.

A. *Feature based Object Detection Techniques*

Table II lists the undersized structure of reviews on feature based object detection methodologies which are released over the years. Automated night-time video analysis is significant for many vision applications which have a huge demand in the field of security and surveillance of defense, protected area, metrological applications. Different researchers have proposed different techniques for detecting the objects over night conditions. Similar to object detection (OD) using visible images, OD using FIR images also consists of feature extraction step. The features developed for day-time OD can also be used for night-time OD. For example, local binary patter (LBP) [28] and its variations, such as the HOG-LBP [29], center-symmetric LBP (CSLBP) [30], and oriented CSLBP (OCSLBP) [31] were also proposed as day-time OD features. However, the LBP-based features have only orientation information of pixel intensity; therefore, they are sensitive to lighting conditions. On the other hand, there are some methods that use the shape of objects as features. Dai et al. [32] utilized the joint shape and appearance cue to find the exact locations of objects. Wang et al. [33] extracted the features using a shape describer and Zhao et al. [34] proposed the shape distribution histogram (SDH). These shape-based features simply used only pixel intensity information and employed background subtraction methods for fixed camera images. Therefore, shape-based features are not suitable for traffic environment where complex background is not fixed.

As a robust feature for object detection, the histograms of oriented gradient (HOG) [35] is one of the most popular OD features and its variations have been proposed [36]. Co-occurrence HOG (CoHOG) is one of the extensions of the HOG and it utilizes pair of orientations for computing histogram feature [37]. Scattered difference of directional gradient (SDDG) that extracts local gradient information along the certain direction is also proposed for IR images [38]. Kim et al. proposed position-intensity HOG (PIHOG) that includes not only HOG but also the detail position and intensity information for vehicle detection [39]. Theses HOG based features utilize only the gradient information based on color images or do not consider the thermal intensity information which is important cue for object detection in night-time. To address these problems of conventional features, J. Baek et al. [40] propose thermal position intensity HOG (TPIHOG). The TPIHOG is the extended version of PIHOG and it is applied for pedestrian detection in night-time.

B. *Background Model based Object Detection Techniques*

A survey on background model based object detection via background subtraction methodologies are listed in Table III. Background subtraction is generally regarded as an effective method for extracting the foreground, and it has moved forward from simply comparing a static background frame with current frame to establishing a sophisticated reference model of the scene with periodic updates. The main objective of reference model construction is to obtain an effective and efficient background model for foreground

TABLE III. REVIEW ON BACKGROUND MODEL BASED OBJECT DETECTION TECHNIQUES

Name of the Publisher	Author	Proposed Method	Used Databases
Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 1999	C. Stauffer, et al [41]	GMM	Video sequences name not mentioned
Proc. 6th European Conf. on Comput. Vis., 2000	A. M. Elgammal, et al [42]	KDE	5 video clips from 'ftp://www.umiacs.umd.edu/pub/elgammal/video/index.htm'
Pattern Recogn., Elsevier, 2007	H. Wang, et al [43]	SACON	Video sequence names not mentioned
IEEE Trans. Image Process., 2011	Droogenbroeck, et al [44]	ViBe	The "house" sequence and PET2001 sequence
Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2012	M. Hofmann, et al. [45]	PBAS	CDnet 2012
IEEE Trans. on Img. Process., 2015	P.L. St-Charles, et al [46]	SuBSENSE	CDnet 2012 and CDnet 2014
IEEE Winter Conf. on Appl. of Computer Vis., 2014	P.L. St-Charles, et al [48]	LOBSTER	CDnet 2012
Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2012	L. Maddalena, et al [49]	SOBS	CDnet 2012
Int. J. Electron. Commun., Elsevier, 2010	M. Wu, et al [50]	CodeBook	Video sequence names not mentioned
IEEE Trans. PAMI, 2000	N.M. Oliver, et al [51]	EigenBackground	Video sequence names not mentioned
IEEE Trans. Broadcasting, 2011	F.C. Cheng, et al [52]	ISBM	The "lobby" and "light switch" sequences

object detection. The motive of these methods was to determine a given pixel intensity value is true positive or not. A very prominent parametric method called Gaussian Mixture Models (GMM) proposed by Stauffer *et al.* [41] is generally modeled each background pixels through a mixture of Gaussian probability density functions by an iterative update rule. Another density based estimation method Kernel Density Estimation (KDE) introduced by Elgammal *et al.* [42] has been successfully applied in background segmentation. The KDE is a non-parametric model which is estimate background probability density functions via directly on local intensity observations.

The history of background sample based strategy is introduced by Wang *et al.* [43] in Sample Consensus (SACON) is defined at each pixel by an account of the *N* most recent pixel intensity samples. This method is also a non-parametric. However, most of these methods are unable to model long-term periodic events because of their observations based on a first-in, first-out strategy. Overcome this problem, random observation replacement strategy introduced in [44][45] to their background models. In [44], Droogenbroeck *et al.* proposed another non-parametric method called VIBE adopt random approach to update background pixels as well as diffuse their current pixel value into neighboring pixel. The main drawback of VIBE method is that it follows a global fixed parameters strategy for model maintenance which is facing problem of dynamic variations of real-world scenes. Hofmann *et al.* [45] proposed a feedback scheme Pixel-based Adaptive Segmenter (PBAS) to monitor background dynamics at the pixel level through adaptive state variables.

In background modeling, there have been numerous attempts to fuse the advantages of pixel-based and spatial-based in generation of background model to control both change detection sensitivity and dynamic background scenes in real world. One of very well-known method in this category is Self-Balanced Sensitivity Segmenter (SuBSENSE) proposed by St-Charles *et al.* [46], where integration of LSBP [47] features and PBAS [45] feed-back model has applied to improve spatiotemporal sensitivity. The authors also extend their work with a few general improvements in method Local Binary Similarity Segmenter (LOBSTER) [48].

There is several other well-known background subtraction for moving object detection available in the literature. Maddalena et al. [49] presented a self-organizing Artificial Neural Network for background subtraction (SOBS) to handle gradual illumination variations and camouflage. The Codebook methods of [50] present cluster observations into code words and store them in local dictionaries. The Eigen value decomposition based background model was first proposed by Oliver *et al.* [51]. An illumination-sensitive background modeling approach introduced by F.C. Cheng *et al.* [52] to analyze the illumination change and detect moving objects.

C. Deep Learning based Object Detection Techniques

A survey on convolutional neural networks based object detection methodologies are listed in Table IV. A popular term i.e. deep learning approaches on detecting moving objects which has been recently studied in the literature over only the last half decades. To the best of our knowledge, M. Braham *et al.* [53] are the first attempt to apply convolutional neural networks (CNN) to the background subtraction problem. This paper is not intended to present a real-time and adaptive technique, but rather to investigate the classification potential of deep features learned with convolutional neural networks (ConvNets) for the background subtraction task. The CNN can only perform satisfying background subtraction on a single scene (that was trained with scene specific data) and also lacks the ability to perform the segmentation in real time. Instead M. Babaee *et al.* [54] proposed an approach which yields a universal network that can handle various scenes without having to retrain it every

TABLE IV. REVIEW ON DEEP LEARNING BASED OBJECT DETECTION TECHNIQUES

Name of the Publisher	Author	Proposed Method	Used Databases
Proc. IWSSIP, 2016	M. Braham and M.V. Droogenbroeck [53]	Convolutional Neural Networks (ConvNets)	CDnet 2014
arXiv preprint arXiv:1702.01731, 2017	M. Babaee, et al [54]	Deep Convolutional Neural Networks, Subsense, Flux Tensor	CDnet 2014
Patt. Recog. Ltt., Elsevier, 2017	Yi Wang, et al [55]	Multiresolution CNN, Cascaded Architecture	CDnet 2014
IEEE Geoscience & Remote Sensing Ltt., 2018	M. R. Jeong, et al [56]	Multiscale Fully Convolutional Network (MFCN) Architecture	IR Image Sequences
Neurocomputing, Elsevier, IF-3.317, 2015	Y. Zhang, et al. [57]	Deep Learning, Hash Method, Binary Hamming Code	CDnet 2014

time the scene changes. First, the author generates the basic background model based on some traditional algorithms (SuBSENSE and Flux Tensor). Then the author constructs a CNN model, followed by some post-processing methods, to get the final foreground mask. Yi Wang *et al.* [55] explore various CNN configurations such as a multiresolution CNN and a cascaded architecture. As well Dongdong Zeng *et al.* [56] also propose a novel multi-scale fully convolutional network (MFCN) architecture for IR foreground object detection. Yaqing Zhang *et al.* [57] proposed a highly effective and efficient method for moving object detection based on binary scene modeling. The proposed method takes a stacked autoencoder-based deep learning scheme to adaptively build a robust feature representation. Furthermore, the binary scene model captures the spatio-temporal scene distribution information in the Hamming space, which ensures the high efficiency of moving object detection.

IV. Conclusion

In this paper, we have provided a review work under four categories. From these reviews, we have concluded with few points. First the reviews on change detection datasets, it has highlighted that there is a need of purely night video dataset to a balanced coverage in atmosphere degraded outdoor scenes for analysing night sequences. Secondly reviews on feature and background subtraction based object detection techniques, the combination of spatial feature and pixel based background model will give robust discriminative power to distinguish between foreground and background. Third, it is a point in time to explore more on convolutional neural network (CNN) in background segmentation for object detection related applications.

Acknowledgment

The work presented here is being conducted in the Biometrics Laboratory of Computer Science and Engineering Department of Tripura University (A central university), Tripura, Suryamaninagar-799022. The research work was supported by Defence Research and Development Organisation (DRDO), Government of India, of Project Grant No. ERIP/ER/DG-MED & CoS/990116507/M/01/1716, Dated: 06/03/2018.

References

[1] D. Gerónimo, A.M. López, A.D. Sappa, T. Graf, ”Survey of pedestrian detection for advanced driver assistance systems,” IEEE Trans. Pattern Anal. Mach. Intell., Vol. 32, pp. 1239–1258, 2010.

[2] M. Enzweiler, D.M. Gavrila, “Monocular pedestrian detection: Survey and experiments,” IEEE Trans. Pattern Anal. Mach. Intell., Vol. 31, pp. 2179–2195, 2009.

[3] P. Dollár, C. Wojek, B. Schiele, P. Perona, “Pedestrian detection: An evaluation of the state of the art,” IEEE Trans. Pattern Anal. Mach. Intell., Vol. 34, pp. 743–761, 2012.

[4] R.T. Tan, "Visibility in bad weather from a single image," IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008), IEEE, pp. 1-8, 2008.

[5] S.G. Narasimhan and S.K. Nayar, "Vision and the atmosphere," International Journal of Computer Vision, Springer, Vol. 48, No. 3, pp. 233-254, 2002.

[6] C.C.Chen, ‘Attenuation of Electromagnetic Radiation by Haze, Fog, Clouds, and Rain’, A report prepared by UNITED STATES AIR FORCE PROJECT RAND, 1975.

[7] J. Davis and M. Keck. A two-stage approach to person detectionin thermal imagery. In Proceeding of Workshop onApplications of Computer Vision (WACV), 2005.

[8] Y. Wang, P.-M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, *CDnet 2014: An Expanded Change Detection Benchmark Dataset*, in Proc. IEEE Workshop on Change Detection (CDW-2014) at CVPR-2014, pp. 387-394. 2014.

[9] S.Mahlke, D. Rösler, K. Seifert, J. F. Krems, andM. Thüring, “Evaluation of six night vision enhancement systems: Qualitative and quantitative support for intelligent image processing,” *Hum. Factors*, vol. 49, no. 3, pp. 518–531, Jun. 2007.

[10] J. Ge, Y. Luo, and G. Tei, “Real-time pedestrian detection and tracking at nighttime for driver-assistance systems,” *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 2, pp. 283–298, Jun. 2009.

[11] J. Han, A. Gaszczak, R. Maciol, S. E. Barnes, and T. P. Breckon, “Human pose classification within the context of near-IR imagery tracking,” in *Proc. 9th SPIE 8901 Opt. Photon. Counterterrorism, Crime Fight. Def.*, 2013, pp. 1–11.

[12] M. Bertozzi *et al.*, “IR pedestrian detection for advanced driver assistance systems,” in *Proc. 25th DAGM Symp. Pattern Recognit.*, Sep. 2003, pp. 582–590.

[13] A. Treptow, G. Cielniak, and T. Duckett. “Real-time people tracking for mobile robots using thermal vision,” *Robot. Autonom. Syst.*, vol. 54, no. 9, pp. 729–739, 2006.

[14] C. Dai, Y. Zheng, and X. Li, “Pedestrian detection and tracking in infrared imagery using shape and appearance,” *Comput. Vis. Image Understand.*, vol. 106, pp. 288–299, 2007.

[15] C.C.Chen, ‘Attenuation of Electromagnetic Radiation by Haze, Fog, Clouds, and Rain’, A report prepared by UNITED STATES AIR FORCE PROJECT RAND, 1975.

[16] J. Davis and M. Keck. A two-stage approach to person detectionin thermal imagery. In Proceeding of Workshop onApplications of Computer Vision (WACV), 2005.

[17] Z. Wu, N. Fuller, D. Theriault, and M. Betke. A thermal infrared video benchmark for visual analysis. In Proceeding of 10th IEEE Workshop on Perception Beyond the Visible Spectrum (PBVS), 2014.

[18] J. Portmann, S. Lynen, M. Chli, and R. Siegwart. People detection and tracking from aerial thermal views. In Proceeding of IEEE International Conference on Robotics and Automation (ICRA), 2014.

[19] A. Berg, J. Ahlberg, and M. Felsberg. A thermal object tracking benchmark. 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2015.

[20] A. Torabi et al. An iterative integrated framework for thermal-visible image registration, sensor fusion, and people tracking for video surveillance applications. Comp. Vis. and Im. Underst., 116(2):210 – 221, 2012.

[21] St-Charles, P.-L., Bilodeau, G.-A., Bergevin, R., Online Multimodal Video Registration Based on Shape Matching, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops)*, Boston, MA, USA, June 7-12, 2015

[22] C. Conaire, N. O'Connor, E. Cooke, and A. Smeaton, "Comparison of fusion methods for thermo-visual surveillance tracking," in *Proc. IEEE Conf. Information Fusion*, 2006.

[23] J. W. Davis and V. Sharma. Background-subtraction using contour-based fusion of thermal and visible imagery. CVIU, 106(2-3):162–182, 2007.

[24] A. González, Z. Fang, Y.Socarras, J.Serrat, D.Vázquez,J.Xu, and A. M. López. Pedestrian Detection at Day/Nighttime with Visible and FIR Cameras: A Comparison. Sensors **2016**, 16, 820.

[25] S. Hwang, J. Park, N. Kim, Y. Choi, and I. S. Kweon. Multispectral Pedestrian Detection: Benchmark Dataset and Baseline. CVPR, pp. 1037-1044, 2015.

[26] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, *changedetection.net: A new change detection benchmark dataset*,
in Proc. IEEE Workshop on Change Detection (CDW-2012) at CVPR-2012, Providence, RI, 16-21 Jun., 2012

[27] Y. Wang, P.-M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, *CDnet 2014: An Expanded Change Detection Benchmark Dataset*, in Proc. IEEE Workshop on Change Detection (CDW-2014) at CVPR-2014, pp. 387-394. 2014

[28] Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Trans. Pattern Anal. Mach. Intell. **2002**, 24, 971–987.

[29] Hurney, P.; Waldron, P.; Morgan, F.; Jonses, E.; Glavin, M. Night-time Pedestrian Classification with Histograms of Oriented Gradients-Local Binary Patterns Vectors. IET Intell. Transp. Syst. **2015**, 9, 75–85.

[30] Ko, B.; Kim, D.; Nam, J. Detecting Humans using Luminance Saliency in Thermal Images. Opt. Lett. **2012**, 37, 4350–4352.

[31] Jeong, M.R.; Kwak, J.; Son, E.; Ko, B.; Nam, J. Fast Pedestrian Detection using a Night Vision System for Safety Driving. In Proceedings of the 2014 International Conference on Computer Graphics, Imaging and Visualization (CGIV), Singapore, 6–8 August 2014.

[32] Dai, C.; Zheng, Y.; Li, X. Pedestrian Detection and Tracking in Infrared Imagery Using Shape and Appearance. Comput. Vis. Image Underst. **2007**, 106, 288–299.

[33] Wang, J.; Chen, D.; Chen, H.; Yang, J. On Pedestrian Detection and Tracking in Infrared Videos. Pattern Recognit. Lett. **2012**, 33, 775–785.

[34] Zhao, X.Y.; He, Z.X.; Zhang, S.Y.; Liang, D. Robust pedestrian detection in thermal infrared imagery using a shape distribution histogram feature and modified sparse representation classification. Pattern Recognit. **2015**, 48, 1947–1960.

[35] Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 20–26 June 2005; pp. 886–893.

[36] Dollar, P.; Wojek, C.; Schiele, B.; Perona, P. Pedestrian Detection: An Evaluation of the State of the Art. IEEE Trans. Pattern Anal. Mach. Intell. **2012**, 34, 743–761.

[37] Watanabe, T.; Ito, S.; Yokoi, K. Co-occurrence histograms of oriented gradients for pedestrian detection. In Proceedings of the 2009 Conference on Pacific-Rim Symposium on Image and Video Technology, Tokyo, Japan, 13–16 January 2009; pp. 37–47.

[38] Qi, B.; John, V.; Liu, Z.; Mita, S. Pedestrian detection from thermal images with a scattered difference of directional gradients feature descriptor. In Proceedings of the 17th IEEE Conference on Intelligent Transportation System (ITSC), Qingdao, China, 8–11 October 2014; pp. 2168–2173.

[39] Kim, J.; Baek, J.; Kim, E. A Novel On-road Vehicle Detection Method Using PIHOG. IEEE Trans. Intell. Transp. Syst. **2015**, 16, 3414–3429.

[40] Baek, J.; Hong, S.; Kim, J.; Kim, E. Efficient Pedestrian Detection at Nighttime Using a Thermal Camera. Sensors **2017**, 17, 1850–1870.

[41] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2, 1999, pp. –252 Vol. 2.

[42] A. M. Elgammal, D. Harwood, and L. S. Davis, "Non-parametric model for background subtraction," in *Proc. 6th European Conf. on Comput. Vis.*, 2000, pp. 751–767.

[43] H. Wang and D. Suter. A consensus-based method for tracking: Modelling background scenario and foreground appearance. *Pattern Recognition*, 40(3):1091–1105, 2007.

[44] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1709–1724, 2011.

[45] M. Hofmann, P. Tiefenbacher, and G. Rigoll, "Background segmentation with feedback: The pixel-based adaptive segmenter," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2012, pp. 38–43.

[46] P.L. St-Charles, G.A.Bilodeau, and R.Bergevin, 'SuBSENSE: A Universal Change Detection Method with Local Adaptive Sensitivity', IEEE Trans. on Image Processing, Vol. 24, pp. 359-373, 2015.

[47] G.-A. Bilodeau, J.-P. Jodoin, and N. Saunier. Change detection in feature space using local binary similarity patterns. In *Computer and Robot Vision (CRV), 2013 International Conference on*, pages 106–112, 2013.

[48] P.L. St-Charles and G.A.Bilodeau, 'Improving Background Subtraction using Local Binary Similarity Patterns', IEEE Winter Conference on Applications of Computer Vision, 2014.

[49] L. Maddalena and A. Petrosino. The SOBS algorithm: What are the limits? In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 21 – 26, 2012.

[50] M. Wu and X. Peng, "Spatio-temporal context for codebook-based dynamic background subtraction," *Int. J. Electron. Commun.*, vol. 64, no. 8, pp. 739 – 747, 2010.

[51] Oliver, N.M., Rosario, B., Pentland, A.: A bayesian computer vision system formodeling human interactions. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8,pp. 831–843, 2000.

[52] F.C. Cheng, S.C. Huang, and S.J. Ruan, 'Illumination-Sensitive Background Modeling Approach for Accurate Moving Object Detection' IEEE Trans. On Broadcasting, Vol. 57, No. 4, pp. 794-801, 2011.

[53] Braham, M.; Droogenbroeck, M.V. Deep Background Subtraction with Scene Specific Convolutional Neural Networks. In IEEE International Conference on Systems, Signals and Image Processing (IWSSIP), Bratislava, Slovakia, pp. 1-4, 2016.

[54] Babaee, M.; Dinh, D.T.; Rigoll, G. A deep convolutional neural network for background subtraction. arXiv preprint arXiv:1702.01731, 2017.

[55] Wang, Y.; Luo, Z.; Jodoin, P.M. Interactive Deep Learning Method for Segmenting Moving Objects, Pattern Recognition Letters, 2016, doi: 10.1016/j.patrec.2016.09.014

[56] Zeng, D.; Zhu. M. Multiscale Fully Convolutional Network for Foreground Object Detection in Infrared Videos, IEEE Geoscience and Remote Sensing Letters, Vol. 15, 2018.

[57] Zhang, Y.; Li, X.; Zhang, Z.; Wu, F.; Zhao, L. Deep Learning Driven Blockwise Moving Object Detection with Binary Scene Modeling, Neurocomputing, 2105
http://dx.doi.org/10.1016/j.neucom.2015.05.082