Low Complexity FPGA Based Background Subtraction Technique for Thermal Imagery

Muhammad Imran, Mattias O'Nils, Huma Munir, Benny Thörnberg
Mid Sweden University
Holmgatan 10, SE-85170
Sundsvall, Sweden
{muhammad.imran,mattias.onils,benny.thornberg}@miun.se, humu1200@student.miun.se

ABSTRACT

Embedded smart camera systems are gaining popularity for a number of real world surveillance applications. However, there are still challenges, i.e. variation in illumination, shadows, occlusion, and weather conditions while employing the vision algorithms in outdoor environments. For safety-critical surveillance applications, the visual sensors can be complemented with beyond-visual-range sensors. This in turn requires analysis, development and modification of existing imaging techniques. In this work, a low complexity background modelling and subtraction technique has been proposed for thermal imagery. The proposed technique has been implemented on Field Programmable Gate Arrays (FPGAs) after in-depth analysis of different sets of images, characterizing poor signal-to-noise ratio challenges, e.g. motion of high frequency background objects, temperature variation and camera jitter etc. The proposed technique dynamically updates the background on pixel level and requires a single frame storage as opposed to existing techniques. The comparison of this approach with two other approaches show that this approach performs better in different environmental conditions. The proposed technique has been modelled in Register Transfer Logic (RTL) and implementation on the latest FPGAs shows that the design requires less than 1 percent logics, 47 percent block RAMs, and consumes 91 mW power consumption on Artix-7 100T FPGA.

Categories and Subject Descriptors

• **Reconfigurable logic and FPGAs** • **High-level language architectures. Computer vision problems.**

Keywords

Background modelling; subtraction; FPGA; architecture; smart camera, thermal imaging.

1. INTRODUCTION

Safety critical surveillance applications require development of robust and high performance embedded smart camera systems which provide easy integration into existing infrastructure and surveillance, irrespective of weather conditions [1][2][3]. The traditional processor based platforms perform serial computation and are memory bound, requiring storage after each imaging task. To achieve real-time performance, higher clock frequency and an increased number of cores will be required [4]. An alternative to a processor based solution is to use hardware platforms such as Field Programmable Gate Arrays (FPGAs) which offer advantages of re-configurability and inherit hardware parallelism in order to better exploit a regular flow of initial data centric tasks at lower clock frequency. This results in an improved performance for FPGA by using a small clock frequency as compared to processor based platforms [5]. However, the development time is greater. Therefore, a low complexity algorithm with small resource utilization needs to be investigated. The algorithms are expected to provide a good working design for FPGA.

In this work, we investigated and implemented a low complexity background modelling and subtraction technique, which is an important step in an imaging flow for surveillance applications. These applications require detection of moving objects that can be isolated from the background by using pre-processing techniques, e.g. temporal filtering, background subtraction and segmentation. In this regard, background subtraction is an important pre-processing task in the imaging dataflow for high level tasks such as classification and pose estimation. The existing background subtraction techniques have mostly been evaluated with respect to visual image sensors [6]. In outdoor surveillance applications, background modelling techniques for visual images are prone to produce false positives because of variation in illumination, the animated background, such as shadows or moving trees, shrubs and electrical wires [2]. These applications require background updating periodically in order to adjust the environmental variables.

The limitation of visual sensors to visible range of electromagnetic spectrum motivates integration of beyond-visual-range sensors such as infrared, x-rays and laser together with visual sensors. The combined information of these different sensors can give better situational awareness of the safety critical areas, such as railway tracks, to avoid accidents [7]. This work aims at integration of Infrared (IR) sensors together with visual sensors for an embedded smart camera in order to have better spectral, spatial and temporal resolutions. The advantage of infrared is that it is insensitive to variation in light and offers better contrast when it comes to separating foreground moving

Figure 1. Visual and thermal images of a person walking along a railway track.

objects from the environment. This will enable the system to detect foreground objects in different weather conditions, e.g. direct light, darkness, rain, snow and haze. For example, Figure 1 shows that in visual image, a person walking along a railway track is difficult to differentiate from background because of poor lighting conditions as compared to thermal image.

In this camera system, pre-processing steps are performed on thermal imaging and the compressed Region-of-Interest images are transmitted frequently to the end user. The compressed visual images are transmitted occasionally for detailed information. On the end user side, both visual and ROI of thermal are blended in order to highlight the activity in the region [3]. This approach relaxes the processing and the transmission requirements of the system.

As discussed earlier that this camera system would be implemented on FPGA, therefore a suitable background subtraction model will be explored for this platform. However, the challenges include design complexity, detection efficiency, memory footprints and resource requirements [8][9]. There is no model addressing all these issues. Therefore a model should be investigated with careful considerations because of aforementioned constraints [10][11].

In this work, two existing low complexity background modelling and subtraction techniques, progressive background generation [12] and background generation with a temporal low pass IIR filter [13][1] have been evaluated with respect to thermal imagery. Based on the analysis, the background modelling techniques have been modified to develop a hybrid technique which is then tested on different data sets. After in-depth analysis, the proposed technique is modelled at RTL level and implemented on FPGA. Following this section, the related work on background modelling and subtraction is presented.

2. RELATED WORK

Tarek *et al.* [2] proposed a background subtraction technique by combining thermal and visible imagery using Gaussian mixture models (GMM). However, the experiments are performed on high-end offline computers. The memory and computational requirements for embedded implementation is missing. GMM is widely used to model complex environments. However, the complexity and memory requirements for such a technique is high. For example, Tessens *et al.* [8] discussed that MOG with K components (K represents the number of mixture components, usually three to seven) requires $4K+11$ arithmetic operations per pixel and a minimum of $2K$ frames of memory storage.

Benezeth *et al.* [14] evaluated seven background subtraction techniques for colour and grayscale images on a standard computer and provides recommendations for real-time implementation. The authors conclude that there is a trade-off between speed, simplicity and efficiency for the investigated seven compression schemes. Reddy *et al.* [9] proposed a sequential technique for background estimation which does not require storage of multiple frames. The performance of the technique is claimed to be invariant for moderate illumination changes.

Chung *et al.* [12] developed a progressive background image generation by using a partial background generation technique. This technique requires histogram storage to record and trace the best background values and is expected to eliminate the unstable effect of physical vibration. Kandhalu *et al.* [13] used a subtraction technique which utilizes a first order recursive filter to integrate new incoming information into current background images. This technique updates the background dynamically in order to incorporate illumination changes. Most of the aforementioned techniques have been investigated on high-end computers by using visual images and require more than one frame for background generation. The FPGA based hardware implementation together with thermal imagery requires investigation of these techniques in the new context.

2.1 Background modelling and subtraction

In this work, we have investigated two low complexity background modelling techniques including progressive background generation [12] and background generation with a temporal low pass IIR filter [13][1]. In this work, we modified the two aforementioned background modelling techniques in order to develop the hybrid technique. The motivation is that the progressive background generation will eliminate the foreground objects because of different intensity regions of moving objects between frames. The low pass IIR filter will update the background image with a specified adaption rate, if the intensity difference is small. The analysis showed that the hybrid technique gives better results as shown in the results section. The proposed background subtraction model is presented here.

2.1.1 Background subtraction model

The hybrid background is represented by Eq.1

$$Bg(t) = \begin{cases} Mbg & \text{if } |In - Bg(t-1)| < \varepsilon \\ Lbg & \text{else} \end{cases} \quad (1)$$

$Mbg = \min(In(t),\ Bg(t-1))$ or $\text{avg}(In(t),\ Bg(t-1))$
$Lbg = \alpha In + (1-\alpha)Bg(t-1)$

Where In is current image, Bg is the generated background, Mbg represents background pixels when the difference between the current and the previous background is smaller than a constant ε. In this work, the thermal energy (heat signature) of moving people sensed by the thermal camera is greater than the temperature of the surrounding static objects. The static objects with heat temperature similar to humans will be considered as a part of the background because of small intensity difference between successive images. Therefore, to generate Mbg, we use a non-linear min filter. Other alternatives for calculating the Mbg value is averaging of the two pixel values, simply the value of the current or previous frame pixel value. Lbg is generated by using low pass first order recursive filter when the difference between the current and the previous background is greater than a constant ε. In this part, the background image is updated by integrating new incoming pixel data into the current background to adapt the temperature variation and motion changes caused by high frequency background objects like trees and shrubs etc. In this case, α is the adaptation coefficient and is selected as 0.05 in order to ensure that artificial tails behind the moving objects are not created [13][1].

2.1.2 Background subtraction

In background subtraction, the current frame is subtracted from the background frame on pixel level. This process will highlight the foreground objects by removing uneven temperature variation and background noise.

2.1.3 Tipping points of failure

In the summer of hot countries, it might become challenging, depending on the sensor size and type, to detect humans where

there are moving tree branches because the small difference in heat signatures. In outdoor environments, a camera jitter because of winds and the vibration of moving trains/vehicles might affect the detection results.

3. EXPERIMENTAL FRAMEWORK

The experimental work consists of two processes, (1) data collection and high level analysis, (2) FPGA implementation setup.

3.1 Data collection and high level analysis

For the experiments, video streams of natural scenes with human movement were recorded in different weather conditions, e.g. summer and winter in the city of Sundsvall, Sweden. The video data represents different challenges, such as poor signal-to-noise ratio because of temperature variation, camera jitter and motion of trees, wires, etc. From the video data, four sets of images (each set with 1000 images with a size of 352×240) were developed to analyse the background subtraction technique. The image data sets can be accessed online [15].

High level investigation was performed on the collected data by using Matlab. This investigation included analysis of individual background subtraction techniques, development and analysis of hybrid techniques, selection of ε in Eq.1 and the integration of segmentation and edge detection in order to identify the difference in foreground objects of the processed images. In relation to this, a number of experiments were performed on image datasets that captured the previously mentioned challenges.

3.2 FPGA implementation setup

The proposed background subtraction technique was modelled in RTL. The functionality was verified in behavioural and post-route simulations and results were compared with high level analysis design. Following this, the design was implemented on real hardware. Nexys™4 is used for implementation because it contains required peripherals such as sufficient pins for camera interface, Video Graphics Array (VGA) and serial interface. Nexys™4 has the latest artix-7 series FPGA, with device name XC7A100T [16]. For image capturing, Tamarisk®320 camera [17] with uncooled VOx microbolometer sensor was interface with the nexus evaluation board [18]. The image sensor is configured for gain and digital video interface by using (Future Technology Device International) FTDI interface. The processed video stream was displayed on a VGA monitor for visual inspection of the processed stream. To perform the experiment, the following test case has been used.

3.3 Test case

The test case used for this work is the surveillance of railway track surroundings in vulnerable areas for pedestrian detection in order to increase security and safety for the people working and living in the vicinity. Every year, hundreds of people lose their lives on the railway tracks [19]. Therefore, this type of safety-critical application requires that the system should work in difficult weather conditions. In addition to this, such systems have a real-time and low latency requirement in order to convey real-time information. To meet these challenges, the envisioned smart camera system is expected to perform processing locally and transmit useful information in real-time to train operators before approaching accident-prone areas. In local processing, background modelling and subtraction are important tasks to identify objects from the background.
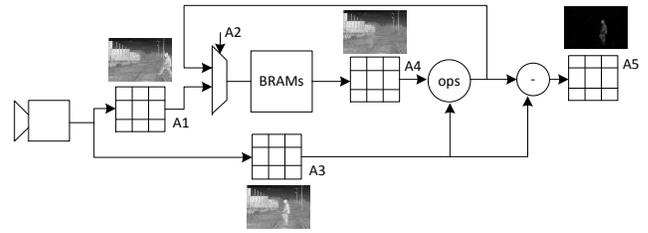

Figure 2. Architecture of hybrid background model and subtraction. (A1) first image. (A2) selection of image for BRAMs. (A3) current image (A4) previous image readout from BRAMs. (A5) subtracted image.

4. ARCHITECTURE FOR BACKGORUND SUBTRACTION

Computational architecture of the proposed hybrid technique for the background modelling and subtraction is illustrated in Figure 2. In the figure, *A1* shows the first frame, *A2* shows the multiplexer to select storage of frame in memory, *A3* is the current frame, *A4* is the modelled background and *A5* is the subtracted image. The multiplexer will allow the first frame *A1* in initial setup configuration and then it will always store new modelled backgrounds to the memory.

In RTL design, the first frame is stored in the (Block Random Access Memory) *BRAMs* without any processing. Following the second frame, the current frame and previous stored frame in *BRAMs* is used to generate a background model *Bg* by using the operations of Eq.1. The current frame is then subtracted from the generated background image in order to extract moving objects and simultaneously, the new generated background is written to the *BRAMs*. These operations are performed on the pixel level.

5. Results

In this section, the performance of the proposed technique, comparison with commonly used background subtraction techniques, resource utilization and power consumption is presented.

5.1 Performance comparison

The proposed background modelling and subtraction technique, referred to as hybrid technique, is compared with commonly used LP IIR filtering background modelling techniques for embedded systems [13][1] and other low complexity progressive background generation techniques [12]. The comparison results are illustrated with the help of three image sequences, captured with 15 frames per second. It is worth noting that Figure 3 shows an image dataset with one object but the subtraction technique is independent of the number of objects present in the image. In Figure 3, the first row (a) shows current images with objects. The second row (b) shows the background modelled using LP IIR technique, the third row (c) shows the background modelled using progressive technique, the forth row (d) shows the background modelled using hybrid technique.

The second row shows that the LP IIR technique models a background image which contains tails of moving objects because of the current image. The third row shows that the progressive technique models a sharp background image without objects and filters out the high frequency details. The fourth row shows that the hybrid technique models a background which contains rich

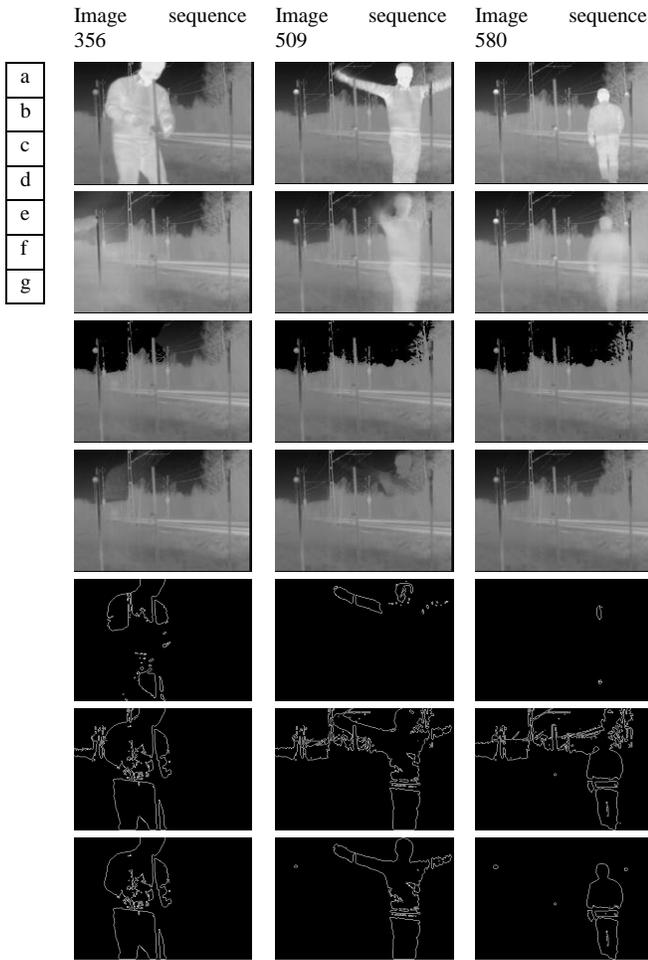|	Image sequence 356	Image sequence 509	Image sequence 580
a			
b			
c			
d			
e			
f			
g			

Figure 3. Comparison of LP IIR filtering, progressive and hybrid techniques for object detection. (a) current frames. (b) background frames with LP IIR technique. (c) background frames with progressive technique. (d) background frames with hybrid technique. (e) Images after backgorund modeling by using the LP IIR technique, subtraction, segmentation and sobe edge detection. (f) Images after backgorund modelling by using the progressive technique, subtraction, segmentation and sobel edge detection. (g) Images after backgorund modelling by using the hybrid technique, subtraction, segmentation, and sobel edge detection.

information of the current image without the moving object tail. The integration of other imaging tasks such as subtraction, segmentation and edge detection will show the effectives of the three techniques. For comparison purposes, we applied sobel edge detection and selected the same value for the threshold in the segmentation operation for the three techniques.

The fifth row (e) shows images after subtraction, segmentation and edge detection when the LP IIR technique is used for background modelling. The sixth row (f) shows images after subtraction, segmentation and edge detection when the progressive technique is used. The seventh row (g) shows images after subtraction, segmentation and edge detection when the hybrid technique is used for background modelling. It is evident that the LP IIR technique is effective for handling the high frequency noise because of the camera vibration and movement of trees. However, moving objects tail to some extent. The

progressive technique cannot handle the higher frequency noise because of the nonlinear min filtering for the background generation. The output images (g) show that the hybrid technique is able to suppress the high frequency noise while still highlighting the moving object without much noise.

Following this, the proposed background modelling and subtraction technique was then investigated with respect to different sets of images with changing weather and temperature variation. Figure 4 shows image sets from the summer season (average temperature 9.1°C to 19.5°C [20]) in the city of Sundsvall in Sweden, and Figure 5 shows image sets from the winter season (average temperature -3.6°C to -12.3°C [20]) in the city of Sundsvall in Sweden. The first row in Figure 4 and Figure 5 represents the sequence of input images, the second row represents modelled background images and the third row represents subtracted images. The subtracted images in Figure 4 and Figure 5 show that the proposed background subtraction performs well in different weather conditions. However, the ε coefficient has to be adjusted according to the summer and winter temperature variation. In the summer, ε was selected as 25/255 whereas in the winter the factor was selected as 200/255 after a number of experiments.

It is worth mentioning that in this technique, the background is updated at each frame on pixel level which in turn will provide a stable background in presence of challenges such as temperature variation and camera jitter. Following this analysis, the resources and power consumption of the proposed approach are discussed.

5.2 Resource utilization and power consumption

Table 1 shows resource utilization and memory requirement for the proposed background subtraction technique. The memory utilization shows that the proposed subtraction technique with a single frame storage requirement consumes 47 percent memory resources on latest artix-7 series FPGA, XC7A100T. The resource utilization in terms of slice LUTs is less than 1 percent. This means that post background subtraction techniques such as segmentation, morphology, labelling and compression schemes can be easily accommodated using the proposed subtraction technique, if we consider the resource requirement of the mentioned tasks for current reconfigurable devices [21]. The power consumption of the proposed background subtraction technique was calculated by using Xilinx Xpower analyzer tool [16] and is shown in Table 2.

The power consumption in Table 2 shows that existing subtraction techniques can be used for embedded systems with reduced power requirements because of little complexity and no requirements for external memory.

Table 1. Resources and memory utilization

Resources	Available	Used	Percentage
Block rams	135	6 4	47%
LUT-FF pairs	348	161	46%
Bonded IOBs	210	44	20%
Slice registers	126800	191	<1 %
Slice LUTs	63400	313	<1%

Table 2. Dynamic and quiescent power consumption.

Total power (mW)	Dynamic (mW)	Quiescent (mW)
91	8	83

6. ACKNOWLEDGEMENT

7. CONCLUSION

In this paper, a low complexity background subtraction technique for thermal imagery is proposed by using low pass IIR filtering and non-linear min filtering with single frame storage requirements. The proposed technique is compared with commonly used existing techniques and the performance is evaluated for a number of test images, captured in different weather conditions. The investigation showed that the proposed background modelling and subtraction technique with only one frame storage requirement perform better than existing techniques. This effectiveness becomes possible because of the updating processes of the modelled background at each frame on pixel level.

The proposed background modelling and subtraction technique is modelled in Register Transfer Logic (RTL) and the functionality is verified on hardware. The register transfer logic implementation shows that the proposed technique requires less than 1 percent logic resources and 64 block rams for the latest seven series FPGA, XC7A100T. The power consumption was estimated to be 91 mW by using the Xilinx Xpower analyzer.

This concludes that the proposed technique can offer robust detection of people for safety critical applications while still having low complexity on FPGA, as well as low resources and memory requirements. This in turn will result in easy integration with other pixel based pre-processing tasks such as segmentation, morphology and labelling on resource constrained embedded platforms.
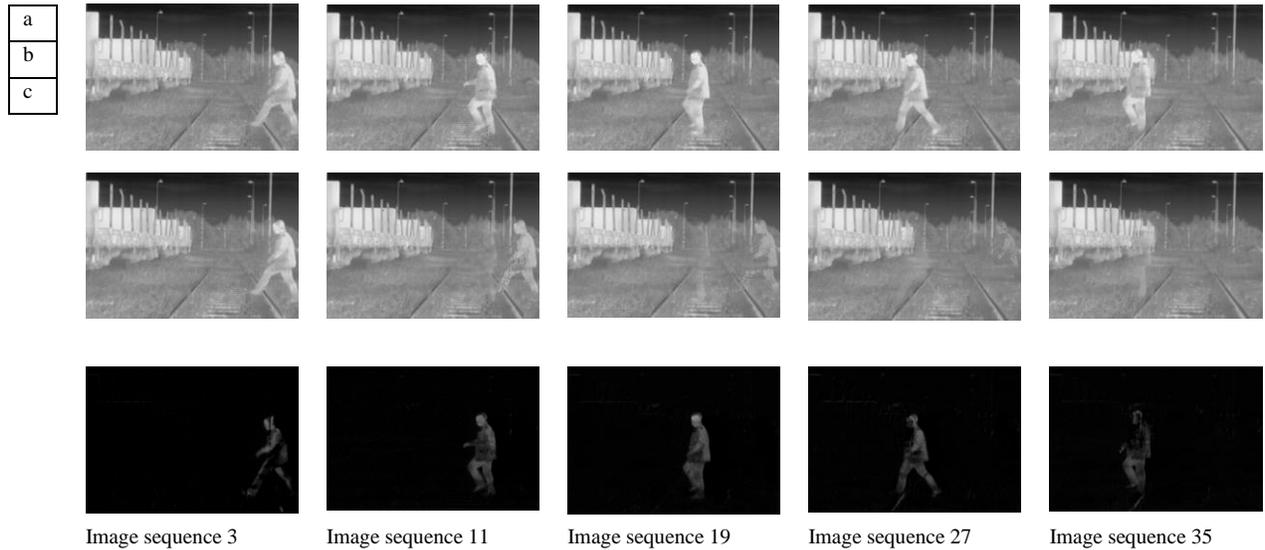

| Image sequence 3 | Image sequence 11 | Image sequence 19 | Image sequence 27 | Image sequence 35 |

Figure 4. Summer (July in Sweden). (a) current images, (b) modelled background, (c) subtracted images.

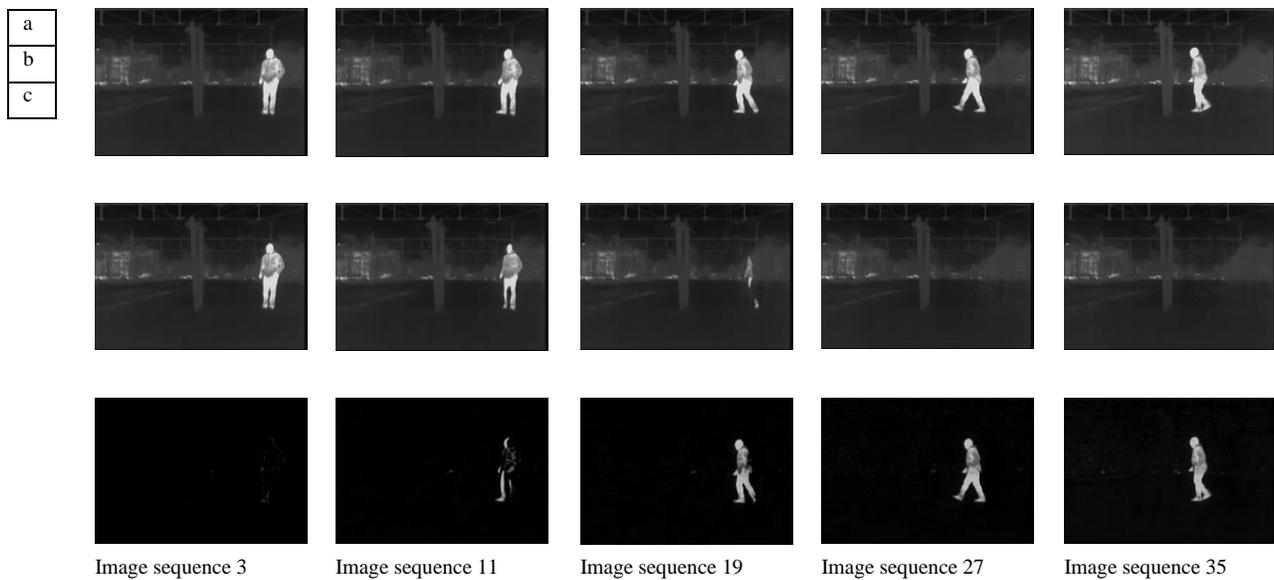| Image sequence 3 | Image sequence 11 | Image sequence 19 | Image sequence 27 | Image sequence 35 |

Figure 5. Winter (January in Sweden). (a) current images, (b) modelled background images, (c) subtracted images.

8. REFERENCES

[1] Heikkila, J. and Silven, O. 1999. A real-time system for monitoring of cyclists and pedestrians. *Second IEEE Workshop on Visual Surveillance*, (Jul 1999), 74-81.

[2] Mouats, T. and Aouf, N. 2014. Fusion of thermal and visible images for day/night moving objects detection. *Sensor Signal Processing for Defence*, (Sept. 2014), 1-5.

[3] Zin, T. T., Takahashi, H., Toriu, T. and Hama, H. 2011. Fusion of Infrared and Visible Images for Robust Person Detection", *Image Fusion, Osamu Ukimura* (Ed.), ISBN: 978-953-307-679-9, InTech, (2011).

[4] Bailey, D. G. 2011. *Design for Embedded Image Processing on FPGAs*. Asia, John Wiley & Sons, (2011), pp. 199-231.

[5] Imran, M., Shahzad, K., Ahmad, N., O'Nils, M., Lawal, N. and Oelmann, B. 2014. Energy-Efficient SRAM FPGA-Based Wireless Vision Sensor Node: SENTIOF-CAM. *IEEE Transactions on Circuits and Systems for Video Technology*, (Dec. 2014), vol.24, no.12, pp.2132,2143.

[6] Andrews Sobral, Antoine Vacavant, A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos, *Computer Vision and Image Understanding*, Volume 122, May 2014, Pages 4-21.

[7] Train accident, 2015 Availble online: http://www.svt.se/nyheter/regionalt/mittnytt/person-med-barnvagn-pakord-av-tag-1.

[8] Tessens, L., Morbee, M., Philips, W., Kleihorst, R. and Aghajan, H. 2009. Efficient approximate foreground detection for low-resource devices. *Proc. ACM/IEEE Intl. Conf. on Distributed Smart Cameras, (*2009), 1 -8.

[9] Reddy, V., Sanderson, C., Lovell, B. C., Bigdeli, A. 2009. An efficient background estimation algorithm for embedded smart cameras. *Third ACM/IEEE International Conference on Distributed Smart Cameras.* (Sept. 2009) pp.1,7.

[10] Imran, M., Ahmad, N., Khursheed, K., O'Nils, M. and Lawal, N. 2013. Low Complexity Background Subtraction for Wireless Vision Sensor Node. *16TH Euromicro Conference on Digital System Design*, (Spain, Sept. 2013).

[11] Chen, P., Ahammad, P., Boyer, C., Huang, S., Lin, L., Lobaton, E., Meingast, M., Oh, S., Wang, S., Yan, P., Yang, A.Y., Yeo, C., Chang, L., Tygar, J. D. and Sastry, S. S. 2008. CITRIC: A low-bandwidth wireless camera network platform. *Second ACM/IEEE International Conference on Distributed Smart Cameras,* (2008).

[12] Chung, Y.-C., Wang, J.-M. and Chen, S.-W. 2002. Progressive background images generation. *15th IPPR Conference on Computer Vision, Graphics and Image Processing*, (Taiwan, Aug 2002).

[13] Kandhalu, A., Rowe, A. and Rajkumar, R. 2009. DSPcam: A camera sensor system for surveillance networks. *Third ACM/IEEE International Conference on Distributed Smart Cameras*, (Sept. 2009), 1-7.

[14] Benezeth, Y., Jodoin, P.-M., Emile, B., Laurent, H. and Rosenberger, C. 2010. Comparative study of background subtraction algorithms. *Journal of Electronic Imaging*, (2010), 19(3):033003.

[15] STC thermal images, 2014. Available online: http://apachepersonal.miun.se/~muhimr/STC_thermal_image _sets.zip.

[16] Xilinx Xpower analyzer, 2014. Available Online: http://www.xilinx.com/products/design_tools/logic_design/v erification/xpower_an.htm/

[17] Tamarisk®320, Document No 1012593, 2012, Available online: http://www.drsinfrared.com/Products/Tamarisk320.aspx

[18] Nexys™4 Artix-7 FPGA Board, 2013, Available online: http://www.digilentinc.com/.

[19] Trafiksäkerheten i Sverige, 2014. Statistik och analys över järnväg, luftfart, sjöfart och väg för 2013.

[20] Sundsvall *Climate & Temperature,* Available Online: http://www.sundsvall.climatemps.com/

[21] Imran, M., Khursheed, K., Lawal, N., O'Nils, M., and Ahmad, N. 2012. Implementation of Wireless Vision Sensor Node for Characterization of Particles in Fluids", *IEEE Transactions on Circuits and Systems for Video Technology*, (Oct. 2012), vol. 22, pp. 1634 – 1643.