A Flexible Method for Detecting Individuals with Color and Thermal Cameras

Laith A. H. Al-Shimaysawee and Ali H. A. Aldabbagh
Robotics and Computer Vision Research Group (RCVRG), Dep. of Electrical Engineering, Faculty of Eng., University of Kufa, Najaf, Iraq
Email: {laitha.alshmesawi, alih.aldabbagh}@uokufa.edu.iq, {laith.anwer, ali.h.aldabbagh}@gmail.com

Nasser Asgari
College of Science and Engineering, Flinders University, Adelaide, Australia
Email: nasser.asgari@flinders.edu.au

***Abstract*—In this paper, we present an algorithm to detect people with both color and thermal images. This technique is an extension of the modified Implicit Shape Model algorithm that we had developed earlier. The algorithm can process both color and thermal images in two modes of White-Hot and Black-Hot using a single codebook generated from samples of thermal images obtained from people. The number of the samples used in the codebook is very small compared with other techniques. In the first step, the location of people is defined based on their proposed centers using the Implicit Shape Model (ISM). Then, an auto generated threshold process is used to detect people from the concentrated proposed center's density. The implementation of this technique does not need high computing power and results in improvements in the speed performance and reduction in the hardware cost. The system was evaluated using 12 image sets, six for indoor environments and six for outdoor environments. Each case included three sets of color images plus three sets of thermal images. This algorithm was implemented on a mobile robot prototype designed for rescue assist missions and the tests conducted provided promising results for detecting individuals in image sets with difficult scenarios.**

***Index Terms*—thermal image, color image, ISM, people detection, rescue**

I. Introduction

With the rapid growth of artificial intelligence and computer vision technologies in recent years, the demand for robust people and object detection techniques is on the rise because of their significance in several applications such as surveillance systems, drive assist systems (DAS), computer and robot interactive applications, rescue assist and military applications, and several others. Many studies on people detection have been carried out using various types of cameras such as normal color camera, night vision or thermal camera and so on. Each camera has its own advantage and disadvantage; however, using normal color camera is very popular in research because of their high resolution, low cost, and ease of availability compared to other types. The disadvantage is that detecting people in color cameras requires more sophisticated approaches because of the clutter in the background.

Some techniques used with thermal cameras depend on the hot spot detection [1]-[4]. These techniques perform very well in indoor environments and some easy outdoor scenarios. These methods assume that the environment is cooler than the people in that environment, but this is not the case in many scenarios and is a limit that stops them to work properly in many scenarios.

There are other algorithms that mainly depend on subtracting the background [5]-[7]. These techniques work very well with surveillance images taken from stationary cameras. However, they do not work properly for processing images from cameras placed on mobile platforms, since the background subtraction process becomes ineffective when the background constantly changes.

One of the significant studies is based on computing the gradients of infrared test images and applying model patches to a search window. In this method, Model Histogram Ratios (MHR) are computed between the patches and classified by a trained Support Vector Machines (SVM) classifier for each window [8]. This approach is similar to the original Histograms of Oriented Gradients (HOG) technique which was first proposed in [9].

Other approaches like the Implicit Shape Model (ISM) algorithm [10]-[12] train the system to detect particular objects depending on the local appearance similarity between input images and the training images. These algorithms have the advantage of not being dependent on the hot spots or background subtraction.

Most of the current studies are based on developing algorithms for a specific type of camera, but it is difficult to use the same algorithm to process images from other camera types. This leads to an increase in the hardware and software complexity of the system, since each camera type requires a different algorithm for processing the relevant image type. This will increase the system size and the power usage.

The technique discussed in this paper has been developed for detecting people in images taken from both

Manuscript received May 8, 2017; revised November 22, 2017.

doi: 10.18178/joig.5.2.39-46

thermal and color cameras. This algorithm is based on our Modified ISM Implementation for Grey Scale Thermal Images [13], [14].

The remarkable point in this research is that the same training samples are used to detect people in both color and thermal images. This approach significantly decreases the amount of training images, especially for a system which uses both thermal and normal cameras simultaneously. This technique has been implemented and tested on a mobile robot prototype for rescue assist missions.

The purpose of the robot is to search for trapped or injured people in hazardous places such as unstable buildings or destroyed structures and inform the rescue team who are following the robot from a safe distance.

The paper is organized as follow: In Section 2, we describe the hardware of the system. In Section 3, we describe our proposed method in details. In Section 4, we explain how the experiments are designed to test the system for people detection. In Section 5 we present the results and evaluation of the system, and finally, in Section 6, we include the conclusion and the future work.

II. System Hardware Description

The prototype that we have built for a rescue assist mobile robot at Flinders University is shown in (Fig. 1). The robot can explore an area autonomously, or can be controlled remotely by an operator. It scans the surrounding by rotating the camera and provides a wireless live video feed to a computer carried by the rescue team. The computer applies the vision process on the received real time video from the robot and displays a live video on its screen in which the detected subjects are highlighted. The details of the vision process have been demonstrated in the next section.

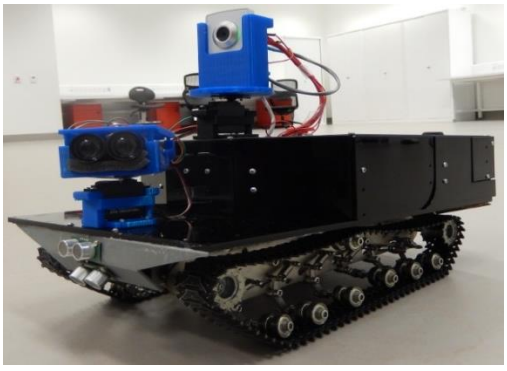


Figure 1. A mobile robot porotype built at Flinders University.

III. The System Implemention

The proposed approach has three main steps: Generating the Codebook, Learning an Implicit Shape Model (ISM), and Detecting Process.

In the first step, we used extremely small number of training people samples with black surrounding as explained in detail in Sections 4 and 5. This is the main difference from the codebook generation step in the standard ISM [10]. The step of learning an Implicit Shape Model is similar to that in the standard ISM [10]. Our main contribution is in the people detection step (see Fig. 2) and is demonstrated below:

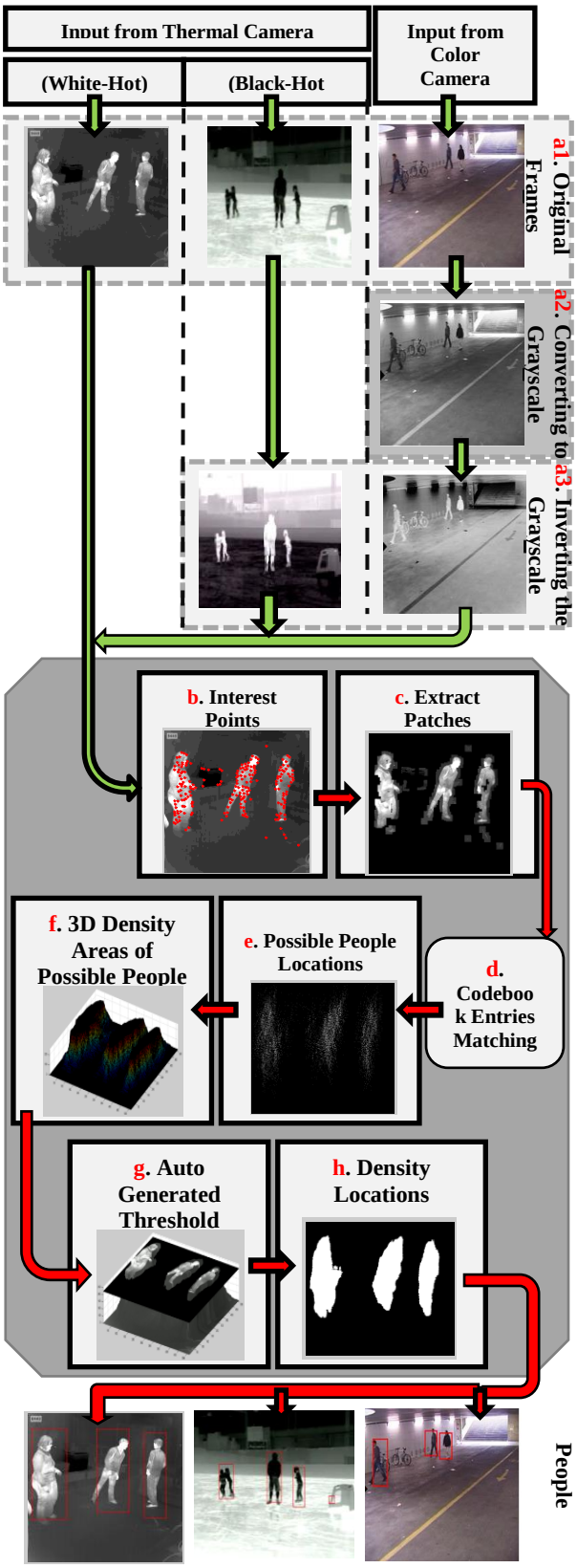


Figure 2. The recognition process for people in thermal and color image frames using a single prepared codebook.

The detecting process can analyze frames taken from color and thermal cameras. In thermal cameras using

White-Hot mode, the video frames can be analyzed directly by the detection process (see Fig. 2 a1). For thermal cameras using Black-Hot mode, we invert the grayscale of each frame by subtracting its pixel values from 255 to get 0 for white and 255 for black as shown in Fig. 2 a3.

On the other hand, in color cameras, the first step as shown in Fig. 2 a2 is to convert the color image to grey scale with 0 for black and 255 for white.

In the next step, the grey scale image is inverted by subtracting its pixel values from 255 to get 0 for white and 255 for black as shown in Fig. 2 a3. This makes a normal image look like a thermal image.

The first stage in the detection process is to use the Harris interest point detector [15] on the input image and extract patches around these points (see Fig. 2b and 2c). The Harris interest points will be mostly concentrated on the people since the thermal images and the processed color images (Fig. 2 a3) do not have the background clutter as normal images. Then, the Normalized Greyscale Correlation (NGC) function [10] is used to compare the extracted patches with the generated codebook. When the similarity exceeds a specified threshold, this codebook entry is activated (Fig. 2d). This leads to include all the possible people centers locations for each activated codebook entry in the detecting process. This is different from the standard ISM version [10] and further versions in [11], [12], [16], [17]. As shown in Fig. 2e, these centers are accumulated on positions associated with individuals.

In the next stage, high density regions are formed by the accumulation of points. This has been done by placing a block around each proposed center where these points correspond to the center of the block base (see Fig. 3). This block has a square base and the height represents the density.

When these blocks overlap, the density increase in the regions that have accumulated points more than other regions (see Fig. 2f). After testing several shapes and sizes of frames, we selected a block shape and size in Fig. 3, which presented better outcomes.

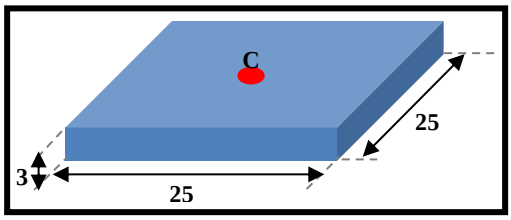


Figure 3. The designed block shape and size. The length and width are in pixels and the height refers to the density.

The final stage includes an implementation of an auto threshold mechanism to detect people from the density area as follow:

After many tests on grey input images of scale 0 to 255, we decided to define the colors range (70-90) as grey, under 70 as black, and above 90 as white.

The pixels of grey range are counted and then, the ratio of black and white pixels regarding to all image pixels is calculated using equation (1):

$$BW = 1 - (g/ (X \times Y)) \quad (1)$$

where BW is the combined black and white pixels, g is the grey pixels, and X and Y are the image dimensions in pixels.

Based on many image tests of different scenarios and situations, it was found that when the average of the grey color in the input image was low, the optimum value of threshold to get the best results has to be high, and vice versa. The threshold value is determined using the following equations:

$$dif = max - min \quad (2)$$

$$TH = min + (dif \times BW) \quad (3)$$

The maximum value (max), the minimum value (min), and their difference (dif) are calculated from the 3D density plot to find the Threshold value (TH).

The calculated threshold value (TH) represents the height of the cross sectional plane (see Fig. 2g). Then, the possible positions of people are located after applying the threshold (see Fig. 2h). To improve the detection accuracy, the areas that is less than 100 pixels, if presented, are discarded in the results.

IV. Experiments

A. Training Process

We selected from the internet several thermal images of people in different situations. The people samples are extracted from these images and the surrounding background are converted to black. The number of training samples is only eight (see Fig. 4). The purpose of the black background is to ensure that the interest points are only distributed on the people samples. This eliminates the need to divide the training samples into positive samples for individuals and negative samples for background as in [10], [11], [12], [16]. In addition, it is worth mentioning, as demonstrated in Section 5, that the training images used in the codebook is very small if compared with other ISM approaches in [10], [11], [12], [16]. Moreover, we used different scales of the training samples in order to be able to detect various scales of people and also to detect people at different distances from the camera.

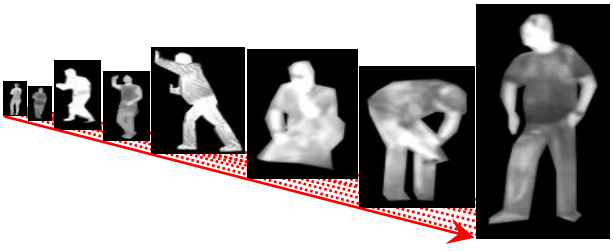


Figure 4. The training samples of people that are used in the algorithm.

B. Experimenting the Proposed Method

The proposed algorithm was tested with color and thermal images for indoor and outdoor environments. Three image sets were used for color and three sets for thermal in indoor scenarios.

In the color image sets, the first set had 150 indoor images captured using the IP camera mounted on the robot

at Tonsley campus of the Flinders University. Each of the other two sets had 250 images in total selected randomly from the sources in [18], [19]. There were 224 people in the first set, 444 people in the second set, and 239 people in the third set.

The thermal image sets were selected from the databases in [18], [20]. 40 images with124 people in them were used, from the first image set [20], to test the proposed method. The people were in low contrast with respect to the surrounding environment. Furthermore, the images in this set belongs to a security camera placed in one of the top wall corners. Therefore, the detection becomes more difficult since the input images were taken from different angle of the training images used in the codebook. From the second set [18], 40 images with a total of 150 people in them were used. The images in this set present people in high contrast with the behind background which make the detecting process easier. However, they were in different scales with some far away from the camera and others closer, which make the detecting process more challenging. From the third image set [18], 100 images with 120 people were used. Some images have individuals with their reflected image on the bright floor, and this is well to evaluate the response of the system to the reflection of people.

Similarly, for outdoor scenarios tests, three color sets and three thermal sets were used.

The first set had 150 images which were taken from the camera on the robot at Flinders University. Each of the other two sets had 250 images in total selected randomly from the databases [18], [21]. There were 441 people in the first image set, 146 people in the second set, and 191 people in the third set.

The thermal image sets were selected from databases in [18], [22], [23]. We used 70 images with 138 people from the first set [22]. The set is showing people in a forest. From the second set [23], we used 100 images with 100 individuals in them. The set shows people walking in a street with various surroundings. The third set has 100 images with a total of 147 individuals [18]. The set shows people walking in a park with very low contrast to their background.

The performance of the proposed approach was verified with different threshold (t) values using equation (4) which measures the similarity between the centers of the codebook or patches, and the patches of the input image [10]:

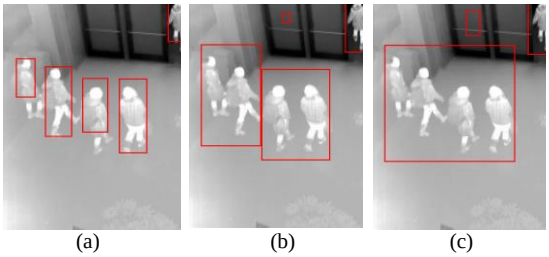
$$similarity\ (C_1, C_2) = \frac{\sum_{p \in C_1, q \in C_2} NGC(p,q)}{|C_1| \times |C_2|} > t \qquad (4)$$

The resolution of the proposed method is measured in indoor and outdoor environments by the number of the correct positives which represents the people detected properly, in addition to the number of the false positives, which represents the false detections. We used the recall-precision curve [24] to evaluate our system. To draw the recall-precision curve, we reduced the similarity threshold (t) gradually, and then we plotted the determined point on the recall-precision curve.

We have used the following rules to decide whether a person is detected or not:

- If at least half or more of the person's body is inside the detection boundary, the person is then detected, otherwise the decision is that s/He is not detected as in Fig. 5.
- A people might be surrounded by more than one small boundaries, as in see Fig. 5, especially at high values of threshold. The decision in this case is that neither a true detecting nor a false one because these small detections are combined to surround individuals when the threshold is reduced.
- If there is a boundary surrounding more than one person, the decision is that all of them are detected (see Fig. 6).

Figure 5. People were considered not detected since more than half of the body is outside the rectangle.



(a) (b) (c)

Figure 6. Detecting individuals with several values of threshold (TH): (a) at TH = 15. (b) at TH = 10. (c) at TH = 7.

V. Results and Discussion

The recall-precision curves of the color and thermal image sets for indoor detection can be seen in Fig. 7 and Fig. 8 respectively. Fig. 9 and Fig. 10 show the recall-precision curves of the color and thermal image sets for outdoor detection respectively. The results of indoor and outdoor scenarios for color and thermal image sets are discussed in the next two sections.

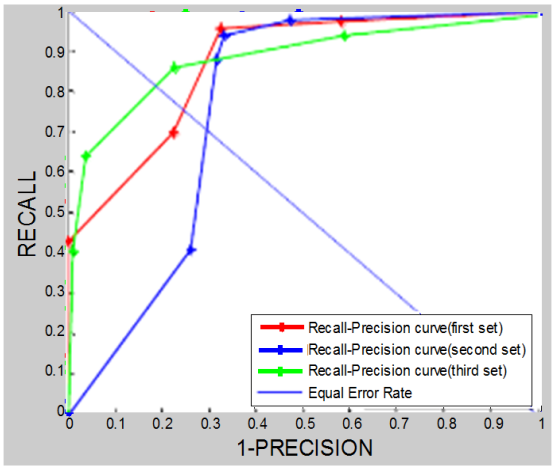


Figure 7. The recall-precision curves of the indoor color image sets.

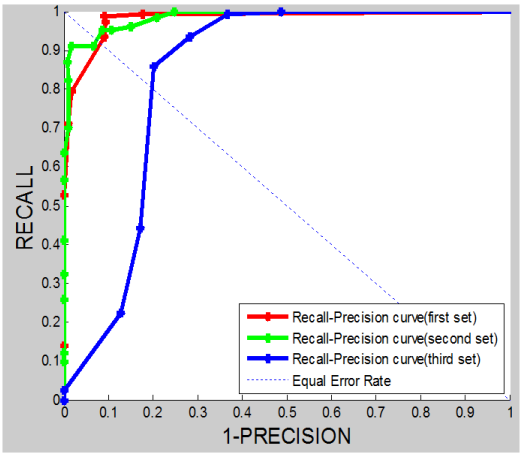


Figure 8. The recall-precision curves of the indoor thermal image sets.

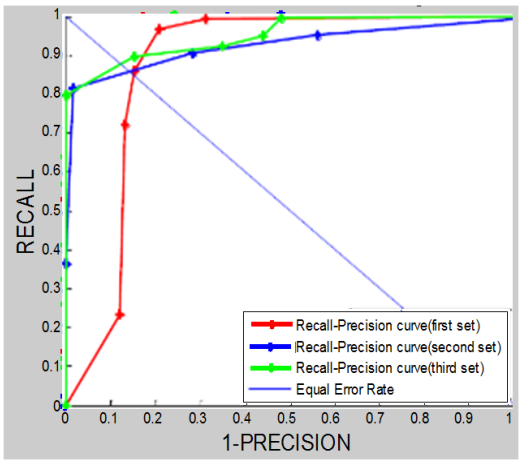


Figure 9. The recall-precision curves of the outdoor color image sets.

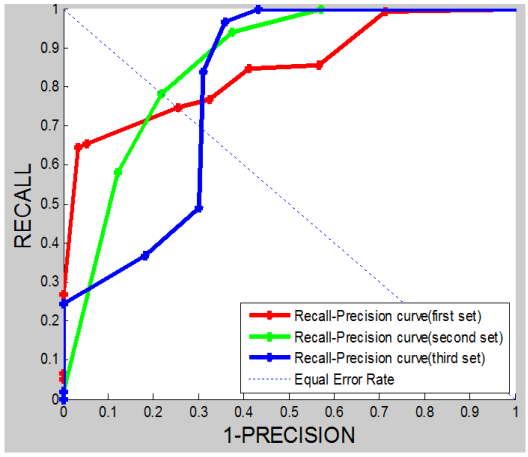


Figure 10. The recall-precision curves of the outdoor thermal image sets.

A. *Indoor Detection Performance*

For the indoor color image sets, the proposed approach achieved a good results in the first and second sets. As shown in the red and blue curves in Fig. 7, the Equal Error Rate (EER) is 75% in the first image set and 70% in the second image set. For the third set, the performance was higher with EER of 81% as shown in green in Fig. 7 and this was in spite of rather difficult conditions due to the poor lighting as mentioned in [19]. Fig. 11 shows samples of the detection results for the indoor scenarios.

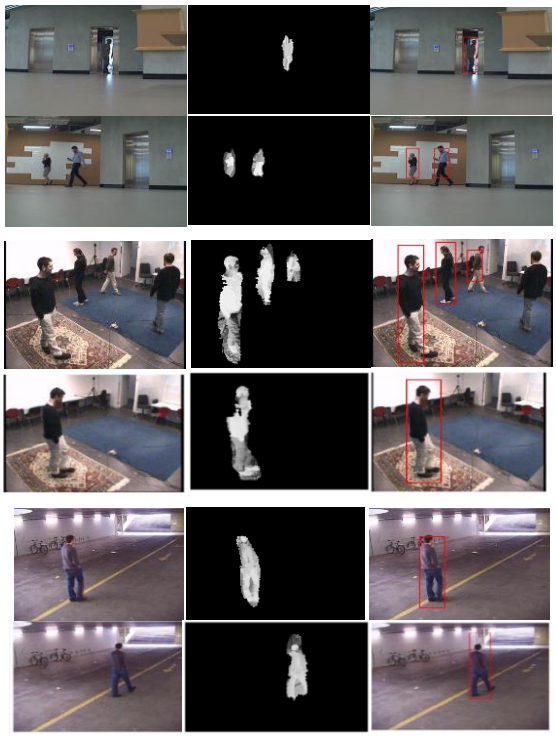


Figure 11. Samples of the detections in indoor environments: the first double-row shows images captured by IP camera placed on our robot, the middle double-row shows people of different scales with complicated background, and the third double-row shows images with poor light condition and low contrast. The first column shows the input images, the second column shows the locations of density, and the last column shows the detections of individuals.

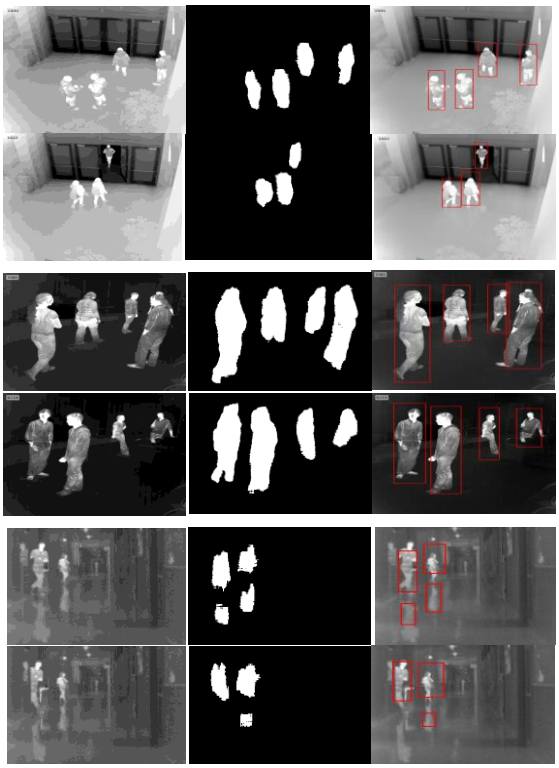


Figure 12. Samples of the detections in indoor environments: the first double-row shows images of a security camera from the top, the middle double-row shows people of different scales, and the third double-row shows people with their visible reflection on the ground. The first column shows the input images, the second column shows the locations of density, and the last column shows the detections of individuals.

For the indoor thermal image sets, the system achieved a very good results in the first and second sets as shown in the red and green curves in Fig. 8. The EER is 92% in the first image set and 91% in the second set. In the last set, there are many images show individuals with their reflective on the floor. When testing such images, at some threshold values, the reflections have been detected as an individual person. The system point of view considers this a correct detecting as it does not distinguish between a person and its reflection; on the other hand, the human point of view, considers this as a false detecting. In order to put more challenge for the system, the decision was to go with the human observation and recording false positives for the detected reflections. This explains why the EER was 80% for this the set (see the blue curve in Fig. 8) which is not as good as in the first and second sets. The EER can be higher if we discard this restriction. Fig. 12 shows samples of the detection results for the indoor scenarios.

B. Outdoor Detection Performance

In the outdoor color image sets, for the first set the system reached a remarkable results of 85% EER as can be seen in the red curve in Fig. 9. The performance of the approach was slightly better in the second set with 86% EER as can be seen in the blue curve in Fig. 9. In the last set, the proposed approach performed marginally better with an EER of 87% as shown in the green curve in Fig. 9. Fig. 13 shows some samples of the detections in outdoor images scenarios.

In the outdoor thermal image sets, while individuals are in small scales and the surroundings have many hot spots in the images of the first set, the system reached a good outcomes of 75% EER as seen in the red curve in Fig. 10. The performance of the approach was better in the second set with 78% EER as seen in the green curve in Fig. 10. In the last set, since the people and the background have similar temperature values, the EER dropped to 70% as shown in the blue curve in Fig. 10. Fig. 14 shows some samples of the detections in outdoor images scenarios.

It is apparent from the results that the system performed better in outdoor color images compared to the indoor ones. This is because in the outdoor environment, especially in the sun light, the contrast between the people and the surrounding background increases to make the people darker than environment.

By inverting the images as part of the detection process, people become brighter than the environment (similar to greyscale white-hot thermal images), and this makes their detection easier. However, for the indoor environment, people and the environment usually have similar brightness, and this makes it challenging for the system to detect people.

In contrast, usually the detection process in thermal outdoor images is more difficult than in indoors. In indoor environments, individuals generally are hotter than the background, so they can be detected easily; however, in outdoor environments, many hot objects and even hotter than the human can be available, and this increases the difficulty of the detection task.

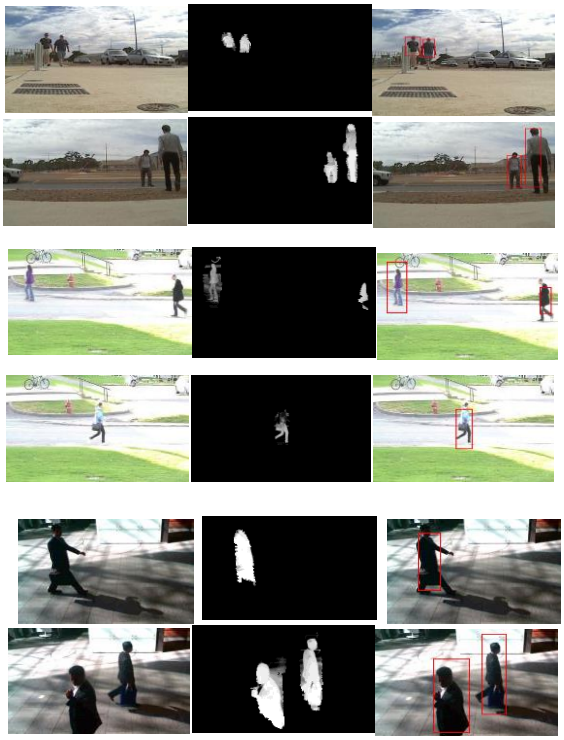


Figure 13. Samples of the detections in outdoor environments: the first double-row shows images from IP-camera placed on our robot, the middle double-row shows people moves in a park, and the third double-row shows people with their shadows. The first column shows the input images, the second column shows the locations of density, and the last column shows the detections of individuals.

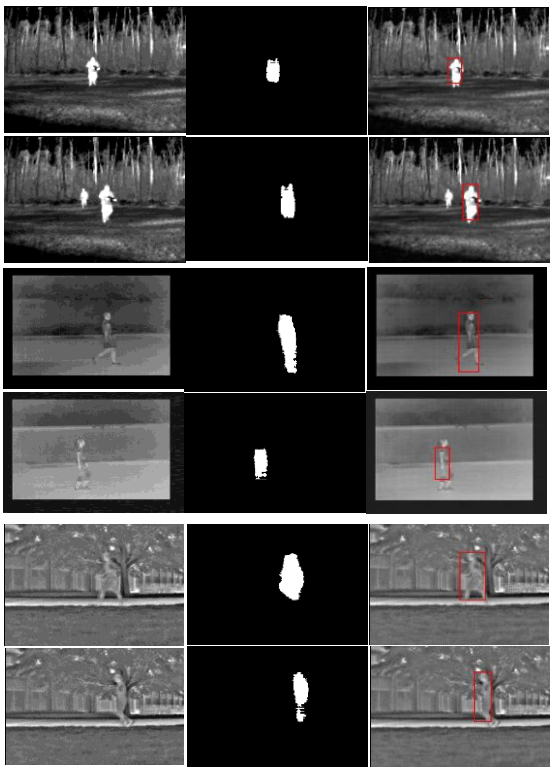


Figure 14. Samples of the detections in outdoor environments: the first double-row shows individuals in a forest, the middle double-row shows people walk in street with several backgrounds, and the third double-row shows people move in a park in a similar contrast with their surroundings. The first column shows the input images, the second column shows the locations of density, and the last column shows the detections of individuals.

C. Overall Evaluation of the System Results

Table I shows the summary of the test results for each indoor/outdoor and color/thermal image set with the number of people in each set, the number of correct and false detections, and the EER%. It is noticeable that for the indoor scenarios, the system performance was much higher for thermal images compared to color images. The reason is that a photo taken with a color camera indoors captures a lot of details, hence, people cannot be detected easily. This is not the case in indoor thermal images, where the complexity of the surroundings is eliminated since it has similar temperature, then the dominant is the hot objects only. On the other hand, as shown in Table I, the situation is opposite in outdoor scenarios and the system performance is higher for color images compared to thermal images. This is because a photo taken with a thermal camera outdoors has a lot of details because of the high temperature in the outdoor scenarios, so people cannot be recognized easily. In contrast, for a color image taken outdoors, the background generally has less details and makes detecting people easier.

The remarkable point is that the training images were not selected from the same sources of the input images, but were taken from unrelated websites. The second point is the very small number of training images in the codebook which was eight only.

It should be pointed out that in [16] the total number of training images was 550 where these systems need positive and negative training images for people and background respectively.

TABLE I. SUMMARY OF THE SYSTEM TESTS RESULTS.

Type of image set		Image set	No. of people	EER (%)	Correct detections	False detections
Indoor scenarios	Color image	1st	224	75	168	56
		2nd	444	70	311	133
		3rd	239	81	194	45
		Total	907	74	673	234
	Thermal image	1st	124	91	113	11
		2nd	150	92	138	12
		3rd	120	80	96	24
		Total	394	88	347	47
Outdoor scenarios	Color image	1st	441	85	375	66
		2nd	146	86	126	20
		3rd	191	88	168	23
		Total	778	86	669	109
	Thermal image	1st	138	75	104	34
		2nd	100	78	78	22
		3rd	147	70	103	44
		Total	385	74	285	100

The proposed method does not need negative training images and the eight training images is sufficient. When using more training images, the process becomes slower. Naturally, using less training images leads to consuming less power and better performance. Consequently, it can be used in small mobile robots with limited processing power. It can be also more suitable for implementing on FPGA to increase the speed.

VI. CONCLUSION AND FUTURE WORK

This algorithm is capable of processing both color and thermal images in White-Hot or Black-Hot modes using a single codebook generated from thermal image samples of people. The number of the codebook's training images is very small if compared with other techniques. The reason is the use of black surroundings in the training images and leads to concentrate the Harris interest points on the individuals. As a result, more details can be extracted from the training images. Therefore, the codebook can be created effectively from small number of training images.

There are many color patterns that complicate a color image and make people recognition difficult. Converting a color image to greyscale will remove most of these patterns. Moreover, inverting a greyscale image in most cases results in people showing in high contrast with their surroundings and this simplifies detecting them.

In the process of detecting, the accumulation of the proposed people centers achieve high density areas that refer to the possible locations of individuals in addition to the creation of auto threshold mechanism helps to detect people from the density areas. This process makes the detections more reliable in difficult environments and the implementation easier.

This was the first time that a system used a codebook generated with thermal image samples to work on both thermal and color images for people detection, resulting in reduced system complexity and increased speed and performance. This system is implemented on a mobile robot prototype for rescue assist missions capable of detecting human in disaster zones. In the future, this system can be used for people tracking in many applications such as a mule robot following solders or rescuers while carrying their equipment. The system can be also modified slightly to be used for detecting parts of a human body in robot interactive applications or similar.

ACKNOWLEDGMENT

This research was funded by Flinders University, University of Kufa and the Higher Committee for Education Development in Iraq.

REFERENCES

[1] G. Cielniak, and T. Duckett, "People recognition by mobile robots," *Journal of Intelligent and Fuzzy Systems*, vol. 15, no. 1, 21-27, 2004.

[2] J. C. Castillo, J. Serrano-Cuerda, A. Fernández-Caballero, and M. T. López, "Segmenting humans from mobile thermal infrared imagery," in *Bioinspired Applications in Artificial and Natural Computation*, Springer Berlin Heidelberg, 2009, pp. 334-343.

[3] J. Serrano-Cuerda, M. T. Lopez, and A. Fernández-Caballero, "Robust human detection and tracking in intelligent environments by information fusion of color and infrared video," in *2011 7th International Conference on Intelligent Environments, IEEE*, July 2011, pp. 354- 357.

[4] M. Teutsch, T. Mueller, M. Huber, and J. Beyerer, "Low resolution person detection with a moving thermal infrared camera by hot spot classification," in *Proc. IEEE Conference on*

Computer Vision and Pattern Recognition Workshops, June 2014, pp. 209-216.
[5] J., Davis, and V. Sharma, "Robust background-subtraction for person detection in thermal imagery," in *IEEE Int. Wkshp. on Object Tracking and Classification Beyond the Visible Spectrum*, 2004.
[6] J. W. Davis and V. Sharma, "Robust detection of people in thermal imagery," in *Proc. 17th International Conference on Pattern Recognition*, 2004, vol. 4, pp. 713-716.
[7] E. J. Carmona, J. Martínez-Cantos, and J. Mira, "A new video segmentation method of moving objects based on blob-level knowledge," *Pattern Recognition Letters*, vol. 29, no. 3, pp. 272-285, 2008.
[8] R. Miezianko and D. Pokrajac, "People detection in low resolution infrared videos," in *Proc. Computer Vision and Pattern Recognition Workshops*, 2008.
[9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2005.
[10] B. Leibe, A. Leonardis, and B. Schiele, "Combined object categorization and segmentation with an implicit shape model," in *Proc. Workshop on Statistical Learning in Computer Vision*, 2004, p. 7.
[11] B. Leibe, A. Leonardis, and B. Schiele, "Robust object detection with interleaved categorization and segmentation," *International Journal of Computer Vision*, vol. 77, no. 1-3, pp. 259-289, 2008.
[12] A. Lehmann, B. Leibe, and L. V. Gool, "Fast prism: Branch and bound hough transform for object class detection," *International Journal of Computer Vision*, vol. 94, no. 2, pp. 175-197, 2011.
[13] L. A. H. Al-Shimaysawee, A. H. A. Aldabbagh, and N. Asgari, "A novel algorithm for people detection in grey scale the rmal images," in *Proc. International Conference on Control, Automation and Robotics, IEEE*, Singapore, May 2015, pp. 238-243.
[14] A. H. A. Aldabbagh, L. A. H. Al-Shimaysawee, and N. Asgari, "A discriminative approach for people detection using color camera for mobile robot platforms," *International Journal of Mechanical Engineering and Robotics Research*, vol. 5, no. 2, pp. 103-108. April, 2016.
[15] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vision Conference*, August 1988, pp. 147-151.
[16] A. Konigs and D. Schulz, "Evaluation of thermal imaging for people detection in outdoor scenarios," in *Proc. IEEE International Symposium on Safety, Security, and Rescue Robotics*, College Station, TX., November 2012, pp. 1-6.
[17] K. Jungling and M. Arens, "Feature based person detection beyond the visible spectrum," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Miami, FL, June 2009, pp. 30-37.
[18] 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 Conference on Computer Vision and Pattern Recognition Workshops*, Columbus, OH, June 2014, pp. 393-400.
[19] F. Fleuret, J. Berclaz, R. Lengagne, and P. Fua, "Multi-Camera people tracking with a probabilistic occupancy map," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.
[20] Z. Wu, N. Fuller, D. Theriault, and M. Betke, "A thermal infrared video benchmark for visual analysis," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Columbus, OH, June 2014, pp. 201-208.
[21] J. Berclaz, F. Fleuret, E. Türetken, and P. Fua, "Multiple object tracking using K-shortest paths optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.
[22] R. Miezianko, "Terravic research infrared database," IEEE OTCBVS WS Series Bench, 2005.
[23] S. Zheng. (2005). CASIA Gait database collected by Institute of Automation, Chinese Academy of Sciences, CASIA Gait Database. [Online]. Available: sinobiometrics.com
[24] S. Agarwal and D. Roth, "Learning a sparse representation for object detection," in *Proceedings of the 7th European Conference on Computer Vision Copenhagen*, Denmark: Springer-Verlag Berlin Heidelberg, vol. 2353, pp. 113-127, May 2002.

Laith A. H. Al-Shimaysawee was born in Najaf-Iraq on 22/03/1988. Al-Shimaysawee awarded masters degree in electronic engineering from Flinders University, Adelaide, Australia, 2015. He has been lecturing at the University of Kufa since 2011. He worked as a Maintenance Engineer in labs of the Electrical Engineering Department in the Faculty of Engineering, University of Kufa. He is a member of the Golden Key International Honour Society since Feb. 2015. He has a Graduate Membership (GradIEAust) in the Institution of Engineer Australia in the Information, Telecommunication and Electronics Engineering College since Nov. 2015. He has publications in the field of robotics, computer vision and digital communication systems.

Ali H. A. Aldabbagh was born in Najaf-Iraq on 28/07/1988. Aldabbagh awarded masters degree in electronic engineering from Flinders University, Adelaide, Australia, 2015. He has been lecturing at the University of Kufa since 2011.
He worked at Najaf governorate as an Adviser Engineer for nine months on electrical foundations at buildings which were under construction. Then, he worked as a Maintenance Engineer in the labs of the Electrical Engineering Department in the Faculty of Engineering, University of Kufa. He is a member of the Golden Key International Honour Society since Feb. 2015. He has a Graduate Membership (GradIEAust) in the Institution of Engineer Australia in the Information, Telecommunication and Electronics Engineering College since Nov. 2015. He has publications in the field of robotics and computer vision.

Nasser Asgari was born on 20/05/1959. Asgari awarded PhD degree in the areas of Computer Engineering from the University of Adelaide, Australia in 2003. He has been lecturing at Flinders University since 1998. He has publications in computer vision, robot localization, and Interprocessor Communications using Multiport Memories.