Background Extraction Technique in Thermal Imagery; Case Studies during Hajj

Adnan Ghazi Abuarafah

Vice Dean for Academic Affairs, Faculty of Computer of and IS, Umm Al- Qura University, Makkah, SA
agabuarafah@uqu.edu.sa

Abstract: Analyzing crowd density in huge crowds like Hajj is very important and background extraction is an effective image processing method to calculate the accurate density of crowd. An advanced and improved approach is introduced based on the background extraction from thermal images to calculate the accurate density of crowd during Hajj 1433H. Firstly, we will get the thermal video frames. They consist of people with background shadows and objects. We then calculate the average of all video frames to get the background image and then the background is removed. The crowd density is then calculated and the robustness of our approach is demonstrated through multiple case studies which gives an accurate results of crowd density calculation and proves to be an enhanced approach.

[Adnan Ghazi Abuarafah. **Background Extraction Technique in Thermal Imagery; Case Studies during Hajj.** *Researcher* 2013;5(5):64-74]. (ISSN: 1553-9865). http://www.sciencepub.net. 10

Keywords: Person detection; thermal imagery; FLIR; background noise; hajj crowd.

1. Introduction

The first step of many computer vision applications is to locate moving objects in a video sequence. Among the various motion-detection techniques, background subtraction methods are commonly implemented as mentioned by Benezeth Y et. al [1]. Background subtraction is an important first step for many vision problems.

A great variety of background subtraction techniques have been proposed in the past years but we are proposing a background removal technique in thermal imagery to get the accurate crowd density during Hajj 1433H. The thermal properties of people is quite different from that of background objects in an image. It is much easier using thermal video cameras to distinguish people from objects in the background. However, the accurate estimation of crowd density is still not possible with just the thermal imagery because of the interference between the temperatures of some of the background objects and the people. We propose multiple case studies to prove practically the accuracy of the background removal technique and to get an accurate density of crowd using CMINS (Crowd Management Intelligent Notification System) module. [2]

In many computer vision applications, the first relevant step of information extraction is the detection of moving objects in video streams. For detecting moving objects, background subtraction is a widely used approach. Over the recent years, many different methods have been proposed. Picardi and Massimo helped researchers to select the best approach for the background subtraction. [3] Background subtraction is also one of the key techniques for automatic video analysis, especially in the domain of video surveillance. [4]

The technique of background removal using thermal cameras gives accurate density of the crowd and it is experimentally proved in this paper. It is also one of the important image processing steps for many computer vision problems such as recognition, activity analysis, tracking and classification. [5]

This paper is divided into five sections. The first section is this introduction. The second section discusses the related work. The third section illustrates the first case study including some parts from the used code. The fourth section shows the second case study and the last section concludes the work and discusses the future of the research.

2. Related Work

Bharti T. says that background subtraction approach is used to detect the moving object from background. Different methods have been proposed to detect object motion by using different background subtraction techniques over recent years. Each technique has its own benefits and limitations. [6] We propose an approach which has no limitation but provides a complete accurate density of crowd.

Gordon et. al illustrated a background estimation method based on a multidimensional (range and color) clustering at each image pixel. They discussed in details the important implementation issues such as low confidence measurements and treatment of shadows. Figure-1 shows their work on background removal technique. [7]

The accurate crowd density cannot be obtained through the above mentioned approach because there may be interference of the colors of the human objects with background leading to inaccurate density. Also, the background removal technique used in the above mentioned paper is useful only during the daytime but not during nighttime due to problems in vision. The proposed work overcome the above mentioned issues of background removal technique. This is done by making use of background extraction technique on thermal (FLIR) video camera frames.

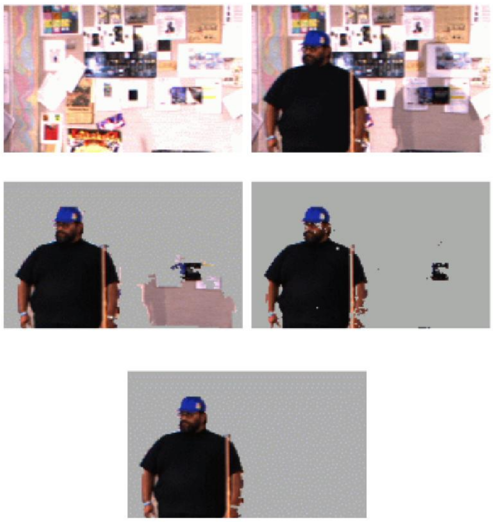

Figure 1: Top: Background image, person in foreground casting a strong shadow. Middle left: Basic color segmentation, shadow remains. Middle right: Effect on color segmentation when using the higher threshold for the entire image: skin tones close to background color are eroded. Bottom: large portions of shadow removed with adaptive (range-based) threshold.

Silveira et. al proposed a concurrent implementation of a background subtraction algorithm suitable for monochromatic video sequences. They used a set of frames to determine an estimate of the scene background, as well as background noise. Each new frame of the video sequence was compared to the background model, and foreground objects are obtained taking into account the background noise estimate. Shadows and highlights are also detected and removed. Figure-2 shows the final output of their proposed method, with shadow/highlight removal and morphological post-processing. [8]

However, a normal camera is not applicable to be used for both day and night scenarios and more problems may rise due to the interference of people with objects in the background. In order to overcome this issue, we are making use of thermal imagery in this paper.

Zhang et. al presented a dynamic background method with three contributions. First they present a local dependency descriptor, called local dependency histogram (LDH), to effectively model the spatial dependencies between a pixel and its neighboring pixels. Second, based on the proposed LDH, an effective approach to dynamic background subtraction was proposed, in which each pixel is modeled as a group of weighted LDHs. Finally, unlike traditional approaches using a fixed threshold to judge whether a pixel matches to its model, an adaptive thresholding technique was also proposed. Figure-3 shows their experimental results. [9]

However, there was no focus on the issues concerning the effects of the background shadow effects and other issues related to the accuracy in crowd density calculations. We will focus on these issues by making use of thermal (FLIR) video camera to get the accurate density of crowd without any interferences.

(a) (b) (c) (d)

Figure 2: (a) Background model. (b) A frame with few foreground pixels. (c) Foreground pixels obtained through condition (3). (d) Final output of the proposed method, with shadow/highlight removal and morphological post-processing.

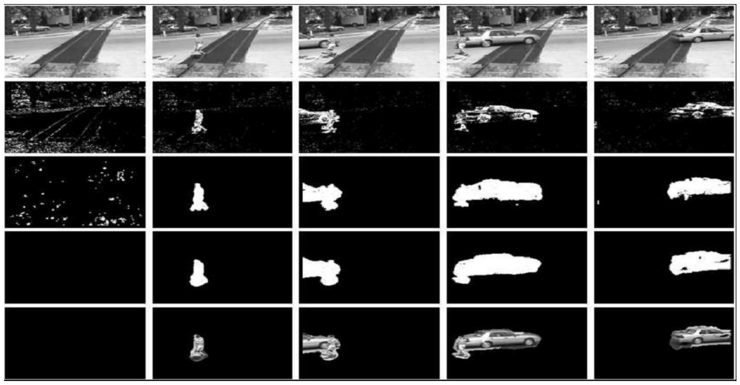

Figure 3: Experimental results on fountain sequence. Morphological operators were not used in the results. The top row are the original images: the 10^{th}, 597^{th}, 877^{th}, 1113^{th} and 1816^{th} frames. The second and third rows are the results obtained by Mixture of Gaussians method (MOG) and Local Binary Patterns based method (LBP), respectively. The fourth row are the results obtained by the proposed method, and the fifth row are the masked original images.

Chan et. al proposed an adaptive model for backgrounds containing significant stochastic motion (e.g. water). The model was based on a generalization of the Stauffer– Grimson background model, where each mixture component is modeled as a dynamic texture. They report on experimental results, which show that the proposed background model both quantitatively and qualitatively outperforms state-of-the-art methods in scenes containing significant background motions. Figure-4 shows some of their work on background subtraction. [10]

Moreover, they did not paid attention to the tribulations in making use of a normal camera which can be overcome by thermal imagery as proved in this paper in case studies.

Figure 4: Foreground detection results on the beach scene from

Davis and Sharma presented a contour-based background-subtraction technique in thermal imagery to extract foreground objects. Statistical background-subtraction was used to identify local regions-of-interest. Their approach was designed to handle the problems typically associated with thermal imagery. [11]

The issues of thermal imagery mentioned in the above paper [11] didn't take account of the new problem of interference of the temperature and colors of human bodies with that of objects in the background. This issue is been carefully examined in this paper later in the case studies.

3. CASE STUDY -1

This is the first model using CMINS module to prove the accuracy of the background removal method on the thermal image by comparing it with the temperature extraction method.

A. Image from a thermal video

The below Figure-5 is an image captured in Makkah Mukarramah, KSA during Hajj 1433H from FLIR thermal video camera.

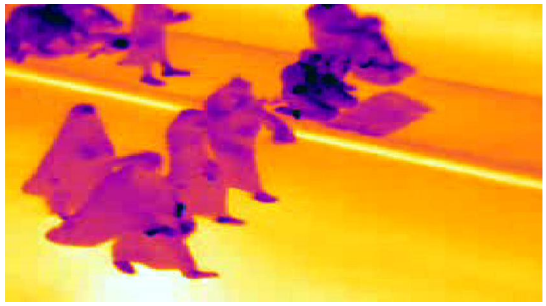

Figure 5: Image from a thermal camera during Hajj1433H

B. Thermal Video frame with ruler

The below Figure-6 is an image with the temperature ruler on the right of the image captured in Makkah Mukarramah, KSA during Hajj 1433H from FLIR thermal video camera.

Figure 6: Thermal image with the temperature ruler on the right during Hajj1433H

C. Density calculation from image (Figure-6) with temperature ruler using CMINS program [6]

The figure below (Figure-7) is a screenshot of a CMINS module to calculate the density of crowd. The land area in figure-7 is 6 m * 5 m (30 m^2) and there are 13 persons. And the average capacity rate is 2 in each 1 m^2. So, the 13 persons will take about 22 % from the total land area. So experimentally, the calculated value should be 22% in order to get the accurate density of crowd.

Now we use a thermal image with temperature rules and colors (as a sample image) to calculate the density and we got 18 % density which is not accurate. This inaccuracy was due to the interference in the color of the people and the background.

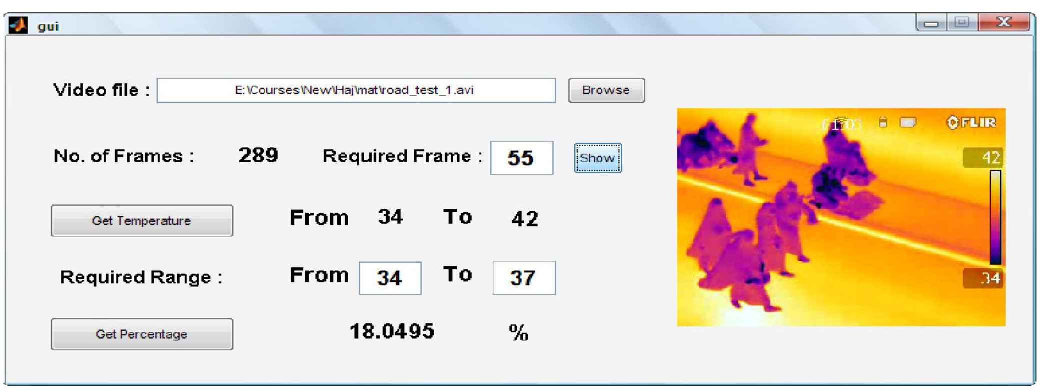

Figure 7: Calculation of crowd density using thermal image with ruler using CMINS module [6]

The code of the program" Matlab tool" is as follows:

```
% get the minimum required range of temperature which is the minimum body
% temperature for normal people which is 34 in this example
r1 = str2num(get(handles.txtMinRange,'String'));
```

```
% get the maximum required range of temperature which is the maximum body
% temperature for normal people which is 37 in this example
r2 = str2num(get(handles.txtMaxRange,'String'));
if isempty(r1) || isempty(r2)
  h = msgbox('You must enter a valid Range to get its percentage','Invalid Temperature Range','error');
  return;
end
% make sure that the entered body temperature range is valid
if (min(r1,r2) < t2) || (max(r1,r2) > t1)
  h = msgbox(strcat('Please Write Temperature Range between ',num2str(t2),' and ', num2str(t1),'.'), 'Invalid
Temperature Range','error');
   return;
end
% make sure that the entered body temperature range is valid
if r1 > r2
    h = msgbox('Maximum Range must be greater than or equal Minimum Range','Invalid Temperature
Range','error');
   return;
end
% calculate the density percentage of people
p = getPercentage(a,r1,r2,t2,t1);
set(handles.txtPer,'String',p);
function p = getPercentage(im,minR,maxR,minT,maxT)
% make sure that the entered body temperature range is valid
if maxR < minR
    p=0;
    return;
end
im=rgb2gray(im);
diffT = maxT-minT;
diffR = maxR-minR;
% get range of colors in the temperature ruler matching the required
% temperature range
range = round((diffR/diffT)*101);
st = round(((maxT - maxR)/diffT)*101 + 70);
high = im(st:(st+range),307:314);
ind = find(((im(:,:,1)<=max(max(high(:,:,1)))) & (im(:,:,1)>=min(min(high(:,:,1))))));
% No. of pixels matching the required range of temperature
sum = length(ind);
[r,c,m] = size(im);
% calculate percentage of matching pixels relative to all frame pixels
p = (sum/(r*c)) * 100;
end
```

D. Background image from Thermal video

After getting an inaccurate result (18%) of the density of crowd from the thermal image with temperature ruler and color, we introduced a new method by removing the background from the thermal images. In this section (D), we are getting a background of the thermal image used in the previous sections (B & C). Firstly, we extracted the background image as an average of all the frames of the captured video. The figure below (Figure-8) is a background of the thermal image.

Figure 8: Background of the thermal image

The below mentioned steps were taken to get the background of the thermal image

If Ik(y, x) is the intensity level at coordinates Y = y and X = x of the kth frame in the sequence and bgk(y; x) is the average background model value at (y, x), then

$$bg^k(y,x) = \frac{1}{N}\sum_{j=k-(N/2)}^{k+(N/2)} I^j(y,x)$$

The code of the program:

```
% get all framed of the entered video file
[vidFrames,frames,rate]=getFrames(fullPath,1,n);
cla(handles.axes2);
cla(handles.axes4);
mov=[];
movsave=[];
gr=[];
clear memory
% convert frames from color to gray
for i=1:n
  gr(:,:,i) = rgb2gray(vidFrames(:,:,:,i));
end
% get average of gray frames
m=mean(gr,3);
```

E. Thermal image after background removal

After extracting the background from the thermal image, we then subtract this background from the thermal image and we got a binary image for the moving objects only in white as shown in the figure below(Figure-9).

Figure 9: Binary image after subtracting background (Figure-8) from thermal image (Figure-6)

The segmented output is:

$$seg1^k(y,x) = \begin{cases} 1 & , \text{if } |bg^k(y,x) - I^k(y,x)| > T1 \\ 0 & , \text{otherwise} \end{cases}$$

Where $T1$ is a predetermined threshold, usually 10.

After subtracting this background model from the current frame and thresholding the difference image, this figure results. The white pixels in the image are the regions that have been segmented as moving pixels in the current frame.

The code of the program:

```
% subtract every frame from the background average frame
gr(:,:,i) = abs(rgb2gray(vidFrames(:,:,:,i))-m);
% convert the subtracted frame to binary
bin = im2bw(gr(:,:,i),graythresh(gr(:,:,i)));
```

F. Density calculation from image without background

After getting the inaccurate density percentage (18%) in section -C, we are now calculating the density again but with deploying a thermal image without background into CMINS module [6].

We got a density of about 22.2 % which is more accurate (as mentioned in the starting of section-C) due to the 13 moving persons in 30 m^2 land area leading to the density of 22%.

This approach is thus proved to be more accurate as it uses the thermal images without background and does not depend on temperature ruler and colors which leads to no interference between the colors of body parts and the background.

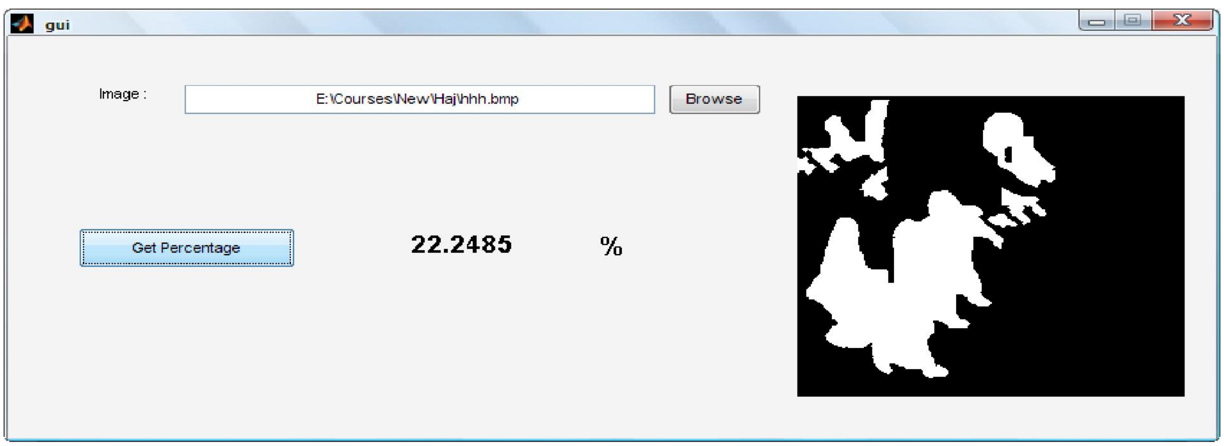

Figure 10: Density calculation using a thermal image without background.

Seg1 is the segmented image which is binary and summation of it is the summation of white pixels in it, so the percentage is calculated as follows:

Per = 1/N * ∑ seg1

The code of the program:

```
[r,c] = size(im);
% The no. of white pixels are the pixels of moving people
w = length(find(im==255));
% calculate percentage of white pixels relative to all frame pixels
p = w / (r*c) * 100;
```

4. CASE STUDY - 2

This is the second model using CMINS module to prove the accuracy of the background removal method on the thermal image by comparing it with the temperature extraction method.

A. Image from a thermal video

The below figure-11 is an image captured in Makkah Mukarramah, KSA during Hajj 1433H from FLIR thermal video camera.

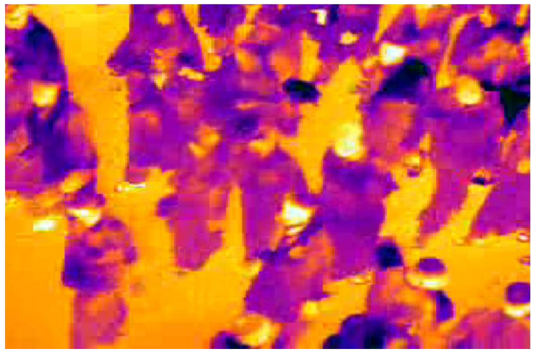

Figure 11: Image from a thermal video camera during Hajj1433H

B. Thermal Video frame with ruler

The below Figure-12 is an image with the temperature ruler on the right of the image captured in Makkah Mukarramah, KSA during Hajj 1433H from FLIR thermal video camera.

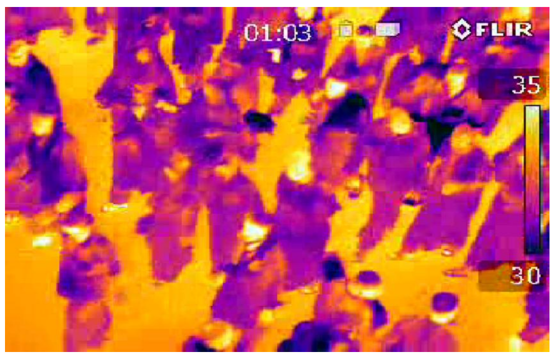

Figure 12: Thermal image with the temperature ruler on the right during Hajj1433H

C. Density calculation from image (Figure-13) with temperature ruler using CMINS program [6]

The figure below (Figure-13) is a screenshot of a CMINS module to calculate the density of crowd. The land area in figure-7 is 6 m * 5 m (30 m^2) and there are 39 persons. And the average capacity rate is 2 in each 1 m^2. So, the 39 persons will take about 65 % from the total land area. So experimentally, the calculated value should be 65% in order to get the accurate density of crowd.

Now we use a thermal image with temperature rules and colors (as a sample image) to calculate the density and we got 68 % density which is not accurate. This inaccuracy was due to the interference in the color of the people colors and the background.

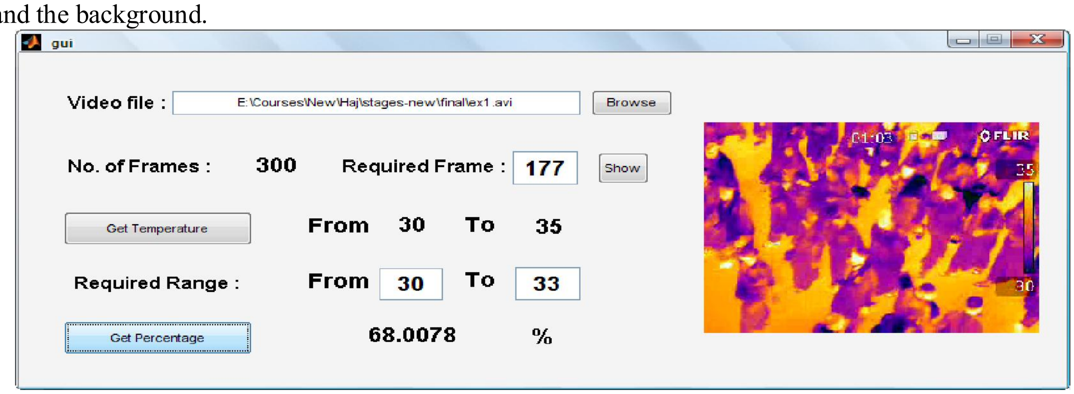

Figure 13: Calculation of crowd density using thermal image with ruler and color using CMINS module

D. Background image from Thermal video

After getting an inaccurate result (18%) of the density of crowd from the thermal image with temperature ruler and color, we introduced a new method by removing the background from the thermal images. In this section (D), we are getting a background of the thermal image used in the previous sections (B & C)

Firstly, we extracted the background image as an average of all the frames of the captured video. The figure below (Figure-14) is a background of the thermal image.

Figure 14: Background of the thermal image

The below mentioned steps were taken to get the background of the thermal image

If Ik(y, x) is the intensity level at coordinates Y = y and X = x of the kth frame in the sequence and bgk(y; x) is the average background model value at (y, x), then

$$bg^k(y,x) = \frac{1}{N}\sum_{j=k-(N/2)}^{k+(N/2)} I^j(y,x)$$

E. Thermal image after background removal

After extracting the background from the thermal image, we then subtract this background from the thermal image and we got a binary image for the moving objects only in white as shown in the figure below(Figure-15).

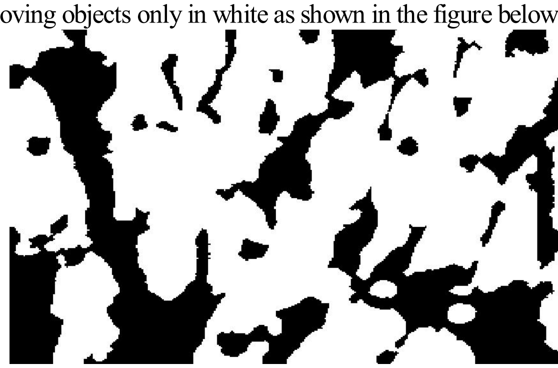

Figure 15: Binary image after subtracting background (Figure-14) from thermal image (Figure-12)

The segmented output is:

$$seg1^k(y,x) = \begin{cases} 1 & , \text{if } |bg^k(y,x) - I^k(y,x)| > T1 \\ 0 & , \text{otherwise} \end{cases}$$

Where *T*1 is a predetermined threshold, usually 10.

After subtracting this background model from the current frame and thresholding the difference image, this figure results. The white pixels in the image are the regions that have been segmented as moving pixels in the current frame.

F. Density calculation from image without background

After getting the inaccurate density percentage (68%) in section-C, we are now calculating the density again but with deploying a thermal image without background into CMINS module [6].

We got a density of about 65.01 % which is more accurate (as mentioned in the starting of section-C) due to the 39 moving persons in 30 m^2 land area leading to the density of 65%.

This approach is thus proved to be more accurate as it uses the thermal images without background and does not depend on temperature ruler and colors which leads to no interference between the colors of body parts and the background.

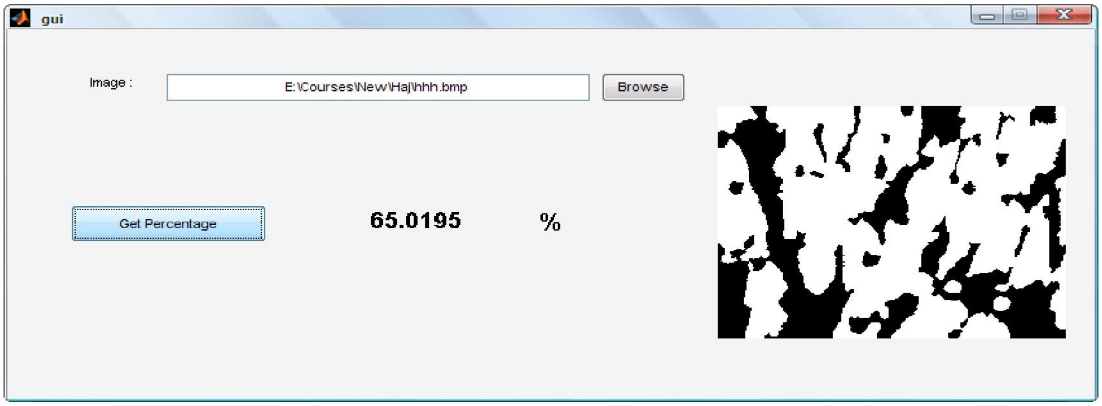

Figure 16: Density calculation using a thermal image without background.

Seg1 is the segmented image which is binary and summation of it is the summation of white pixels in it, so the percentage is calculated as follows:

$$Per = 1/N * \sum seg1$$

5. Conclusion

We have proposed a novel technique in getting an accurate density of crowd by deploying background subtraction technique onto thermal imagery. We had shown two different case studies done during Hajj 1433H and proved the accuracy of the crowd density calculation and the robustness of our approach.

The results of the case studies had proved our approach to be an enhanced one in this research paper and the finals results are discussed below.

For Case Study-1

(Note: 22% is the accurate crowd density of the sample image used in the experiment)

- Crowd density percentage is **not accurate** using the thermal image. It was 18.05% by using the CMINS program[6]. This is because of the interference of colors of the people and the background of the image.
- Crowd density percentage **is accurate** using the thermal image with the background removal. It was 22.25% by using the CMINS program. This is because of the extraction of background from the thermal image.

For Case Study-2

(Note: 65% is the accurate crowd density of the sample image used in the experiment).

- Crowd density percentage is **not accurate** using the thermal image. It was 68.01% by using the CMINS program[6]. This is because of the interference of colors of the people and the background of the image.
- Crowd density percentage **is accurate** using the thermal image with the background removal. It was 65.02% by using the CMINS program. This is because of the extraction of background from the thermal image.

In the future, we may apply this triumphant proved approach in different sectors/applications in order to get the exact crowd density and give an enhanced surveillance or monitoring of overcrowded areas.

References

[1] Benezeth, Y.; Jodoin, P.-M.; Emile, B.; Laurent, H.; Rosenberger, C., 2008 "Review and evaluation of commonly-implemented background subtraction algorithms," Pattern Recognition, ICPR 2008. 19th International Conference on, vol., no., pp.1 (4): 8-11 Dec. 2008.

[2] Khozium M Osama, 2012. "A hybrid intelligent information system for the administration of massive mass of Hajjis". Life Science Journal 9(2): 377-383, 2012.

[3] Piccardi, M., "Background subtraction techniques: a review, 2004. " Systems, Man and Cybernetics, 2004 IEEE International Conference on , vol.4, no., pp.:3099-3104 (2004) vol.4, 10-13 Oct. 2004)4.

[4] Brutzer, S.; Hoferlin, B.; Heidemann, G., "Evaluation of background subtraction techniques for video surveillance," Computer

Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, vol., no., pp.1937-1944, 20-25 June 2011.

[5] Komagal, E.; Vinodhini, A.; Srinivasan, A.; Ekava, B., 2012 "Real time Background Subtraction techniques for detection of moving objects in video surveillance system," Computing, Communication and Applications (ICCCA), 2012 International Conference on , vol., no., pp.1 (5): 22-24 Feb. 2012.

[6] Bharti, Tejinder Thind, 2013, "Background Subtraction Technique review". International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278 -3075, Volume - 2, Issue –(3):…., February 2013.

[7] Gordon G., T. Darrell, M. Harville, J. Woodfill, 1999. "Background estimation and removal based on range and color". In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (Fort Collins, CO), June 1999.

[8] Silveria Cesar, Cavalheiro Gerson, Jung Claudio and Musse Julio. 2010. "An improved background subtraction algorithm and concurrent implementations" Parallel Processing Letters, Vol. 20, No.(1): (2010) 71-89.

[9] Zhang Shengping, Yao Hongxun and Liu Shaohui. 2009. "Dynamic background subtraction based on local dependancy and histogram". International Journal of Pattern Recognition and Artificial Intelligence Vol. 23, No.(7): (2009) 1397–1419.

[10] Chan Antony, Mahadevan Vijay and Vasconcelos Nuno. 2011. "Generalized Stauffer–Grimson background subtraction for dynamic scenes". Machine Vision and Applications (2011) 22:751–766.

[11] Davis James and Sharma Vinay. 2007. "Background-Subtraction in Thermal Imagery Using Contour Saliency". International Journal of Computer Vision 71(2), 161–181, 2007.

3/4/2013