Crowd Reckoning towards Preventing the Repeat of ‘2015 Hajj Pilgrims Stampede’

Blinded for Review

***Abstract*— The issue of crowd reckoning has as of late gotten noteworthy consideration in the current era. Given an image of a crowded environment, the objective is to evaluate the density of the crowd, by counting the quantity of individuals that it contains. There are some approaches to solve crowd reckoning problem. They intend to reckon crowd for a medium sized place. But crowd detection at Hajj Pilgrimage site has never been tried before as far we have studied. Hajj Pilgrimage sites are different than most of the places. The aim of this work is to prevent any future stampede at Hajj Pilgrimage sites. This paper proposes a method where head detection technique is used for the crowd reckoning. It uses the segmentation of head regions from still images. The crowd reckoning will determine the possibility of a stampede and can alarm the people to take necessary steps. Assessment of the proposed model demonstrates worthy performance in realistic operating conditions.**

Keywords—Crowd Reckoning; Pilgrims; Stampede; Erosion; Image Processing.

I. Introduction

The objective of crowd reckoning, is to counting the number of people that it contains by a given video or an image of a crowded environment. Nowadays, this is certainly a great concern about human safety where overcrowded is a common scenario [1], [2]. On 24 September 2015, a crowd catastrophe caused the losses of life measurably 2,236 hajjis who were stampeded during the annual Hajj pilgrimage in Mina of Mecca, making it the deadliest Hajj tragedy ever. Estimation of the number of death may vary; the Associated Press reported 2,411 deaths. The definite reason for the congestion that led to an enormous crush on Mina is yet to be determined. To overcome this issue, we need to develop a method which can warn us of overcrowded scenario. The key idea is to count people and warn if the selected place is overcrowded.

Because of the extensive application of people reckoning, most computer-based systems undergo from the following problems First, mutual barrier among people causes momentous distinctions in their presences and the loss of wrest features. It often results in a miscalculation of the quantity of individuals. Second, the difficulties initiated by images which have low resolution or blurry condition, especially for people who are far distant from the camera, generally reduce the strength of a counting system. Thirdly, women wearing hijab or people having umbrella may cause wrong counting. Finally, large variations in the appearance of individuals and lighting conditions, and in addition chaotic backgrounds, make crowd reckoning more challenging [1]. Crowd counting problem has been tried to evaluate many times. There are some approaches to solve crowd counting problem. Crowd detection at Hajj Pilgrimage site has never been tried before.

This paper proposes a new method which aims to develop an efficient method that will reckon crowd at Hajj Pilgrimage sites. The goal is to prevent the repeat of 2015 Hajj Pilgrims Stampede. The proposed method uses head detection for the crowd counting. It uses the segmentation of head regions from still images. The result of crowd reckoning will determine the possibility of a stampede and can alarm the people to take necessary steps.

The rest of the paper is organized as follows. Section II includes some related works of associated fields, the proposed method is illustrated in section III. Experimental analysis is shown in Section IV. Finally, Section V concludes the paper.

II. Related Work

In this section, we survey the research related to the advancement of the proposed methodology.

Ankan at el. [5] introduced a crowed counting method that can detect high density crowds. But it is not so efficient for Hajj pilgrimage sites because mutual occlusion among peoples causes loss of information.

Fu at el. [6] proposed a crowd counting approach which can only estimate crowd in low density. It captures images in low area segments but it can’t work in high density area.

Huiyuan Fu [7] proposed another promising crowd counting method which uses depth camera to find the possible head regions. But it cannot detect crowd in large region. The depth camera is costly which incurs the cost overhead of the total system.

III. The Proposed Methodology

To accomplish the research, a sample image is taken and it is then converted to gray scale image which is shown in Fig. 1. The several steps of the proposed methodology are then carried out and these are outlined shortly as follows.

A. Thresholding

Thresholding is the easiest technique for image segmentation and building binary images from grayscale images. We apply thresholding on our image to segment head region from other body parts. We have chosen thresholding value 127. If a pixel intensity is less than 127, then it is considered to be a head region, otherwise it is considered background or other body parts. The resulting binary thresholding image is displayed in Fig. 2.

Fig. 1: Gray-scale Image converted Image sample.

Fig. 2: Convert Binary Image from Gray-scale Image using Thresholding.

B. Segmentation

For dense crowds, people are standing closely to each other. That's why it is tough to detect individual's head. So, we applied erosion to segment head regions of individuals. Erosion is a morphological image processing technique. The key idea in binary morphology is to probe an image with a template image, conclude on how this template fits or misses the shapes in the image. Template itself is a binary image (i.e., a subset of the space or grid). It is a filtering process where image details that are smaller than the template are filtered.

If two heads are connected as shown in Fig. 3(a), then by applying erosion, we are removing those connecting lines and separate the two heads as shown in Fig. 3(b). We are using 3x3 kernel for this purpose. By applying this procedure two or three times, we get our resulted image where we can apply our head detection technique. We take a 3x3 kernel to perform erosion which will move over the image. If all the pixels inside the kernel equals to high density (binary value = 1), then the center pixel is selected, otherwise the center pixel is ignored. The ignored pixel will have value zero at the end of the process.

Fig. 3. Segmentation of two heads using erosion.

At first, we set our kernel over the image. Now before we examine all contained pixels, we see that the center pixel is already zero, so we ignore this pixel, as shown in Fig.4(a). Now we move the kernel to one column right. Again, we check where this kernel contained pixels, as shown in Fig.4(b). Here not all of the pixels are of high density, so we ignore the center pixel too. We set this pixel intensity to zero. Fig. 4(c) and Fig. 4(d) follow the same scenario. After some iterations, we have a situation where all the kernel contained pixels are equal to one, as shown in Fig. 4(e). So, we select the center pixel and move the kernel to the right as shown in Fig. 4(f). Fig. 4(g) shows the image with the selected three pixels after all the iterations are completed. Fig. 4(h) is the final result where ignored pixels are set to zero.

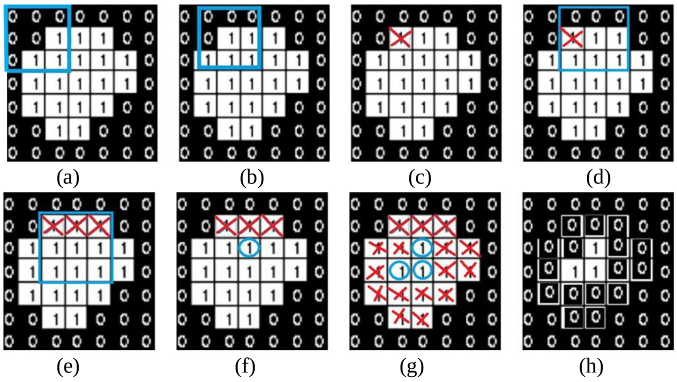


Fig. 4. Working procedure of Erosion.

C. Head Detection

After exceeding the erosion method, we apply head detection technique. Fig. 5 describes the process of head region counting. At first, we invert the image to make our work easier, as shown in Fig. 5(a). Then we select the first pixel as shown in Fig. 5(b) whose intensity is high, because this can be a possible head region. After that we traverse the eight connected neighbor in recursive manner. If the neighboring pixels' intensity is also high, then we mark these pixels as visited as shown in Fig. 5(c). We terminating the recursive process for this pixel when no more neighbor pixels can be selected. So after this iteration with a pixel, we get the result as shown in Fig. 5(d). Then we calculate the area of the marked region. If this area value is not sufficient to be a predefined head region, then we discard this considering as a noise. Otherwise we increment our counter of head region.

Now the process is again on loop with the next pixel whose intensity is high and the pixel is not visited as shown in Fig.5(e). Finally, after traversing all the pixels in the image, the result is shown in Fig. 5(f). We have found two possible head regions with this sample image.

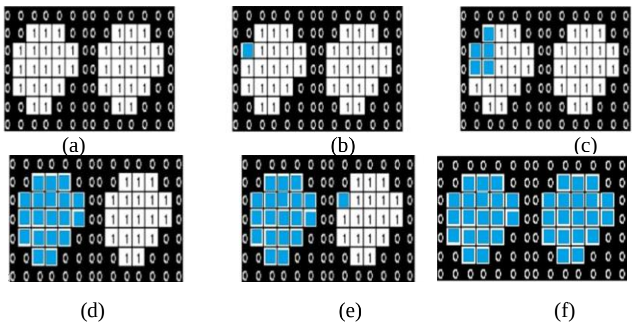


Fig. 5. Detection of head region using DFS.

Then we calculate the area of the marked region. If this area value is not sufficient to be a predefined head region, then we discard this considering as a noise. Otherwise we increment our counter of head region. This process is continued for all high intensity pixels in the image.

IV. EXPERIMENTAL ANALYSIS

The experimental setup and the analysis are outlined shortly as follows.

A. Experimental Tools

In order to the accomplish the proposed research, the following tools are used.

- **Language**: C++.
- **Library**: Open CV.
- **Processor**: Intel corei3 with 4GB RAM.
- **Image size**: 500 x 300.

B. Results

From Fig. 6 it can be observed that at our first iteration, the image has a lot of noise elements indicated by big orange circles. Then in the next iteration, we search head regions, and count them. In Fig. 7 and Fig. 8, small circles indicate individuals head regions and large circles indicate noise elements.

Fig. 6. 1st Iteration Image Contains too much noise.

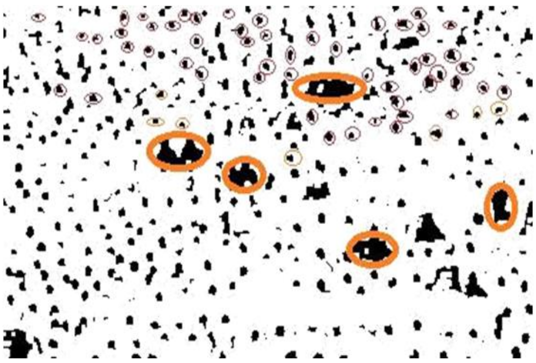


Fig. 7: Noise element marked on 2nd iteration.

Fig. 10 demonstrates the result of crowd reckoning for 12 images. For each image, there are four sets of data which show the result of original crowd counting (in this case, it was done manually), Crowd counting after 1st iteration, crowd counting after 2nd iteration and crowd counting after 3rd iteration.

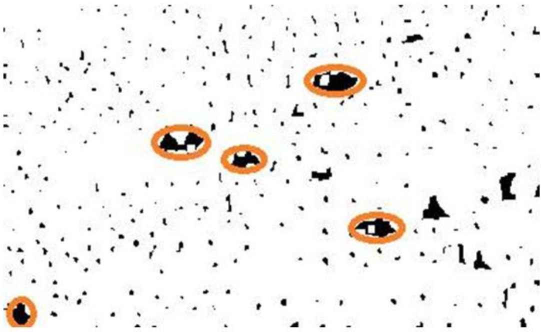
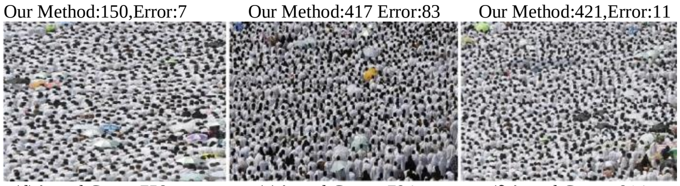


Fig. 8. Noise element marked on 3rd iteration

(a)Actual Count: 143 Our Method:150,Error:7

(b)Actual Count: 500 Our Method:417 Error:83

(c)Actual Count: 410 Our Method:421,Error:11



(d)Actual Count:758 Our Method:719,Error:39 (g)

(e)Actual Count: 794 Our Method:741,Error:53 (h)

(f)Actual Count: 914 Our Method:761,Error:153 (i)

Fig. 9. Sample Images for result anaysis

In Fig. 10, it is clear that head detection result on 1st iteration is long distant from original count due to heavy noisy element. But 2nd and 3rd iterations are showing promising results. So we take the average of them as our estimated original count using the following equation.

$$F = \frac{1}{n}\sum_{i=1}^{n} P_i$$

Here, F means the calculated average result and P_i is the result from each iteration, i = 1, 2, 3, ..., n (n is number of iteration).

We use absolute error (AE) and normalized absolute error (NAE) for evaluating the performance. Here η_{AE} and η_{NAE} denote the mean of AE and NAE respectively, η is the estimated count and i is the actual ground based count, and N is the number of iterations.

$$\eta_{AE} = \frac{1}{N}\sum_{i=1}^{N} |\eta_i - \hat{\eta}_i|$$

$$\eta_{NAE} = \frac{1}{N}\sum_{i=1}^{N} \frac{|\eta_i - \hat{\eta}_i|}{\eta_i}$$

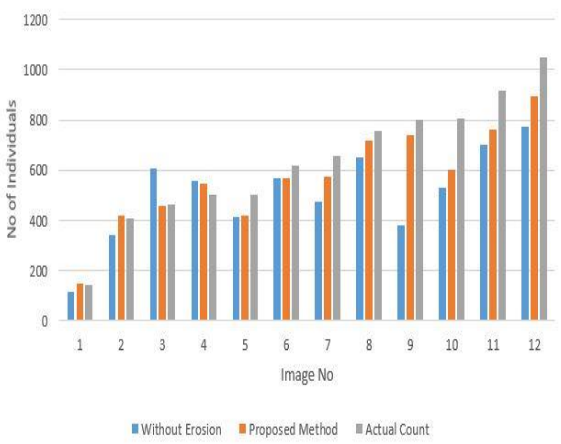


Fig. 10. Crowd counting Results for 12 images

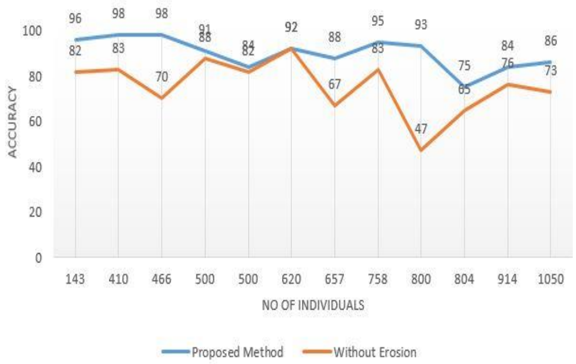


Fig. 11. Crowd counting accuracy results for 13 images

Fig. 11 shows that the accuracy of our proposed method is increased with the increase of number of people in images.

Table I shows the absolute errors and normalized absolute errors for three different sample images (Fig. 9). It can be observed from this table that we get high normalized absolute error for Fig. 9 (c) which has very high density crowds. Due to high density, it becomes a tedious job to separate two or more heads by the erosion technique. Hence, the denser the image is, the more possibility of noise is there.

TABLE I. THE ERRORS OBSERVED FOR THREE DIFFERENT SAMPLE IMAGES

	Actual Count	1st Iteration	2nd Iteration	3rd Iteration	AE	NAE
Image 1	143	32	150	151	7.5	.05
Image 2	410	104	421	422	11.5	.03
Image 3	466	305	448	464	10	.02
Image 4	500	162	626	467	79.5	.02
Image 5	500	53	397	438	85.5	.17
Image 6	620	113	592	543	52.5	.08
Image 7	557	74	558	592	18	.03
Image 8	758	78	680	758	39	.05
Image 9	800	136	749	734	58.5	.07
Image10	804	71	565	635	204	.25
Image11	914	72	720	802	153	.16
Image12	1050	102	840	946	157	.14

V. CONCLUSIONS

We considered a method for estimating the number of hajj pilgrims in extremely dense crowds from still images. The counting problem at this scale has barely been tackled before. We evaluate our method with respect to the head count estimation and the experimental results are convincing. Since the model is very simple, it can be applied for real time head counting in pilgrimage sites to prevent any possibility of a stampede. For low density crowds, our method is not efficient. The limitation increases if there are huge noisy elements or other objects. We plan to enhance its capabilities to detect head regions in case of larger noisy areas.

REFERENCES

[1] Nick C. Tang, Yen-Yu Lin, Ming-Fang Weng, and Hong-Yuan Mark Liao," Cross-Camera Knowledge Transfer for Multiview People Counting" *IEEE Transactions On Image Processing*, Vol. 24, No. 1, January 2015.

[2] B. Liu and N. Vasconcelos, "Bayesian Model Adaptation for Crowd Counts," *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, 2015, pp. 4175-4183.

[3] H.D. Ma, C.B. Zeng, and C.X. Ling, "A reliable people counting system via multiple cameras," *ACM Transactions on Intelligent Systems and Technology*, vol. 3,pp. 1-22, 2012

[4] B. Antic, D. Letic, D. Culibrk, and V. Crnojevic, "K- means based segmentation for real time zenithal people counting," *IEEE International Conference on Image Processing*, pp 2565-2568, 2009.

[5] Ankan Bansal and K S Venkatesh, "People counting in High Density crowds from still images," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp 1093-1100, 2009.

[6] Huiyuan Fu, Huadong Ma, Hongtian Xiao, "Crowd Counting via head detection and motion ow estimation," *IEEE International Conference on Image Processing*, 2014.

[7] Huiyuan Fu, Huadong Ma, Hongtian Xiao, "Real-time accurate crowd counting based on rgb-d information," IEEE International Conference on Image Processing, pp 2585-2568, 2012.

.