Automatic Estimation of Crowd Density

Jugal Kishor Gupta
Research Scholar, IFTM University,
Moradabad, India

S. K. Gupta
Bundelkhand Institute of Engineering and
Technology, Jhansi, India

ABSTRACT
This paper considers the problem of automatic estimation of crowd densities, an important part of the problem of automatic crowd monitoring and control. Anew technique based on texture description of the images of the area under surveillance is proposed. Two methods based on different approaches of texture analysis, one statistical and another spectral, are applied on real images captured in an area of Liverpool Street Railway Station, London, UK. The results obtained show that both methods present similar general rates of correct estimation, and that the potential use of texture description for the problem of automatic estimation of crowd densities is encouraging

Keywords
Crowd, Image, Surveillance

1. INTRODUCTION
The management and control of crowds is a crucial problem for human safety, since when an accident happens where there is congestion of people many lives can be lost [1].Two important aspects of the problem of correct management and control of crowds are the design of environments where crowd congestion is expected to arise and the real-time monitoring of crowds within existing, typically urban, structures. The development of models of crowd behaviour provides a basis for informing architects and town planners to design safer buildings. Some reviews crowd psychology in terms of its relationship to engineering and crowd safety and stresses the need to validate computer simulations of crowd movement and escape behaviour against psychological as well as engineering criteria [2]. For the problem of real-time crowd monitoring there is an established practice of using extensive closed circuit television systems. As routine crowd monitoring is tedious, human observers, responsible for the simultaneous monitoring of many different areas through an array of television monitors, are likely to lose concentration. The advantages and necessity of automatic surveillance for routine crowd monitoring are, therefore, clear. This paper describes a new technique based on texture description for the problem of automatic estimation of crowd density.

2. PREVIOUS TECHNIQUE FOR AUTOMATIC ESTIMATION OF CROWD DENSITY
Davies et. al. [3] have proposed a technique to estimate crowd densities based on two measures extracted from the input image of the area under surveillance. The first measure is the number of for ground picture elements computed by subtracting the input image from a reference image containing no people. The second measure is the number of edge picture elements of the image computed by an edge detection followed by a thinning operation.

3. OUR METHOD
Firstly, after inputting video, we extract the foreground image through Gaussian mixture model, and then process foreground image; the processing includes filtering noise through median filtering and morphological operations. Secondly, estimate the number of people after dividing the region of interest. The specific statistical methods are: give a preliminary judgment for the crowd density through the number of foreground pixel. For different density, we use the method of pixel statistics or GLCM texture analysis to extract crowd feature. And then use linear regression to estimate the number of the crowd. Finally, add the number of each region and then estimate crowd density. It is worth mentioning that we estimate the crowd size for the high density and extremely high-density crowd.

3.1 Pixel Feature
The property of the pixel statistic is the earliest feature to be used for crowd density estimation, and it is a very effective feature. The basic idea of this algorithm is: the denser crowd, the greater proportion of the foreground image. Researchers considered that there is a linear relationship between the number of foreground pixel and number of people in the scene. Pixel features usually are: the foreground image area, perimeter, and edge pixels, and so on.

Pixel statistical algorithm is relatively intuitive, easy to understand, low computational complexity, the relationship between the number of people and pixel feature is relatively simple after pre-processed, easy to train, and the generalization ability of classifier or function relationship is very well after training. However, the pixel statistical algorithm has some problems: foreground image segmentation algorithm is not ideal, and needs to correct the weight of extracted pixel due to the impact of perspective distortion, has bad result in high-density crowd.

In this paper, this pixel statistical method is used to give the initial judgment of crowd density, and estimate the crowd density of extremely low density, low density, and medium density.

3.2 Texture feature
The pixel is very important feature among crowd density estimation, but the accuracy is very low for more serious occlusion area. To solve this problem, Marana proposed texture analysis algorithm. Different density crowd has different texture pattern for texture analysis. Images of low density crowds show coarse texture, while images of high density crowds show fine texture. The calculation of GLCM texture features is a common and effective method. In this paper, this texture method is used to estimate the crowd density of extremely high-density and high density.

3.3 Definition of Classification
Polus [9] proposed crowd density from low to high is divided into five levels. In this paper, we reference to this definition, crowd density levels are defined as shown inTable 1.

Table 1: Crowd Density of each category

Classi-fied level	Extremely low Density	Low Density	Moderate Density	High Density	Extremely High Density
Crowd boun-dary (people)	0-10	11-30	31-60	61-100	>100

3.4 Proposed Method and Result Analysis

Figure 1 shows the details of proposed method in this paper. This method:

a) Capture video, and use Gaussian mixture model to extract video foregrounds. For the foreground image, we use method of binary process, noise elimination by median filtering, and morphological operation.

b) Set the region of interest. Since the presence of abnormal projection, especially in the process of large-scale monitoring, the effect of abnormal projection is particularly evident brought by the perspective effect. To solve this problem, this paper divides into four different sub-regions for each scene image. The sub-region division effects showed in Figure 1.

3.4.1 How the learning is done?

We evade the hard task of learning to detect and localize individual object instances. Instead, we cast the problem as that of estimating a continuous density function whose integral over any image region gives the count of objects within that region.

We start with a set of dot-annotated training images and a set of features, so that each pixel in those images is assigned a real-valued feature vector describing the local appearance (in our experiments, these features were defined using either using SIFT-based visual dictionaries or a set of randomized trees). Our system then learn a linear mapping, that transforms the feature vector at each pixel to a density value, obtaining density function value in that pixel.

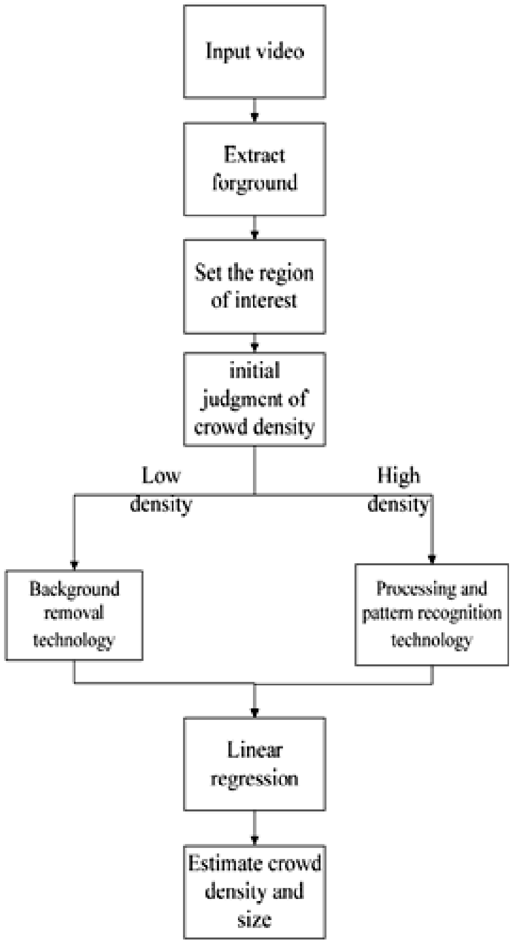

Fig 1: Proposed Method

During the training, we optimize the coefficients of this linear mapping, so that the density functions resulted from the mapping matches the ground truth density functions (defined as a sum of Gaussian kernels cantered at the user-provided dots) as closely as possible. Additionally, we impose a quadratic regularization on the mapping coefficients to prevent over fitting.

Table 2: Result Obtained

Classi-fied level	Extre-mely low Density	Low Density	Mode-rate Density	High Density	Extremely High Density
Number of people dected	4	13	45	72	200
Actual Num-ber of people dected	4	13	40	65	240

Table 2 shows the number of people detected on running our proposed algorithm against he actual number of people. It clearly shows that proposed method gives almost correct results when the crowd density is low but as the crowd density increases the variation from the original data increases and this way the accuracy decreases. Therefore this proposed system proves best when the number of crowd is less.

4. CROWD DENSITY ESTIMATION

Not all events with large gathering of people are conducted in an enclosed venue with turnstiles where crowd density estimation can be administered seamlessly. And for some events such as parades or political protest, employing professionals to conduct human counting is infeasible. Nevertheless, estimating density of crowd is of utmost importance to better administer the well-being of crowd as a whole, development of public space design and accurate documentation of historical events.

The Hillsborough disaster [3] is an example of the consequences of overcrowding. On the contrary to the former two aspects of crowd behaviour analysis, crowd density estimation is independent of the 'thinking' component of each entity in crowd. Existing work on crowd density estimation depends mainly on collective motion and appearance cues, with respect to the type of inputs (*i.e*., crowd video sequences or single crowd image). Different techniques are adopted to cope with crowd scene of varying density. The greater density of crowd in a scene, the more complicated the task to estimate crowd density where dynamic occlusions come into picture. It is infeasible to discerned different person and ones' body parts when a person may only be occupying few pixels [6] and further rendered by background clutter.

For instance, framework that performs clustering of coherent trajectories to represent a moving entity, and inferring number of individual in the scene by Rabaud and Belongie [15], is limited to crowd scenes of sparse crowd where continuous sets of image frames are accessible. The results presented in their work illustrated that for some crowd scene where individuals are closely positioned with each other, trajectories are incorrectly merged. This is due to the phenomenon of collective motion occurring between moving interacting entities. Using an analogous perception, Li et. al. [11] estimate the numbers of people in crowd by implementing foreground segmentation and head–shoulder detection approach. The proposed method was intended to address stationary crowd, where subtle motions of individual is crucial and deeply relied on in defining foreground segments. Nonetheless, the proposed framework is susceptible to inter-occlusion between individuals, particularly prominent in a dense crowd scene. Ge and Collins [12], proposed a Bayesian marked point process to detect individuals in crowd where clear silhouette of individuals is required for accurate projection to a trained set for accurate detection and counting of individuals.

In another study, Ge and Collins [13] uses a generative sampling-based approach that leverage on multi-view geometry to achieve estimation of density of individuals in crowd. The work assumes that individuals in a crowd retain a certain space with each other (i.e. separation), which is one of the rules of interaction between entities in the crowd. Hence, individuals should not be occluded from all viewing angle. Alleviating the need to detect each person in a crowd, some works uses low level crowd features formed based on the collectives of crowd to estimate crowd density. Marana et. al. [14] presented a method based on texture analysis to estimate crowd density, where the estimation is given in terms of discrete ranges (*i.e*., very low, low, moderate, high and very high). Their objective was to challenge scenes of dense crowd where each individual is greatly occluded. They assumed that crowd scene of high density tend to illustrate fine textures, whereas crowd scene of low density are mostly made up of coarse patterns.

5. THE FORTHCOMING CROWD BEHAVIOR ANALYSIS

There are several aspects of crowd behaviour analysis which the authors believe are at their infancy and have the potential to develop further.

Stationary crowd: Crowds may essentially develop into two types, *i.e*., stationary or dynamic crowds. Stationary crowds are usually found as spectators or audiences at concerts, rallies, performances and speeches. Dynamic crowds are defined as crowd which is on the move, such as pilgrims that walk around the Kaaba during Hajj. Most of the existing work on crowd focuses on moving patterns of individuals in the scene to infer their activities. Motion is often detected by using standard approaches such as frame-differencing to more complicated techniques such as dense optical flow. The estimated motion patterns are then analyzed to deduce various suggestions on the crowd activities. On the other hand, stationary crowd analysis has never been sufficiently investigated although the non-motion characteristics can provide rich information. This counter-intuitive approach of stationary crowd analysis is based on the notion that individuals or groups that remain in a particular area for a long time are worthy of attention. The system is able to cope with hundreds of people moving around in a busy scene, to detect abandoned object as long as the object is visible for 50% of the time. In a more advanced and recent work [20], a stationary crowd analysis method is proposed to detect four major activities; group gathering, stopping by, relocating and deforming. This work alludes to the findings of [16], where their simulation on groups in crowd shows that stationary groups have greater impact on the dynamics of the scene than moving groups in some cases. This is justified further by simulating individuals forming stationary groups. The formation of stationary groups acts as an obstruction that changes the motion directions and dynamics of other individuals in the scene. Stationary crowd analysis is still at its early stage of research and is definitely worthy of upcoming investigations for a broader degree of scene under-standing and traffic pattern analysis, in particular.

6. CONCLUSION

This paper proposes an approach for crowd density estimation, which combines the pixel statistical feature and texture feature. The proposed method removed background with Gaussian mixture model and gave preliminary judgment for the crowd density through pixel feature, meanwhile reduced the impact of perspective distortion by dividing the region of interest. The texture features were extracted using GLCM, and selected Contrast 0° and Homogeneity 0° as texture feature. Experimental and comparative results (as discussed in section 3.4) show that the method is an effective, universal method which can be used in a real-time crowd density estimation system. And this paper estimated the crowd size for high density and extremely high density, which was more conducive to group events analysis. As a future work of this study, different other estimation techniques like neural network, textures, etc. can be used and a comparison between them can be done.

7. REFERENCES

[1] A. C. Davies, J. H. Yin, S. A. Velastin, 1995. Crowd monitoring using image processing. Electronics and Communications Engineering Journal, February , pp. 37-47.

[2] Chow T. W. S. Cho, S. Y. C. T. Leung, 1999. A Neural based Crowd Estimation by Hybrid Global Learning

Algorithm. IEEE Transaction on Systems, Man, and Cybernetics, pp. 535-541.

[3] Da Fontoura Costa L. Lotufo R. Velastin S Marana, A., 1999. Estimating Crowd Density with Minkowski Fractal Dimension, Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 3521-3524.

[4] Velastin S. Costa L. Lotufo R. Marana, A., 1998. Automatic estimation of crowd density using texture. Safety Sci, Volume 28, Issue 3, pp. 165-175.

[5] Velastin, S., Yin, J., Davies, A., Vicencio-Silva, M., Allsop, R., Penn, A., 1994. Automated Measurement of Crowd Density and Motion Using Image Processing. Road Traffic Monitoring and Control, In Seventh International Conference, pp. 127–132.

[6] Antonio Albiol, Maria Julia Silla, Alberto Albiol and Jos´e Manuel Mossi, 2009. Video Analysis using Corner Motion Statistics. Proceedings 11th IEEE International Workshop on PETS, Miami,.

[7] Donatello Conte, Pasquale Foggia, Gennaro Percannella, 2010. A Method for Counting Moving People in Video Surveillance Videos, EURASIP Journal on Advances inSignal Processing.

[8] A. N. Marana, L. F. Costa, R. A. Lotufo, S. A.Velastin, 1998. On the efficacy of texture analysis for crowdmonitoring. Computer Graphics, Image Processing, andVision, pp. 354-361.

[9] Schofer J. Ushpiz A. Polus, A., 1983. Pedestrian Flow and Levelof Service. Journal Transportation Eng, Volume 109, Issue 1, pp. 46-56.

[10] Zi Ye, Jinqiao Wang, Zhenchong Wang, HanqingLu, 2012.Multiple features fusion for crowd density estimation.Proceeding ICIMCS'12 Proceedings of the 4thInternational Conference on Internet MultimediaComputing and Service, pp. 42-45.

[11] Subburaman V B, Descamps A,2012. Carincotte C.Counting People in the Crowd Using a Generic HeadDetector[C]. Proceedings of 2012 IEEE 9thInternationalConference on Advanced Video and Signal-BasedSurveillance (AVSS): September 18-21, 2012. Beijing,China, pp. 470-475.

[12] M. Thida, Y. L. Yong, P. Climent-Perez, P. Remagnino, E. L. How, 2013. A LiteratureReview on Video Analytics of Crowded Scenes, Intelligent Multimedia Surveillance,Springer.

[13] B. Zhan, D.N. Monekosso, S.A. Remagnino, P. Velastin, L.-Q. Xu, Crowd analysis:a survey, Mach. Vis. Appl. Volume 19, Isue 5–6, 345–357.

[14] J. C. S. Jacques Junior, S. Raupp Musse, C. R. Jung, 2010, Crowd Analysis using Computervision Techniques, Signal Process. Mag., Volume 27, Issue 5, pp. 66–77.

[15] N. N. A. Sjarif, S. M. Shamsuddin, S. Z. Hashim, 2012. Detection of Abnormal Behaviorsin Crowd Scene: A Review, International Journal Advance Soft Comput. Applications, Volume 4, Issue 1, pp. 1-33.

[16] C. C. Loy, K. Chen, S. Gong, X. Tao, 2013. Crowd Counting and Profiling: Methodology and Evaluation, in Modeling, Simulation and Visual Analysis of Crowds, Springer, pp. 347–382.

[17] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, S. Yan, 2014. Crowded scene analysis: asurvey, Circuits Syst. Video Technol. (CSVT), Volume 25, Issue 3, pp. 367–386.

[18] R. L. Hughes, 2003. The Flow of Human Crowd, Annual Review Fluid Mech., Volume 35. Issue 1, pp. 169–182.

[19] Richard Leggett, 2004. Real-Time Crowd Simulation: A Review, R. Leggett.

[20] V. Alexiadis, K. Jeannotte, A. Chandra, 2004. Traffic Analysis Toolbox Volume i:Traffic Analysis Tools Primer, Technical Report.