Approach Based Study of Crowd Size Analysis

Jugal Kishor Gupta[1], Dr. S K Gupta[2]
Research Scholar[1], Assistant Professor[2],
Department of Computer Science & Engineering[1, 2]
IFTM University[1]
Bundelkhand Institute of Engineering & Technology[2]
Email: .jugalkg33@gmail.com[1], guptask_biet@rediffmail.com[2]

Abstract- We are going to analyze crowd by different approaches like public tracking, mob density estimation, and validation, simulation and event detection. In the current paper we are going to put light on how we can analyze crowd using computer vision and computer graphics. This technique is found to be very helpful and appropriate for the points covered.

Index Terms- Computer Vision; Crowd Behavior; CGT.

1. INTRODUCTION

In recent past some systems were designed and developed to recognize, understand and track the behavior of various variety objects, commercially using one or multiple cameras on objects for processing the information from computer to computer. As human behavior is proving to be an area for never ending scientific interest and research. One person could be occluded in many ways and could be detected and identified by analyzing behavior, reorganization and continuous monitoring and could prove to be an asset for research [1]. When analyzing at high level of details, we could segment body parts (eg.. Head, face, hand and arms) and could use in gesture recognition or human-machine interaction [2,3]. We can also deploy the same technique for group behavior analysis, once people can be part of high-level structures for example groups or crowds and off-course behavioral analysis of crowded scenes are of great interest in a large number of applications[4]; such as :-

- ✓ Crowd Management: To ensure public safety and avoid crowd related disasters we could deploy crowd management strategies.
- ✓ Public Space Design: They help in tracking guidelines for creating designs for public space.
- ✓ Virtual Environment: The simulation performance of mathematical models could be enhanced and validate in Virtual Environment.
- ✓ Visual Surveillance: Anomalies and alarms could be recognized automatically using Visual Surveillance.
- ✓ Intelligent Environment: Based on behavior, Intelligent Environment could be used for taking decision to split a crowd in a museum.

The approach number two facilitates us to assume the crowd as a standalone entity and individual analysis could be avoided, crowded video sequences are used to extract information using computer vision and this information could benefit many applications(such as;

Surveillance, design of densely populated public space and analysis of crowds safety such as crowd entering sports arena)[6-8].

Reproduction of realistic crowds using computer graphic techniques for simulation is another challenge for crowd phenomenon. Animations of crowds find applications in many areas, including entertainment (e.g., animation of large numbers of people in movies and games), creation of immersive virtual environments, and evaluation of crowd management techniques (for instance, simulation of the flow of people leaving a football stadium after a match). As related by Zhan et al. [4], crowd information can be better exploited to indicate the status of the crowd so that crowd events can be inferred. Crowd models have been built to represent this status, either implicitly or explicitly. Computer vision and graphics when used together proves to be an asset for integrated synthesis and analysis of crowd, which was not possible with the systems available. Though validation remains a question mark when we analyze crowd using computer vision alone because gathering realistic data for real crowd scenes is not an easy task (specially in emergency conditions), data gathered using computer graphic techniques would be used by computer vision algorithms for validation and training. By far crowd simulation models can prove to be more realistic if they are provided with information captured from real world.[12,14] As said by "Musse At All"[12] computer vision technique could automatically capture trajectories of individuals in non-crowded areas and this data could be used in generating extrapolated velocity maps that provides main directions and velocities to simulation models in analyzing filmed environments, though research provides many simulation models but non is a perfect fit because real life scenes with freedom of mobility around 360 degree and complex human behavior proves to be more challenging. Using computer vision with crowd simulation techniques arises questions like; what degree/level of realism can be achieved?

What behavior can be realistically simulated using crowd simulation? These questions are required to be answered any how.

2. CROWD DYNAMICS

An important problem in crowd analysis is people counting/density estimation (either in still images or video sequences). For instance, crowd density analysis could be used to measure the comfort level in public spaces, or to detect potentially dangerous situations [15].

The work described in [5] was one of the pioneers to use computer vision techniques for obtaining automatically some kind of information from crowds. In their work, the authors proposed an approach to estimate the density of the crowd using pixellevel information.

Regazzoni et al. [21] proposed an approach to estimate the crowd density in images. In their work, features extracted from each acquired image (basically the results of an edge detection algorithm and finding vertical edges) are related to the number of people present in the monitored scene by using nonlinear models obtained by means of dynamic programming in an offline training phase.

Cho and collaborators [23] presented a method based on neural networks to estimate the crowd density in subway stations. Their objective was to detect high-density situations and to provide statistical analysis about the flow of people across time for activity planning. They proposed a simple edge detector based on binary image thresholding and explored the length of detected edges as a feature for people counting through a single hidden layer neural network.

Yang et al. [24] proposed a multicamera-based method to segment people and to estimate the number of people in crowded video sequences. In their work, groups of image sensors were used to segment foreground objects from the background, aggregate the resulting silhouettes over a network, and compute a planar projection of the scene's visual hull.

Ma and colleagues [25] proposed a crowd density estimation algorithm for surveillance purposes. In their work, an approach based on pixel counting of foreground objects is used to estimate the crowd density, combined with a projective correction using a calibrated camera.

The density over time is also monitored, aiming to detect some unusual behavior. Their approach suffers the same occlusion problems faced in [5], since a linear model is used to estimate the people count. Kong and collaborators [7] presented a method based on learning to estimate the number of people in crowds. Edge orientation and the histogram of the object areas (extracted from foreground objects through a background subtraction algorithm) are used as image features.

3. CROWD ANALYSIS

Estimating people density/count is another issue in analyzing crowd; As this count could be a measure for potentially dangerous situations and comfort zone in public space. Pixel-level information was utilized by authors[5] to estimate the crowd density using Computer Vision Techniques. Images were used to estimate crowd density by "Regazzoni At All"[21], he explained the concept for extracting features from images and relating them to number of people present in monitored scene by using non-linear models and dynamic programming in off-line training phases.

A model to analyze statically and detecting high density conditions about flow of people across time for planning activities using Neural-Networks for crowd estimation was purposed by Cho and Collaborators[23].

In another approach Yang Et AL[24] has used multiple cameras for segmenting people and calculating individuals in crowded video sequences, they have used the scene's visual hull to compute a planar projection by using sensors and segmenting foreground objects from background objects.

A crowd density estimation algorithm for surveillance purposes was proposed by Ma and colleagues [25]. They defined an approach based on pixel counting combined with a projective correction using a calibrated camera of foreground objects and used that to estimate density of crowd, unusual behaviors were detected by monitoring density over time. Their approach suffers the same occlusion problems faced in [5], since a linear model is used to estimate the people count. Kong and collaborators [7] presented a method based on learning to estimate the number of people in crowds. Edge orientation and the histogram of the object areas (extracted from foreground objects through a background subtraction algorithm) are used as image features.

4. TEXTURE ANALYSIS

Four methods were used by Marana et al. [26] in texture analysis and three classifiers to deal with the crowd density estimation problem. Regarding texture analysis, they compared the following four methods: gray-level dependence matrix, straight lines segments, Fourier analysis, and fractal dimension. Regarding the classifiers, they compared the following three methods: neural network, statistics (Bayesian), and a fitting function-based approach. They found better results when using the gray-level dependence matrix-based method, providing better contrast and homogeneity as texture features, combined with a

Bayesian classifier. However, it should be noted that they generated ground-truth information empirically, which could affect the comparison. They estimated the crowd density in one of the five following classes: very low density, low density, moderate density, high density, and very high density. The authors mentioned that the method can not discriminate very well the difference between high and very high densities. support vector machines (SVMs) and texture analysis were used by Wu et al. [27] in estimating crowd density. A series of multi-resolution cells were generated using perspective projection model, textural information within these cells were extracted by gray-level dependence matrix method. Actual density of the scene was related to textural features using SVM and multi scale texture vector, a maximum estimated error of below 5% was measured for each cell , and proposed as future work the possibility of including a background subtraction method in the feature extraction stage.

5. OBJECT LEVEL ANALYSIS

Object level analysis tries to identify individual object in a scene. More accurate result produce when compared to pixel level analysis or texture based approaches, but identifying individual in a single image or a video sequence is mostly of erasable in lower density crowds.

Denser crowds clutter and severe occlusions make the individual country problem almost impossible to solve despite the recent advances of computer vision and pattern recognition techniques. Lin et al. [30]

Proposed an algorithm to estimate the crowd density in three stages

1. They searched for objects with head like contour in the image space, using the haar wavelet transform.
2. The feature of the object is analyzed, using an SVM aiming to classify it as a head or not.
3. Finally, a perspective transformation is done aiming better estimate the density of the entire crowd. Zhao and Nevatia [31] proposed a Bayesian approach to segment people in crowds.

Brostow and Cipolla [35] presented an unsupervised Bayesian clustering method to detect independent movements in a crowd. Their hypothesis is that a pair of points that move together should be part of the same entity. An interesting characteristic of [35] is that it does not require any training stage or appearance model to track individuals.

6. TRACKING IN CROWD SCENES

Tracking people and identification of individual's position in frames of sequence in crowd is another problem, exploring individual trajectories and crowd flow can be helpful in detecting abnormal behaviors. However existing problems like object tracking, cluster and particle or complete omission make individual tracking a challenge in denser crowds. Reference about generic tracking algorithms can be found in [4]. An extension for tracking is presented in people counting approach [33], for an auto-regressive estimation of foreground information they have used color signature, an appearance template and probabilistic target mask in their approach. People tracking approach in structured high density scenario is given by ALI and SHAH [28], they divided video sequence frames into cells presenting one particle. Every individual is consists of particles and individuals are affected by scene's layout. They have used crowd modeling concept in their approach.

7. CROWD ANALYSIS HELPING CROWD SYNTHESIS

Crowd analysis and crowd simulation are complimentary to each other. For instance, to analyze how physical changes in environment would affect people flow virtual world is required to create real scenarios and using that we can analyze people in specific scenario (density, directions, and velocity field). Extraction of data using traditional scenarios (manually) is very time consuming and error prone, computer vision algorithms proves to be very motivating for the same. Courty and Corpetti [13] gives an approach to capture information from real crowds and feed a crowd simulator. In their work, an optical flow-based method is used to capture the movements of people in a given environment.

Musse et al. [12] and Paravisi et al. [14] also presented different approaches to capture information from the real world; they have used computer vision techniques to create realism by simulating crowd. However, above given concepts tracking based on the individuals trajectories in non-crowded scenarios are provided instead of global movements. Musse and collaborators [12] presented some quantitative comparisons between the real and the virtual scenarios, such as the average speed of people. Paravisi et al. [14] proposed a "spatial occupancy map" to measure the spatial distribution of people, which was used for comparison purposes. However, evaluating simulation results of a synthetic scenario that does not relate to a real one is a very complex task. The motion and/or behavior of a crowd can vary depending on several aspects, such as culture and context. It is important to mention that defining quantitative metrics to evaluate the realism of a crowd simulation result is still an open problem.

8. CONCLUSION

Computer Vision based crowd analysis is done in this paper through survey. We have taken in account some major aspects of crowd analysis; Understanding behavior of crowd, tracking crowded scenes and estimating people density/count along with detection of unusual situations, estimating velocity, temporal evolution of different approaches to analyze crowd was revived and we found that Computer Vision and Computer Graphics based techniques are better for crowd analysis but still after analyzing above mention point we came to conclude that serious work is required to be carried out in future to develop techniques for crowd analysis

REFERENCES

[1] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-basedhuman motion capture and analysis," *Comput. Vis. Image Understand.*, vol. 104, no. 2, pp. 90–126, 2006.

[2] J. C. S. Jacques, Jr., A. Braun, J. Soldera, S. R. Musse, and C. R. Jung, "Understanding people motion in video sequences using voronoi diagrams," *Pattern Anal. Applicat.*, vol. 10, no. 4, pp. 321–332, 2007.

[3] P. Kilambi, E. Ribnick, A. J. Joshi, O. Masoud, and N. Papanikolopoulos, "Estimating pedestrian counts in groups," *Comput. Vis. Image Understand.*, vol. 110, no. 1, pp. 43–59, 2008.

[4] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, and L.-Q. Xu, "Crowd analysis: A survey," *Machine Vis. Applicat.*, vol. 19, no. 2, pp. 345–357, 2008.

[5] A. C. Davies, J. H. Yin, and S. A. Velastin, "Crowd monitoring using image processing," *IEE Electron. Commun. Eng. J.*, vol. 7, no. 1, pp. 37–47, 1995.

[6] H. Rahmalan, M. Nixon, and J. Carter, "On crowd density estimation for surveillance," in *Proc. Institution of Engineering and Technology Conf. Crime and Security*, 2006, pp. 540–545.

[7] D. Kong, D. Gray, and H. Tao, "A viewpoint invariant approach for crowd counting," in *Proc. Int. Conf. Pattern Recognition*, 2006, vol. 3, pp. 1187–1190.

[8] S. Ali and M. Shah, "A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2007, pp. 1–6.

[9] S. R. Musse, M. Paravisi, R. Rodrigues, J. C. S. Jacques, Jr., and C. R. Jung, "Using synthetic ground truth data to evaluate computer vision techniques," in *Proc. IEEE Int. Workshop on Performance Evaluation of Tracking and Surveillance*, 2007, pp. 25–32.

[10] G. Taylor, A. Chosak, and P. Brewer, "OVVV: Using virtual worlds to design and evaluate surveillance systems," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, June 2007, pp. 1–8.

[11] E. L. Andrade and R. B. Fisher, "Simulation of crowd problems for computer vision," in *Proc. 1st Int. Workshop on Crowd Simulation*, 2005, pp. 71–80.

[12] S. R. Musse, C. R. Jung, J. C. S. Jacques, Jr., and A. Braun, "Using computer vision to simulate the motion of virtual agents," *Comput. Animation Virtual Worlds*, vol. 18, no. 2, pp. 83–93, 2007.

[13] N. Courty and T. Corpetti, "Crowd motion capture," *Comput. Animation Virtual Worlds*, vol. 18, no. 4–5, pp. 361–370, 2007.

[14] M. Paravisi, A. Werhli, J. C. S. Jacques, Jr., R. Rodrigues, A. Bicho, C. Jung, and S. R. Musse, "Continuum crowds with local control," *in Proc. Computer GraphicsInt. (CGI'08)*, Istambul, Turquia, June 2008, pp. 108–115.

[15] J. Fruin, *Pedestrian and Planning Design.* Mobile, AL: Elevator World Inc. 1971.

[16] B. A. Boghossian and S. A. Velastin, "Motion-based machine vision technique for the management of large crowds," in *Proc. 6th IEEE Int. Conf. Electronics, Circuits and Systems*, 1999, vol. 2, pp. 961–964.

[17] H. Benesch, *Atlas de la Psychologie.* France: Encyclopdies d'Aujourd'huik, 1995.

[18] D. Thalmann and S. R. Musse, *Crowd Simulation.* New York: Springer- Verlag, 2007.

[19] K. Still, "Crowd dynamics," Ph.D. dissertation, Univ. Warwick, England, 2000.

[20] E. T. Hall, *The Silent Language.* New York: Doubleday, 1959.

[21] C. S. Regazzoni, A. Tesei, and V. Murino, "A real-time vision system for crowding monitoring," in *Proc. Int. Conf. Industrial Electronics, Control, and Instrumentation*, 1993, vol. 3, pp. 1860–1864.

[22] C. Ottonello, M. Peri, C. Regazzoni, and A. Tesei, "Integration of multisensory data for overcrowding estimation," in *Proc. IEEE Int. Conf. Systems, Man and Cybernetics*, Oct. 1992, vol. 1, pp. 791–796.

[23] S.-Y. Cho, T. W. S. Chow, and C.-T. Leung, "A neural-based crowd estimation by hybrid global learning algorithm," *IEEE Trans. Syst., Man, Cybern.*, 1999, vol. 29, no. 4, pp. 535–541.

[24] D. B. Yang, N. Héctor, H. González-Ba, and L. J. Guibas, "Counting people in crowds with a real-time network of simple image sensors," in *Proc. IEEE Int. Conf. Computer Vision*, Washington, DC, USA, 2003, p. 122.

[25] R. Ma, L. Li, W. Huang, and Q. Tian, "On pixel count based crowd density estimation for visual surveillance," in *Proc. IEEE Conf. Cybernetics and Intelligent Systems*, 2004, vol. 1, pp. 170–173.

[26] A. Marana, L. da Costa, R. Lotufo, and S. Velastin, "On the efficacy of texture analysis for crowd monitoring," in *Proc. Int. Symp. Computer

Graphics, Image Processing, and Vision (SIBGRAPI'98), Washington, DC, 1998, p. 354.

[27] X. Wu, G. Liang, K. K. Lee, and Y. Xu, "Crowd density estimation using texture analysis and learning," in *Proc. IEEE Int. Conf. Robotics and Biomimetics*, 2006, pp. 214–219.

[28] A. Chan, Z. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2008, pp. 1–7.

[29] A. B. Chan and N. Vasconcelos, "Bayesian Poisson regression for crowd counting," in *Proc. IEEE Int. Conf. Computer Vision*, 2009, pp. 1–7.

[30] S.-F. Lin, J.-Y. Chen, and H.-X. Chao, "Estimation of number of people in crowded scenes using perspective transformation," *IEEE Trans. Systems, Man, Cybern. A*, vol. 31, no. 6, pp. 645–654, 2001.

[31] T. Zhao and R. Nevatia, "Bayesian human segmentation in crowded situations," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2003, vol. 2, pp. 459–466.

[32] E. Leibe, B. Seemann, and B. Schiele, "Pedestrian detection in crowded scenes," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Washington, DC, 2005, pp. 878–885.

[33] J. Rittsche, P. H. Tu, and N. Krahnstoeve, "Simultaneous estimation of segmentation and shape," in *Proc. Computer Vision and Pattern Recognition*, Washington, DC, 2005, vol. 2, pp. 486–493.

[34] V. Rabaud and S. Belongie, "Counting crowded moving objects," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2006, pp. 705–711.

[35] G. J. Brostow and R. Cipolla, "Unsupervised Bayesian detection of independent motion in crowds," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Washington, DC, 2006, pp. 594–601D. B. Yang, N. Héctor, H. González-Ba, and L. J. Guibas, "Counting people in crowds with a real-time network of simple image sensors," in *Proc. IEEE Int. Conf. Computer Vision*, Washington, DC, USA, 2003, p. 122.

[36] M. J. Jones and D. Snow, "Pedestrian detection using boosted features over many frames," in *Proc. Int. Conf. Pattern Recognition*, 2008, pp. 1–4.

[37] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Comput. Surv.*, vol. 38, no. 4, 2006, pp. 13.1–13.45.

[38] S. Ali and M. Shah, "Floor fields for tracking in high density crowd scenes," in *Proc. European Conf. Computer Vision*, 2008, pp. II:1–14.

[39] M. Rodriguez, S. Ali, and T. Kanade, "Tracking in unstructured crowded scenes," in *Proc. IEEE Int. Conf. Computer Vision*, Kyoto, Japan, 2009, pp. 1389–1396.

[40] A. M. Cheriyadat and R. Radke "Detecting dominant motions in dense crowds," *IEEE J. Select. Topics Signal Process.*, vol. 2, no. 4, pp. 568–581, Aug. 2008.