An Innovative Crowd Segmentation Approach based on Social Force Modelling

Yu Hao[1,2], Zhijie Xu[2], Ying Liu[1], Jing Wang[3], Jiulun Fan[1]
Xi'an University of Posts and Telecommunications[1], University of Huddersfield[2], Sheffield Hallam University[3]
Xi'an, China[1], Huddersfield , UK[2], Sheffield, UK[3]
E-mails: haoyu@xupt.edu.cn, z.xu@hud.ac.uk, ly_yolanda@sina.com, jing.wang@shu.ac.uk, jiulunf@xupt.edu.cn

***Abstract*—The effective segmentation and separation of groups in a crowd serves a crucial role in the video analysis for crowd behavior understanding. Conventional crowd segmentation techniques rely on the spatial distribution or temporal trajectories of individual agent in the crowd. However, psychological influences among subjects in a high density crowd are often ignored. In this research, an enhanced Social Force Model (SFM) is developed to define in-crowd motions among agents and its affiliated groups. An innovative SFM descriptor - the Grouping Center - is modeled as a key feature for aiding accurate crowd segmentation. During the process, extracted optical flows are mapped to the detected agents, then the SFM components such as the desired force and the repulsive force are calculated. Experiments show accurate segmentation can be achieved even when agents from different groups are heavily (spatially) mixed up. Random crowd scenarios generated by a commercial game engine are used for crowd behavior analysis and model evaluation. Evaluations indicate that the new SFM model and its related-techniques are capable of handling complex crowd scenarios with latent semantic meanings.**

Keywords- crowd segmentation; motion prediction; social force model.

I. INTRODUCTION

The analysis of crowd behaviors is becoming a crucial research issue to support the maintenance of public safety. With the wide installation of Close Circuit Television Inspection (CCTV) cameras in public area, the massive amount of recorded video data could be exploited for the prediction of potential danger, real-time anomaly alerts and forensic evidences. The segmentation of the crowd in video data usually serves as the pre-processing stage in the framework of crowd analysis, yet it is a vital component of the entire pipeline. The precise segmentation of the crowd would has significant influence on the performance of feature extraction, model training and behavior recognition.

The technique of Crowd Segmentation [1][6] usually consists two fundamental procedures, which are Individual Detection and Motion Prediction. The aim of Individual Detection is to verify as many agents from the crowd as possible. In most crowded videos, the density of the crowd is high. The high density will cause the difficulty to maintain the accuracy of individual detection, for example the issue of frequent occlusion will make the tracking and locating of the pedestrian extremely difficult. In order to address this issue, approaches of individual detection under heavy occlusion are proposed by various researches [1]-[7]. Some other segmentation approaches attempt to consider the crowd as a single entity to avoid the individual detection. By analyzing the flow-based information of this entity, the crowd is segmented according to its spatial and temporal relationship. However, this kind of approach is not capable of handling the following situation. Assuming agents with two different destination are randomly distributed, the approach cannot segment these agents into two groups until they are spatially separated, or the temporal information such as trajectory are calculated after some amount of time. In order to achieve the successful segmentation of these two groups in the early stage, this research devised a segmentation approach based on Individual Detection and the concept of Social Force. This approach will be able to segment the randomly distributed crowd according to the proposed Group Attraction Force. The detailed algorithm will be introduced in following section. The aim of Motion Prediction is to predict the long-term behaviors of detected individuals according to their spatial and temporal information. In the procedure of Motion Prediction, states such as position and trajectory are fed to a behavioral model to estimate the individual's next motion. Once the long-term behavior of each detected individual is obtained, a classifier could be implemented to cluster the estimated behavior for the segmentation of the crowd.

The techniques of individual detection consist of two approaches. In the first approach, the statistical or shape information is extracted from the image and feed to the trained classifier for the detection result [1]-[5]. In the second approach, a rough estimation such as head/shoulder detector is implemented to generate hypotheses. Then the concept of Expectation Maximization (EM) is exploited to verify these hypotheses based on various iterations [6][7].

In the research of Lan [1], the image is firstly preprocessed into a foreground image. Then the silhouette of the foreground image is sampled and transformed with Discrete Fourier Transform as a Fourier Descriptor. The silhouette can be reconstructed anytime with Fourier coefficients. The Euclidean distance between two Fourier Descriptors is utilized to measure the similarity of silhouettes. Next, KNN and locally-weighted regression are used to find the closest match of target image among the trained sample sets. The machine learning techniques are frequently utilized as well in recent researches. Kang [3] mentioned for crowd with high density, appearance-based approaches with patterns such as HOG, Scale-invariant feature transform (SIFT) and Linear Binary Pattern (LBP) do not have good performance. Techniques of Deep Learning are widely implemented in computer vision field. In the research, the so-called Fully Convolutional Neural Network (FCNN) is used for crowd segmentation with both appearance and motion features. In the FCNN, a multiple stage deep learning architecture with features of appearance, motion and structure is devised. In this research, the FCNN claims to have advantages than the conventional CNN.

In the research of Tu [6], the concept of EM is utilized to verify the segmentation hypothesis, and it attempts to

 ISBN: 978-1-61208-677-4

achieve the correct detection under serious occlusion situation. In the first step of this approach, a rough head and shoulder detection are applied on the image, each detected pedestrian is named hypotheses. However, due to the complexity of scene and the occlusion problem, many false positive detections will occur. For the second step, the image is divided into a grid of patches. Next, this research proposed an algorithm to calculate the so-called Affinity between each patch and hypotheses. In order to further optimize the segmentation hypotheses, the concept of EM is imported to address this issue. An approach with similar methodology, but different details is proposed in [7]. In this approach, a scale-invariant interest point operator [8] is firstly used to locate the Points of Interests. Around these points, patches with certain radius are extracted from the image. Next, an agglomerative clustering scheme [9] is implemented on each patch to obtain a representation of its structure as the signature for the hypothesis's verification. Then the spatial occurrence distribution is used to measure the distance between different signatures. For each iteration of the E-step, patches are extracted from the interest points, and the distance is measured by the trained codebook using spatial occurrence distribution to verify the hypotheses. And for the M-step, the measured result will be used to update the codebook. After several iterations, the updated hypotheses are recognized as detected pedestrians.

With the premise of well devised agent behavioral model and successfully detected agent motion state, it is expected to have an accurate prediction of agent's next movement. However, in real-life cases, this premise is very difficult to fulfill. It will make conventional motion prediction approaches do not show a satisfying performance in real-life scenario. Furthermore, the conventional approach is only capable of making the prediction of short-term movement. As long as the approach is applied to the longer-term prediction, the accuracy decays drastically. In order to address this issue, two alternative approaches are proposed by the following researches. 1) Instead of detecting the motion state of an agent, a grid is placed onto the current image. For each cell within this grid, the transition probability is calculated for an agent moving to next cell [10] [11]. 2) All trajectories inside scene are firstly detected. Then trajectories with similar patterns are clustered. The clustered trajectory will be used for prediction of next motion [12] [13].

The Social Force Model [15] is primarily implemented in the field of crowd simulation, and sometime crowd analysis. In most of the segmentation algorithm, spatial information such as co-ordinate, size and shape, temporal information such as trajectory are exploited. However, the behavior of pedestrians within highly dense environment could be easily influenced by the psychological factor with its neighbors. In this research, the Social Force Model along with the improved concepts such as personal space and perception field are utilized to devise an innovative signature for crowd segmentation namely Grouping Center. The result of following experiments indicates the segmentation approach with psychological concept is capable of handling the mixed crowd from multiple groups.

The structure of this contribution is composed as follows. Section 2 gives a detailed introduction of the proposed crowd segmentation approach. To be specific, the innovative concept of Group Attraction Force is explained and the extraction algorithm is introduced. In Section 3, the proposed algorithm is formulated to achieve a more explicit explanation. In Section 4, the proposed approach is implemented on the simulated crowd video to assess the performance of segmentation. Section 5 provides the conclusion and discusses the future work.

II. Crowd Segmentation Approach at Highly Mixed State

In the previous work [14], a simulation model based on the Social Force Model (SFM) is devised to generate crowd behavior with visual realism. The concept of SFM is firstly proposed by Helbing [15], and widely exploited on crowd simulation and behavioral modeling. The fundamental concept of SFM is the motion of every pedestrian among the crowd is affected by three types of psychological-related forces: Desired Force, Repulsive Force and Avoidance Force, which is formulated as (1). Which sf_i is the final force determines the motion of pedestrian i. f_d is the desired force, f_{ji} is the repulsive force from neighbor j, and f_{oi} is the avoidance force from obstacle o.

$$sf_i = f_d + \sum_{j \neq i} f_{ji} + \sum f_{oi} \qquad (1)$$

In the proposed simulation model, an innovative group attraction force is devised to achieve a better visual realism of the simulated crowd. In the simulation scenario, sixty agents from three different groups are randomly distributed on the stage. Each agent is mapped with a behavioral model which controls the agent's motion. The behavioral model contains parameters such as personal space, radius, angle of perception. The radius describes the size of each agent. The personal space and angle of perception determine the repulsive force affected on the agent. If a neighboring agent exists in the range of personal space and perception angle, the repulsive force will be generated, otherwise it will be omitted. In order to further increase the visual realism, based on the assumption that simulated agents always attempt to stay closer to the cluster with agents from same group, the concept of Group Attraction Force is devised to enhance the conventional SFM. To model the Group Attraction Force, agents with same group number appeared in the perception field of current agent are firstly allocated. Then the average position of the allocated agents is calculated as the Grouping Center. Next, a force from current agent to the grouping center will be estimated as the Group Attraction Force. The simulation result shows the group attraction force embedded behavioral model presents the better performance than the conventional SFM on the visual realism. The enhanced SFM with Group Attraction Force could be adjusted as in (2), in which f_{Gi} is the Group Attraction Force.

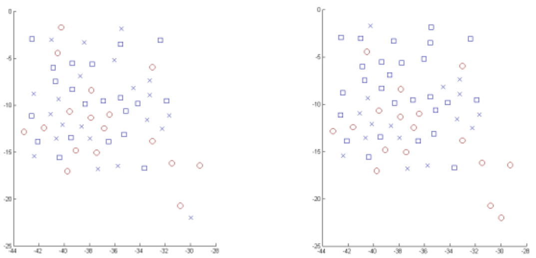
$$sf_i = f_d + \sum_{j \neq i} f_{ji} + \sum f_{oi} + f_{Gi} \qquad (2)$$

By further investigating the distribution of Grouping Center, it could be observed that despite the spatial distribution of agents from three different groups is sparse and random, the Grouping Centers from different groups are naturally clustered. As illustrated in Fig. 1(a), the spatial distribution of agents is very random. It is almost impossible to use classifiers such as KNN to segment. However, according to the Group Attraction Force Model, the calculated distribution of group centers exhibits significant patterns, as illustrated in Fig. 1(b). The Fig.1(c) shows the distribution of grouping centers in large scale. It could be explicitly observed that grouping centers are naturally clustered.

(a) (b) (c)

Figure 1. The comparison of distribution between spatial position and group center. (a) Spatial distribution. (b) Grouping center distribution. (c) Magnified grouping center distribution

Upon this observation, it is worthy to try segmenting agents according to the clustering result of group centers. Therefore, the simple KNN clustering algorithm is exploited for the segmentation. It's assumed that the value of *K* is correctly determined, in this case *K* equals to 3. As illustrated in Fig. 2(a) shows the Ground Truth spatial distribution of agents from three different groups. Agents from same group are labeled with same shape. Fig. 2(b) shows the segmented result using proposed approach. In this case, 53 out of 60 agents are correctly segmented comparing to the ground truth.



(a) (b)

Figure 2. Comparison between ground truth and segmentation result. (a) Ground Truth. (b) Segmentation Result using Grouping Center

The previous paragraphs proved that the Grouping Center is capable of segmenting crowd consists of multiple groups at randomly distributed state. In order to exploit the Grouping Center to achieve desired segmentation performance, a fundamental issue still needs to be addressed. In the previous paragraph, it is assumed that agent knows the group number of other agents. Based on this assumption, the Grouping Center could be calculated. Nevertheless, in the process of segmentation on the crowd video obtained from the CCTV cameras, the group number of each pedestrian is unknown. Thus, the Group Centers for pedestrians cannot be obtained directly. Therefore, an innovative algorithm to estimate Group Attraction Force from two consecutive frames is introduced. The architecture of the algorithm is illustrated in Fig. 3.

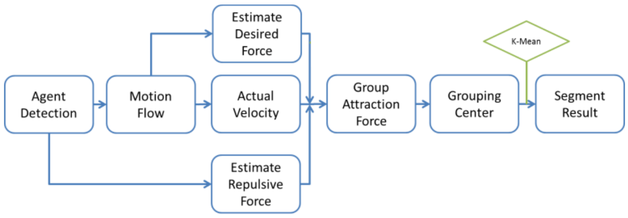


Figure 3. The framework of the proposed crowd segmentation approach

For each agent, the group attraction force f_{Gi} is unknown. if the actual force sf_i, desired force f_d, repulsive force $\sum_{j\neq i} f_{ji}$ and avoidance force $\sum f_{oi}$ are obtained, the group attraction force could be calculated according to (2). With the obtained group attraction force, the Grouping Center will be calculated.

In the procedure illustrated in Fig. 3, every agent needs to be detected at the first step using the main stream pedestrian detection algorithm such as [3]. For each detected agent, the parameters of SFM such as personal space, radius and angle of perception are estimated according to the current environment. The radius of agent could be estimated upon the average pixel number of all detected agents. Once the radius is determined, the personal space could be decided empirically. The angle of perception could be set to the same value of the simulation model. Therefore, the repulsive force $\sum_{j\neq i} f_{ji}$ of all detected agents i could be calculated.

On the other hand, the global flow-based information between two consecutive frames is extracted, and mapped to each agent. The flow-based information such as optical flow could be transformed into the actual velocity of mapped agent and set as the actual social force sf_i. The estimation of desired force f_d could be a difficult task, since the group number of current agents is unknown. Therefore, the destination cannot be determined. In order to address this issue, several constrains have been made to simplify the problem. Firstly, agents are assigned into two groups. The destinations of these groups located at opposite directions on the scene. Secondly, if the actual force of agent is large enough, the direction of desired force is identical to the actual force, since the magnitude of repulsive and group attraction force are usually much smaller than the desired force. Thirdly, if the actual force is not large enough, it will be close to the repulsive force. In this case, the direction of agent's desired force will be randomly determined. Considering the complexity of scenario, the avoidance force $\sum f_{oi}$ is omitted. In the following experiment, the scene only consists with floor, agents and destination. The obstacle is removed from scene.

With the estimated desired force, repulsive force and actual social force, the group attraction force could be estimated according to (2). Since the group attraction force is a vector, the grouping center could be calculated using current position of the agent. With the estimated grouping centers, the conventional classifier such as KNN could be utilized to segment the crowd.

In next section, the proposed algorithm is formulated in order to help better understanding the proposed approach.

III. FORMULATION OF THE PROPOSED APPROACH

According to the framework introduced in Fig. 3, the group attraction force f_{Gi} of agent i could be derived from the (2). As explained in previous section, the avoidance force is omitted in order to simplify the model. The estimated group attraction force f_{Gi} of agent i is presented as (3).

$$f_{Gi} = sf_i - f_d - \sum_{j \neq i} f_{ji} \qquad (3)$$

in which sf_i is the actual social force obtained from the optical flow information. f_d is the estimated desired force, the magnitude of f_d is a constant α, and its direction could be presented as (4).

$$dir(f_d) = \begin{cases} dir(sf_i) & ||sf_i||_2 > \tau \\ \theta & otherwise \end{cases} \qquad (4)$$

where τ is a threshold constant, if the magnitude of actual force sf_i is greater than the threshold τ, the estimated direction of f_d will be same as actual force. Otherwise, the direction of f_d will be randomly determined between two goals.

The calculation of actual force sf_i, is based on the flow-based information. In this research, the conventional Horn Schunck optical flow patterns [17] are first extracted from two consecutive frames. Next, the standard pedestrian detection algorithm is applied to obtain the position of each agent. By mapping the optical flow to each agent, the actual force could be determined. The actual force sf_i could be represented as (5).

$$sf_i^{x,y} = \sum k f_o^{x\pm1,y\pm1} \qquad (5)$$

where $sf_i^{x,y}$ is the actual force of agent i at position xy, $f_o^{x\pm1,y\pm1}$ is the optical flow vector on position xy, k is the weight factor, when $x=0$ and $y=0$, $k=1$, otherwise $k=0.4$ in this case.

The calculation of repulsive force f_{ji} is imported from Qingge's Velocity Perception Based SFM [16]. In the conventional SFM, the repulsive force is set as a constant magnitude value β. If the distance between current agent and any neighboring agents is less than personal space ρ, the repulsive force will be applied to the agent, otherwise no repulsive force will be applied. This solution will generate the significant 'vibration' phenomenon between agents. Instead of using a constant value, (6) is devised to describe the most realistic repulsive force between agent i and j.

$$f_{ji} = A_i e^{(r_{ij} - d_{ij})/B_i} n_{ij} \qquad (6)$$

where A_i and B_i are the constant values to control the magnitude scale of the repulsive force. r_{ij} is the sum of two agents' radius. d_{ij} is the distance between two agents. The exponential function ensures the fast dispersing when the distance between two agents increase and vice versa. n_{ij} describes the direction of the repulsive force. In order to further increase the realism of simulated repulsive force, the personal space ρ and perception field are also considered. The extracted repulsive force is further regulated using (7). If the distance between two agents is smaller than the summation of agent i's personal space and agent j's radius, the repulsive force remains the same. Otherwise the repulsive force equals to zero when the distance between two agents is not in the range of the personal space.

$$f_{ji} = \begin{cases} f_{ji} & d_{ij} \leq \rho_i + r_j \\ 0 & otherwise \end{cases} \qquad (7)$$

Even if the distance satisfied the range of personal space, the perception of agent and the motion direction could still need to be specifically modeled. In real-life, the repulsive force exhibits different patterns on three different circumstances. While two agents are moving along similar direction, the agent in the front should not be affected by the repulsive force from the one behind. While two agents are moving along opposite direction and about to collide, they should be both affected by the repulsive force. While two agents are moving along opposite direction but back to back, both of them should not be affected by the repulsive force. According to these details, the repulsive force f_{ji} could be further constrained using (8).

$$f'_{ji} = \begin{cases} \theta((v_j - v_i)n_{ij})G_{ij}f_{ji} & d_{ij} > \rho_i + r_j \\ (1 + \theta((v_j - v_i)n_{ij})G_{ij})f_{ji} & d_{ij} \leq \rho_i + r_j \end{cases} \qquad (8)$$

where $\theta(z)$ equals to zero if $z < 0$, and equals to z otherwise. The range of G_{ij} is from 0 to 1, which is used to impact the magnitude of repulsive force when two agents are moving along opposite directions.

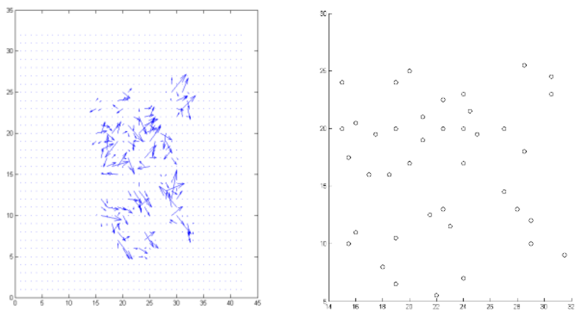
With the proposed equations, the Group Attraction Force f_{Gi} of agent i could be estimated. Since f_{Gi} is a vector, the co-ordinate of Grouping Center could be finally obtained using f_{Gi} and agent i's co-ordinate (x_i, y_i).

Figure 4. A snapshot of the simulated video

IV. EXPERIMENTAL SETTINGS AND RESULTS

The video data used for experiment is generated by 3D game engine Unity. The scene is composed with 40 agents, and each agent is replaced with a sphere for easier individual detection. Agents are assigned to two different groups, 20 agents for each. Agents from different groups are mapped with different textures. As illustrated in Fig. 4, agents from group 1 are mapped with dark green color and those from group 2 are mapped with bright white color. The two hexagons mark destinations of these two groups, agents from group 1 attempt to move toward to the upper destination and agents from group 2 to the lower one. Each agent in the stage is mapped with a behavioral model consists of the desired force, repulsive force and group attraction force strictly defined using previously proposed formulations.

Fig. 4 illustrates the randomly distribution of two groups. Agents from both groups are mixed up. It's very difficult to segment the crowd in current state using spatial information. According to the architecture of the proposed segmentation procedure in Fig. 3, the first step is to extract the optical flow using the Horn Schunk algorithm. In the experiment, frames of number 20 and 22 are selected for the optical flow extraction. The reason of using the early frames is that agents at this time are still in randomly distributed state. Since the distribution of agents would quickly become clustered under the impact of the group attraction force. The reason of skipping one frame instead of using two consecutive ones is to ensure the magnitude of extracted optical flow is large enough. Because if the magnitude is too small, the optical flow field would be easily affected by factors such as deviation and noise. Fig. 5(a) illustrates the extracted HS optical flow field from Fig. 4. In this experiment, the grid size is set to 10 pixels. Since the ultimate goal of this research is to achieve real-time detection, the overly condensed sampling grid would have the chance to affect the performance of the system. It is necessary to declare that the magnitude of the optical flow is amplified by 3 times to provide a better visual experience. However, the actual motion of each agent is not as drastic as the figure shows.

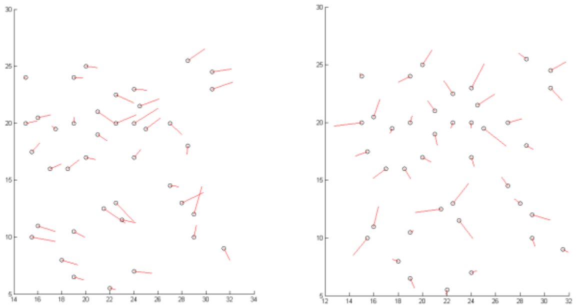


(a) (b)

Figure 5. The extracted optical flow and detect agents. (a) The extracted optical flow field using frame 20 and 22. (b) The spatial distribution of detected agents

According to the framework proposed in Fig. 3, agents in the scene need to be correctly detected. In most of the case, conventional people detectors could satisfy the requirement unless the crowd density is too high. Since the agent is represented with a sphere, shape detection algorithm such as Hough Circles could be more reliable. The agent detection result is illustrated in Fig. 5(b).

Next step of the procedure is to map the detected agents with extracted optical flow flied to obtain the actual social force. According to the equation (5) proposed in previous section, the average of all nine neighboring optical flow vectors is exploited as the final optical flow for current agent, and the parameter K is still set to 0.4. The obtained actual social forces are illustrated in Fig. 6(a). By comparing the detected agent's motion with the group truth, most behaviors of the agent are basically matched, despite with some deviations.



(a) (b)

Figure 6. (a) The mapping result of actual social force. (b) The estimated repulsive force

In next step, the repulsive force of each agent is calculated according to (7)(8) and (9). The estimation result of the repulsive force is illustrated in Fig. 8. In this case, the value of personal space ρ is set to 5 pixels, radius of agent r is set to 0.5 pixel. The value of parameter A_i and B_i should be 2 and 0.5. In Fig. 6(b), the value of B_i is set as 1 for a better visual presentation, the actual calculation would still be using 0.5.

Fig. 7 illustrated the estimation result of desired force of each agent. In the experiment, it is assumed the two destinations are successfully detected using Points of Interest detection techniques. Therefore, the only estimation in this step is to determine which destination the agent belongs to using (4). In the experiment, the magnitude of the desired force is set to the average amount of actual force affected from agents.

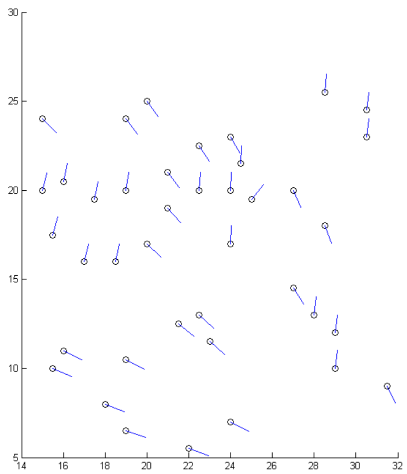


Figure 7. The estimated desired force

Once the actual force, repulsive force and desired force are extracted, the final Group Attraction Force could be estimated with (3). According to the definition of Group Attraction Force, the Grouping Center of each agent could be calculated. Next, the ordinary classifier is applied to the grouping center for the clustering. The clustering result is illustrated in Fig. 8. Agents from different groups are marked with different labels. Comparing the ground truth illustrated in Fig. 4, most of agents are correctly segmented. The accuracy is 87.5% in this example.

 ISBN: 978-1-61208-677-4

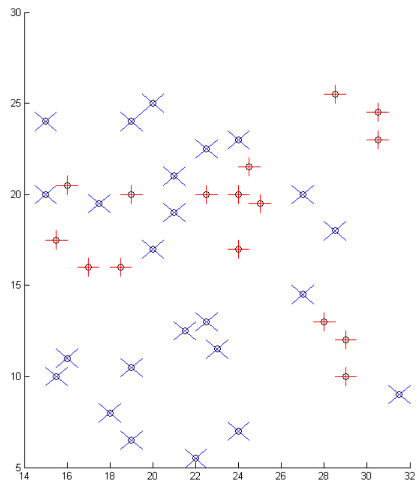


Figure 8. The segmentation result using the proposed approach

V. CONCLUSION AND FUTURE WORKS

In order to achieve effective crowd segmentation under complex sub-grouping states, an innovative Social Force Model is devised in this research. It is based on the assumption that crowd agents from the same behavioral/action group, i.e. a family unit, would make due effort to stay close within a crowd. A novel interaction force, named Group Attraction Force, is devised to populate the SFM feature set for accurately describing crowds formed by sub-groups. The concept of Grouping Center can effectively aid crowd segmentation even when agents from different groups are heavily mixed up.

Unlike conventional crowd segmentation techniques that often rely on spatial and temporal patterns only, the proposed SFM approach achieves superior results with improved semantic interpretation. Random crowd scenarios generated by a commercial game engine are used for crowd behavior analysis and model evaluation. Experiments show accurate segmentation can be achieved even when agents from different groups are heavily (spatially) mixed up. Future work will see test and evaluations being performed on real-life video feeds when factors such as occlusions and illumination variations need to be considered.

TABLE I. PARAMETER SETTING

Parameter Name	Parameter Value
weight factor k	0.4
Relax parameter G_{ij}	0.5
Frames selection	20 and 22
Grid size in optical flow	10 pixels
personal space ρ	5 pixels
radius of agent r	0.5 pixel
A_i	2
B_i	0.5

ACKNOWLEDGMENT

This research is sponsored by Chinese National Natural Science Foundation No. 61671377, and Shaanxi Smart City Technology Project of Xianyang No.2017k01-25-5

REFERENCES

[1] L. Dong, V. Parameswaran, V. Ramesh, and I. Zoghlami. "Fast Crowd Segmentation Using Shape Indexing". In 2007 IEEE 11th International Conference on Computer Vision. pp. 14-21. 2007

[2] N. Dalal and B. Triggs. "Histograms of Oriented Gradients for Human Detection". In International Conference on Computer Vision Pattern Recognition, pp. 886–893. 2005.

[3] K. Kang and X. Wang. "Fully Convolutional Neural Networks for Crowd Segmentation". In arXiv:1411.4464v1 [cs.CV]. 2014

[4] N. Dalal and B. Triggs. "Histograms of Oriented Gradients for Human Detection". In International Conference on Computer Vision Pattern Recognition 2005, pp. 886–893. 2005.

[5] P. Sinha. "Object Recognition via Image Invariants: A Case Study". In Investigative Ophthalmology and Visual Science. Volume 35, pp. 1735-1740. 1994.

[6] P. Tu, T. Sebastian, G. Doretto, N. Krahnstoever, J. Rittscher, and T. Yu. "Unified Crowd Segmentation". In Computer Vision – ECCV 2008. Lecture Notes in Computer Science, vol. 5305. Springer, Berlin, Heidelberg. 2008

[7] B. Leibe, E. Seemann, and B. Schiele. "Pedestrian Detection in Crowded Scenes". In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (2005). San Diego, California, pp. 878-885, June 2005

[8] D. Lowe. Distinctive image features from scale-invariant keypoints. IJCV, 60(2). pp. 91–110, 2004.

[9] B. Leibe, A. Leonardis, and B. Schiele. Combined object categorization and segmentation with an implicit shape model. In ECCV'04 Workshop on Stat. Learn. in Comp. Vis., pp. 17–32, 2004

[10] K. Tanaka, "Detecting collision-free paths by observing walking people," in Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne , Switzerland, pp. 55-60, 2002.

[11] S. Tadokoro, M. Hayashi, Y. Manabe, Y. Nakami, and T. Takamori, "Motion planner of mobile robots which avoid moving human obstacles on the basis of stochastic prediction," in IEEE International Conference on Systems, Man and Cybernetics, pp. 3286-3291, 1995.

[12] S. Gaffney and P. Smyth, "Trajectory clustering with mixtures of regression models", in 99 Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, University of Californa, Irvine. pp. 99-15, 1999.

[13] M. Bennewitz, W. Burgard, and S. Thrun, "Learning motion patterns of persons for mobile service robots," in Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3601-3606, 2002.

[14] Y. Hao, Zhijie Xu, Ying Liu, Jing Wang, and Jiulun Fan. "Crowd Synthesis Based on Hybrid Simulation Rules for Complex Behaviour Analysis". In Proceedings of the 24th International Conference on Automation & Computing, in press.

[15] H. Dirk, and P. Molnar. "Social force model for pedestrian dynamics", Physical review, E 51, no. 5, 4282. 1995.

[16] Q. Ji, F. Wang, and T. Zhu, "VPBS: A Velocity-Perception-Based SFM Approach for Crowd Simulation", In 2016 International Conference on Virtual Reality and Visualization. Hangzhou, China. 2016

[17] B. Horn, and B. Schunck. "Determining optical flow". Artificial Intelligence, vol 17, pp. 185–203. 1981