

Novel Group Detection and Analysis Method Based on Automatic and Fast Density Clustering

Ke Jin, WeiWei Xing
School of Software Engineering
Beijing Jiaotong University
Beijing 100044, China
www.xing@bjtu.edu.cn

Peng Bao
School of Software Engineering
Beijing Jiaotong University
Beijing 100044, China

Abstract—Group-level crowd behavior analysis is a new and promising method with important applications for the video surveillance and understanding of crowds. However, a specific definition for a group in a crowd field has rarely been investigated. This paper proposes a complete and reasonable definition for a group in a crowd field and presents a fast and automatic group detection method. First, automatic and fast density clustering (AFDC) is used to find the group core, which is then refined based on the property of coherent neighbor invariance. This detection method is more adaptive to groups with arbitrary shapes and varying densities because the group core is refined with coherent neighbors. Experiments on hundreds of video clips of public scenes showed that the method achieved an excellent detection performance and attractive statistical results. In particular, the number of people in a group exhibits a power-law distribution truncated by an exponential tail; this is significant to understanding crowd scenes and crowd simulation.

Keywords—group detection; group level; crowd behavior analysis; density clustering; power-law distribution

I. INTRODUCTION

A crowd is a large number of persons gathered closely together. The behavior analysis of crowds has important applications in public areas of high security interest. Because groups are the primary entities that make up a crowd and impact crowd dynamics [1, 2], group-level behavior analysis is a very important method for characterizing and comparing the dynamics of different crowded scenes. Understanding group dynamics is scientifically important and practically useful for crowd behavior analysis [3, 4]. However, a specific definition for a group in a crowd has rarely been given, even though groups have been extensively studied in the socio-psychological and biological fields [5–7].

There is a rich body of literature [1, 8–10] suggesting that a crowd generally maintains the important property of collectiveness, i.e., the majority of pedestrians tends to move in groups with friends and family members. The state of a group can be greatly different: it may be stationary at some point but become mobile following crowd dynamics with diverse densities and varying shapes other times. The collective phenomenon is more prominent in dense crowds [2, 11]. As a set of members with a common goal and collective behavior, a

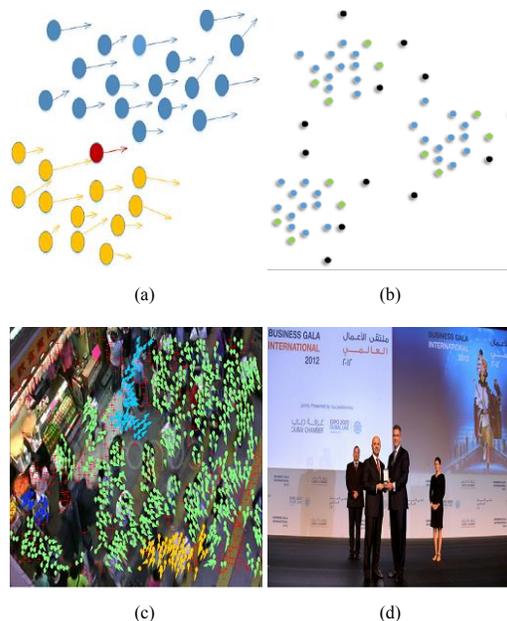


Figure 1. Examples of overlapping and hierarchical properties of a group. (a) The red individual was originally part of the blue group and now belongs to the orange group. (b) The group can be viewed as hierarchical: the stable individuals comprise the group core (blue), which collects other coherent individuals (green) into a group. The remaining individuals not belongs to any group are outliers (black). (c) The collectiveness of a group in the real world; the two groups in opposite directions merge into a bigger group. (d) The important people (person giving the award and person receiving it) may be part of the group core (all four people).

group often exhibits two other properties in the real world: overlapping and hierarchy. These can be abstracted from the common phenomenon where some individuals of a group move towards the same goal while being surrounded by other coherent individuals. This is important when analyzing crowd behavior at the group level. Fig. 1 shows some examples. Overlapping means that some individuals belonging to one group may be part of another group at another time, especially when the group

splits or merges in extremely crowded situations. Depending on the level of importance and correlation with the group center, individuals in a group can be divided into two levels: the group core and coherent individuals. Thus, groups can be viewed as hierarchical: stable individuals comprise the group core, which collects other coherent individuals into a group.

To better understand these properties of a group in a crowd, the problem of detecting a group must be solved first. Many methods have been proposed to detect groups [1, 12–14]. However, these methods cannot be applied to complex patterns. Later works [6, 15] have developed models by abstracting the spatiotemporal coherent relationships to provide more satisfying results. These emphasized finding the group without accounting for the hierarchical attributes.

Group detection can be a complex problem in crowded scenes [16], especially when considering groups with arbitrary shapes and varying densities. However, hierarchical clustering (HC) [8] and the density peaks method [17] cluster tracklets under the assumption that the cluster center is surrounded by neighbors at a lower density or by using velocity and spatial constraints. It is difficult for these methods to handle crowd problems with varying densities. Other methods [18, 19] that use a motion flow field are limited to specifying the clustering number manually. The most recent methods [17, 7] require a factor given manually, which needs field experience. Shuai et al. [2, 11] tried to analyze group attributes but could only handle the problem when the groups are stationary.

For application to video surveillance, another key challenge is efficiency. The diverse scenes from large amounts of video require a much more efficient way to quickly detect group dynamics. Approaches based on optimal flow [4, 13, 20, 21] consider large amounts of tracking points, and methods based on local coherent relationships [3, 6, 8, 15] need to maintain the topological structure and update it with iterative computations. The time cost for all of these is so high that they cannot be used in real-time applications.

The goal of our research was to profile a group in a crowd, give a specific visual definition, detect the group, study the statistical properties, and consider the impact on crowd dynamics. We considered a group to be a set of members with a common goal and collective behavior and having two common properties: overlapping and hierarchy. Groups are the primary entities that make up a crowd. Their dynamics are influenced by the crowd dynamics, and their impact on crowd cannot be ignored either. To the best of our knowledge, we are the first to try to comprehensively and systematically give a relatively specific definition based on previous research. The main contributions of our approach include the following:

- A novel automatic and fast density clustering algorithm is proposed. There are three salient properties must not be ignored: (a) recognizing outliers accurately, (b) automatically determining the group number and group center without any manual intervention, and (c) accurate, efficient, and easy real-time application.
- A new fast and robust group detection method is developed that is based on the Automatic and Fast Density Clustering (AFDC) algorithm and combined

with spatiotemporal neighbor invariance to characterize the underlying priors between neighbors. This detection method is more robust against complex scenes and it is also presented to uncover both local and global attributes. It can be effectively applied to groups with arbitrary shapes and varying densities.

- A short but comprehensive definition is presented for depicting groups in a crowd along with interesting statistical findings. Since it has attracted a great number of attention, there is no proper definition for a group. Based on a great deal of research, we limited a “group” to a specific meaning in a crowd field. This served as an inspiration for group understanding and crowd behavior simulation. Similar to collective phenomena in colonies [5], human groups also show interesting statistical results.

II. RELATED WORKS

For a long time, crowd analysis has been conducted at the macroscopic or microscopic level. At the macroscopic level, the main research focus has been the global motions of massive numbers of people; crowd behavior analysis did not consider the movements of individuals [4, 22–24]. At the microscopic level, the crowd is treated as a collection of individuals [25, 26], and analysis is based on the collective information of individuals. In contrast, some studies tried to analyze crowds at the group level [3, 8]. There are two main methods for recognizing groups in a crowd: shape classification and trajectory analysis. There are no fine methods for group detection by shape classification, while methods based on trajectory/tracklets greatly rely on object detection and tracking.

Coherent motion is an important phenomenon that has attracted a great deal of attention in a wide range of fields, from biology [5, 27] to physics [28], and cannot be ignored in video analysis. There are many similarities between group dynamics and coherent motion; both have the same goal and demonstrate collectiveness in a crowd.

State-of-the-art methods [3, 6, 7] have achieved group detection through tracklet analysis. Shao et al. [3] proposed the novel Collective Transition (CT) prior to capture the underlying dynamics of a group and devised a set of visual descriptors to quantify the universal properties of groups in a crowd. Wu et al. [7] developed the Collective Density Clustering (CDC) approach, and Zhou et al. [6] presented Coherent Filtering (CF) to recognize coherent motion. However, the above methods are unfeasible for real-time application or require manual intervention. More importantly, none of them analyze crowd behavior at the group level nor learn groups with both overlapping and hierarchical properties.

Other works [23, 29] focusing on scene-specific analysis may not be applied to groups with arbitrary shapes and varying densities. Our work is significantly different from the above research as follows: (i) We analyzed crowd behaviors at the group level and considered the overlapping and hierarchical properties when profiling a group, (ii) and our approach can handle groups both in crowded scenes and in low-density or disconnected areas.

III. GROUP DETECTION METHOD

We define a group as a collection of individuals who gather together and move towards the same goal. Examples in the real world have shown us that groups are usually overlapping and hierarchical. A group is generally a special part of crowd but goes beyond a simple collection of several individuals. The members of a group are usually spatially proximate or temporally coherent, and a group is a dynamic and independent unit that demonstrates various fundamental properties. Its activities can greatly influence the crowd in diverse modes.

In this paper, we propose a group detection method based on this definition. First, we detect the group core automatically by quickly finding the peak density of individuals. The accuracy and efficiency of our method guarantee its effectiveness for real-time application and make it robust against arbitrary shapes and varying densities in changing scenes. We then refine the group members with the coherent neighbor invariance prior. This reflects the overlapping property of the group and our approach achieves detection with both local and global properties. Groups vary in different scenes. Examples are shown in Fig. 2.

Motion features are detected with the gKLT feature point tracker [15]. Each group $\{G_i\}$ encompasses a set of tracklets $\{z\}$. For better performance, the expectation-maximization (EM) algorithm is used to fit the tracklets.

A. Automatic and Fast Density Clustering

The AFDC algorithm is the main part of the group detection method. First, we set the parameters for the algorithm, and we explain the differences between our algorithm and the normal one [17]. We then find the group center for each group and solve for the group members. Here, the group members comprise the group core.

1) Initial Parameters

In contrast to the normal density calculation method [17], our method can estimate both the position and orientation to solve the problem.

In two-dimensional space, consider a crowd $\{C\}$ possessing N points, where $\{C_i\}$ is the set of points and $\|C\| = N$. We consider every point C_i ($i = 1 \dots N$) as a vector with the spatial position (x_i, y_i) and orientation α_i . The orientation for each point at the certain frame can be calculated by neighbouring several frames. The similarities (including distance and orientation) between two points C_i and C_j are denoted as

$$d_{xy}(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (1)$$

$$\theta(i, j) = |\alpha_i - \alpha_j| \quad (2)$$

Now, we can calculate the density ρ_i as a Gaussian kernel function:

$$\rho_i = \sum_j \exp\left(-\frac{d_{xy}(i, j) * (\cos \theta(i, j) + 1)}{d_{cut}}\right) \quad (3)$$

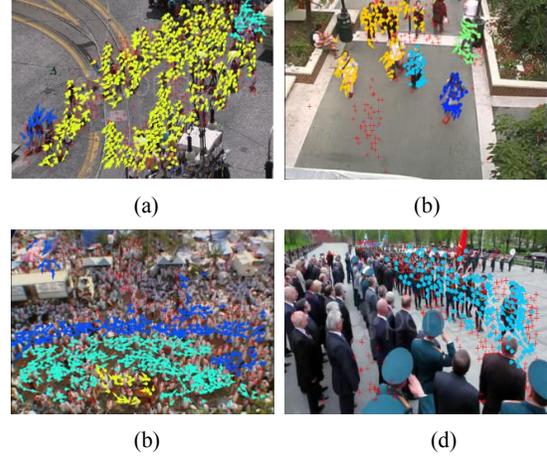


Figure 2. Groups in several representative crowd scenes. Scatters and arrows with different colors indicate different groups, and red crosses indicate the outliers. (a) Non-uniform group. (b) Crowd with low density. (c) Groups in an extremely crowded scene. (d) Marching soldiers and stationary nearby visitors.

The salient difference between this parameter and the primary one [17] is that the position and orientation can be synthetically estimated at the same time. As we know, $\cos \theta$ is a monotonic decreasing function from 1 to 0, where the independent variable θ changes from 0 to $\frac{\pi}{2}$. Thus, the orientation θ can adjust the error grouping for only the distance, but a small θ will be in the same group. In addition, δ_i represents the closest distance from other points with a higher density ρ_i . For the point C_i , δ_i is denoted as

$$\delta_i = \min_{j: \rho_j > \rho_i} d_{xy}(i, j) \quad (4)$$

For the point with the highest density, $\delta_i = \max d_{xy}(i, j)$.

Because the accuracy of ρ_i heavily depends on the threshold d_{cut} , we propose an automatic method to handle this inevitable problem by using entropy without depending on subjective experience [7, 17]. We were inspired by the Shannon entropy; in information theory, entropy is often used to quantify the uncertainty of a probability distribution when a specific event is observed. We calculate the entropy H for the point C_i ($i = 1 \dots N$) with the density δ_i as follows:

$$H = -\sum_i^N \frac{\rho_i}{W} \log \frac{\rho_i}{W} \quad (5)$$

where $W = \sum_i^N \rho_i$ is the normalization factor. We then establish the function H with the independent variable d_{cut} . We performed experiments to prove the effectiveness of this method.

2) Detecting the Group Core

Inspired by the new clustering method [17], the first step after the density ρ_i and distance δ_i are defined is determining the group center for each subgroup. A larger value of ρ_i indicates that the point C_i has more neighboring points, and δ_i is much

Algorithm 1 Automatic and Fast Density Clustering

Input: Number of points N , set of points $\{C_i\}$ ($i = 1 \dots N$)**Output:** Number of group M , the clustering result $\{Group_j \mid (j = 1 \dots M)\}$

```
1: FOR  $i = 1$  TO  $N$ 
2:   FOR  $j = i+1$  TO  $N$ 
3:      $d_{xy}(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ 
4:      $\theta(i, j) = |\alpha_i - \alpha_j|$ 
5:   END
6: END
7: FOR  $i = 1$  TO  $N$ 
8:    $\rho_i = \sum_j \exp\left(-\frac{d_{xy}(i, j) * (\cos \theta(i, j) + 1)}{d_{cut}}\right)$ 
9: END
10: FOR  $i = 1$  TO  $N$ 
11:   IF  $\rho_i$  is larger than any other points
12:      $\delta_i = \max d_{xy}(i, j)$ 
13:   ELSE  $\delta_i = \min_{j: \rho_j > \rho_i} d_{xy}(i, j)$ 
14: END
15: FOR  $i = 1$  TO  $N$ 
16:    $product_i = \rho_i * \delta_i$ 
17: END
18: SORT( $product_i$ , DESC)/sort the  $product_i$ *
19: FOR  $i = 2$  TO  $N$ 
20:    $gap_i = product_{i-1} - product_i$ 
21:   IF  $gap_i > TH$ 
22:      $Group_i = i-1$ 
23:   ELSE BREAK
24: END
25: FOR  $i = 1$  TO  $N$ 
26:   Assign point according its distance to group center.
27: END
```

Return: $\{Group_j \mid (j = 1 \dots M)\}$

larger than the nearest neighbor distance only when C_i has the maximum density. We determine the group centers subjectively according to the larger ρ_i and δ_i . In contrast to the typical method in [17], the initial group center is determined by $\max_i(\rho_i \times \delta_i)$. We select the value for which the gap between the neighboring two ($\rho_i \times \delta_i$) is larger than a threshold TH, which is far above the other gap. $Group_i = i$ is used to number the group centers. Other group members are assigned later in

descending order of ρ_i . Except for the initial group centers, other points are assigned as follows. For the points C_i and C_j with $\rho_i < \rho_j$, the point C_i is assigned to the nearest point C_j when $d_{xy} < \frac{d_{cut}}{\cos \theta(i, j) + 1}$. Our method has two features: (a) points should in the same group with the nearest neighbor of higher density, and (b) points in one group show no great difference in orientation. Thus, it can be used to determine the number of groups.

After we define the center and core of groups, we can also determine outliers, including those inside groups. Outliers often occur when the group contains too few points or the orientations of points in one group are inconsistent. The outliers are grouped into a single point. The method can correctly identify outliers.

Algorithm 1 presents the details of the AFDC algorithm.

B. Group Refinement with Coherent Neighbor Invariance

We considered the collectiveness of a group combined with the spatiotemporal relationship between individuals to refine the group core. Experiments on synthetic data have shown the existence of Coherent Neighbor Invariance, including both spatiotemporal and velocity relationships [6].

Based on the underlying prior in the dynamics of crowd, we tried to finding volatile individuals with coherent behavior mainly from observing the time series for group cores. For every point C_i ($i = 1 \dots N$) at time t , where the frame t contains N individuals, we denote the K nearest neighbors as N_t^i . From the time t to $t + d$, the invariant neighbor set $M_{t \rightarrow d}^i$ contains the invariant neighbors among the K nearest neighbor set. For a certain individual A in the same group as another individual B, A tends to keep in the KNN set of B [6] over time. We define the new group G as the union of $M_{t \rightarrow d}^i$ and the group core, where i indicates the group center individual.

IV. EXPERIMENTAL RESULTS

We evaluated our proposed group detection algorithm for diverse and varying crowd scenes in the real world, such as escalators, crossroads, marathons, and railway stations. Some fascinating group characteristics were observed in the statistical experiments, and the application to crowd simulation demonstrated the effectiveness of our method.

A. Experimental Setup

Evaluations were mainly conducted on the CUHK Crowd Dataset for group detection and statistical analysis. This dataset contains crowd videos with various densities and perspective scales, including 474 video clips from 215 scenes. Shao et al. [3] manually annotated tracklets from 300 video clips. Tracklets not belonging to any group were annotated as outliers.

B. Parameter Setup

We used the entropy method to automatically determine the only parameter of the AFDC algorithm. We confirmed the effectiveness of the entropy method by designing an experiment with the CUHK Crowd Dataset [3] to investigate how the density threshold influences the quality of the summaries. We tuned the density threshold by varying k from 1% to 5% of the

total N . We calculated d_{cut} with k and N from the experiment and chose the best parameter, as we did before. In general, a large number of experiments should be done to obtain the parameter. Fig. 4 indicates that our group detection method performed the best for specific scenes where the density threshold d_{cut} was set to 60 around $k = 2\%$ and started to drop significantly before 50 and after 80. This proved that our method can detect groups automatically. The density threshold d_{cut} is calculated automatically by choosing the independent variable d_{cut} of the function H when the dependent variable H mutates. Fig. 3 shows an example.

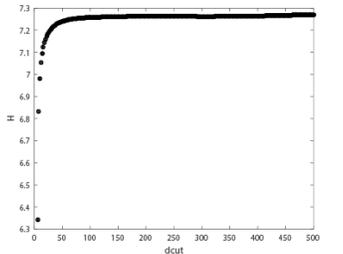


Figure 3. Function H with varying d_{cut} . We select the right d_{cut} at the inflection point.

C. Group Detection Performance

We compared our method with state-of-the-art group detection approaches using Collective Transition priors [3] and Coherent Filtering (CF) [6]. Owing to the expensive time cost of handling video clips of various lengths, we only took the first 30 frames from each clip to analyze. Experiments have shown

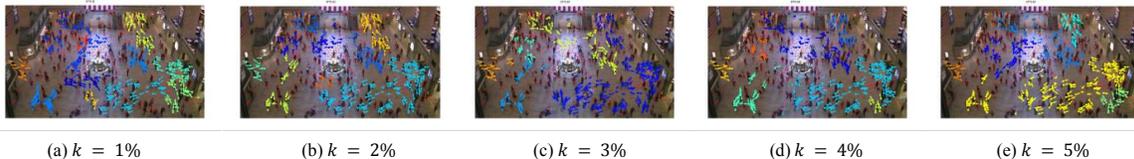


Figure 4. Crowd scene: Station. k changes from 1% to 5% of total N

that considering longer frames does not make a significant difference [3]. The ground truth was based on the CUHK dataset [3]. Fig. 5 compares examples of the detected results and ground truth.

CF is sensitive to tracking errors because it detects coherent motions just by considering neighbor relationships, not the whole dynamics of the whole group. Another reason is that CF first detects coherent motions among neighbors in several frames that it associates with the group. Therefore, its errors accumulate. CF works well in crowded scenes but not in situations when the group merges or splits. CT is a state-of-the-art group detection method with novel universal properties for groups in crowds. It works well in both sparse and crowded scenes. However, it cannot identify outliers well, especially those within groups. Because of the preprocessing method needing to maintain a neighbor set and the later iterative process, CT is extremely time-consuming and inapplicable to scenes in real-time. Table I compares the efficiencies of CT and our method in detail. We chose five representative scenes with different frames, sizes, and tracking points. Both CT and our method are based on the gKLT [30] tracking strategy. The efficiency is related to the number of tracking points. The experiments showed that our method works more than ten times faster than CT and can easily be applied to real time scenes.

To some extent, the group detection method can be seen as a clustering process, which is convenient for quantitative evaluation. We adopted three widely used measurements for evaluation: Normalized Mutual Information (NMI) [31], Purity [32], and the Rand Index (RI) [33]. Table II compares the values. Our method and CT performed at the same level, but our method was much more efficient than CT.

TABLE I. EFFICIENCY COMPARISON WITH CT AND OUR DETECTION METHOD

Methods				CT	Ours
	Frames	Tracking points	Size	Time per frame	Time per frame
Scene 1	184	1500	856×480	20.713 s	1.189 s
Scene 2	152	1000	856×480	13.564 s	0.649 s
Scene 3	201	1000	1920×1080	37.687 s	1.463 s
Scene 4	1001	500	960×540	31.498 s	0.501 s
Scene 5	123	1500	1920×1080	39.613 s	1.764 s

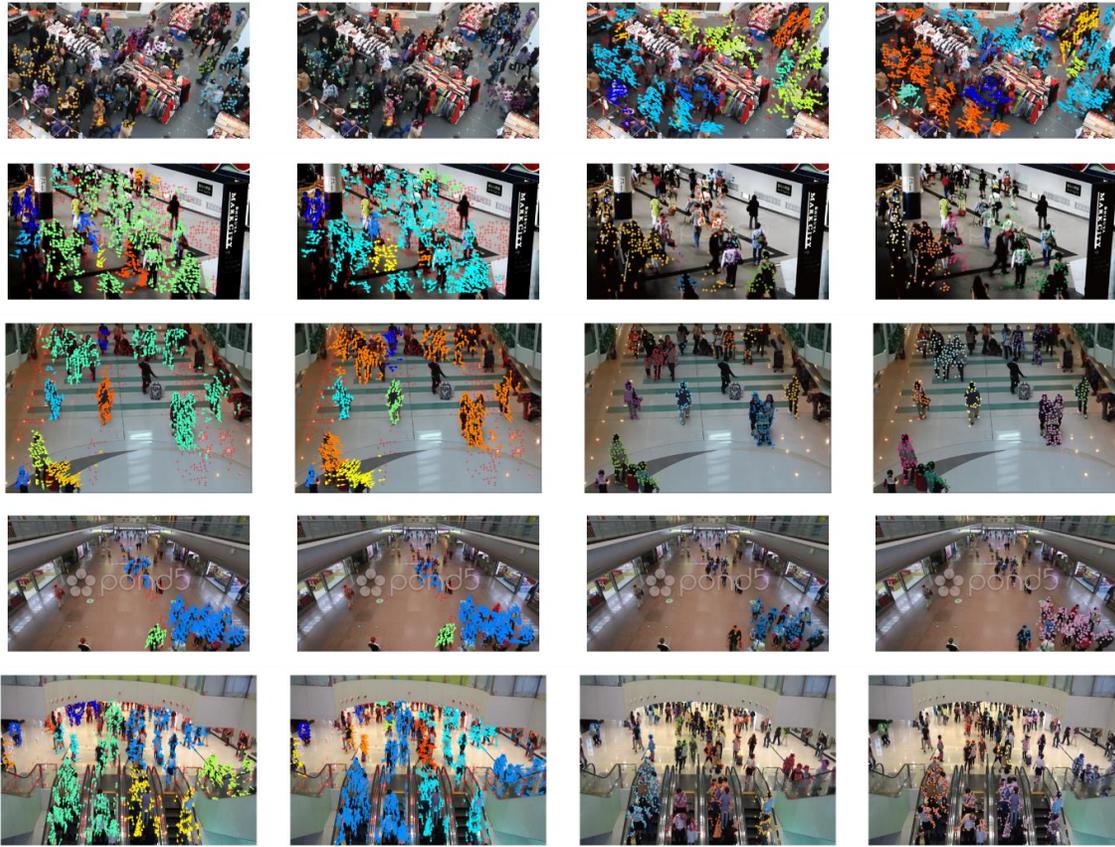


Figure 5. Comparative results of group detection with three methods in five diverse crowd scenes. Groups are distinguished with colors. Arrows indicate moving directions, and red crosses indicate the outliers. Best viewed in color.

D. Group Statistical Characteristics

Inspired by the collective motion of bacterial colonies [5], we observed some interesting results that will play an important role in crowd simulation and crowd scene understanding. Our statistics for a large number of groups showed that the number of individuals in a group exhibits a power-law distribution truncated by an exponential tail. The statistical properties become more obvious with enough groups. This can be a great guide for crowd simulation and understanding scenes.

With regard to the degree of the power-law distribution, this is a common phenomenon, especially for human beings. This is a distinctive property of complex networks and is of interest in activities involving human beings. Fig. 6 shows an example.

TABLE II. COMPARISON OF GROUP DETECTION METHODS

Methods	NMI	Purity	RI
CF [6]	0.43	0.72	0.76
CT [3]	0.49	0.78	0.82
HC [8]	0.27	0.62	0.73
Ours	0.52	0.83	0.86

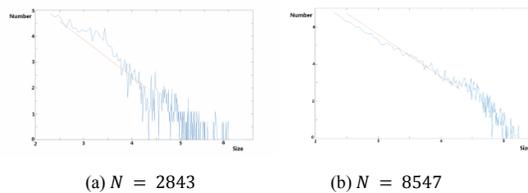


Figure 6. The abscissa is the group size, and the ordinate is the number of groups. (a) The total group number N is 2843. (b) The total group number N is 8547. The red straight line fits the power-law distribution, and an exponential tail can be observed.

V. CONCLUSION

In this paper, we systematically analyzed group behavior. We used the AFDC algorithm to detect groups in a crowd. We obtained important and interesting attributes intrinsic to a group through a statistical approach. These findings are vital to group

state analysis, crowd simulation, and crowd motion understanding. The advantage of our approach is that it can work with human intervention and provides excellent performance in terms of both accuracy and efficiency. The much higher efficiency of our method due to algorithm optimization and more efficient programming language means that it can be applied in real time to extremely complex scenes.

This work may open an interesting field. Further research is needed, such as video surveillance to define and detect group events and characterizing people to divide a group into three level. In the future, the following improvements will be made to the method: developing a more robust and accurate multi-object detection algorithm for tracklets, developing a better preprocessing algorithm for the tracklets, and reducing the repetitive computation for a faster AFDC algorithm

ACKNOWLEDGMENT

This work is supported in part by National Natural Science Foundation of China (No. 61100143, 61272353, 61370128), Program for New Century Excellent Talents in University (NCET-13-0659), Beijing Higher Education Young Elite Teacher Project (YETP0583), and Fundamental Research Funds for the Central Universities (2014JBZ004, 2015RC031).

REFERENCES

- [1] M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz, "The walking behaviour of pedestrian social groups and its impact on crowd dynamics," *PLoS One*, vol. 5, no. 4, e10047, 2010.
- [2] S. Yi and X. Wang, "Profiling stationary crowd groups," in *IEEE Int. Conf. Multimedia and Expo*, Chengdu, China, Jul. 14-18, 2014, pp. 1-6.
- [3] J. Shao, C. C. Loy, and X. Wang, "Scene-independent group profiling in crowd," in *IEEE Conf. Computer Vision and Pattern Recognition*, Columbus, OH, Jun. 23-28, 2014, pp. 2227-2234.
- [4] S. Ali and M. Shah, "A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in *IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Minneapolis, MN, Jun. 18-23, 2007, pp. 1-6.
- [5] H. P. Zhang, A. Beer, E.-L. Florin, and H. L. Swinney, "Collective motion and density fluctuations in bacterial colonies," *Proc. Natl. Acad. Sci.*, Vol. 107, No. 31, pp. 13626-13630, 2010.
- [6] B. Zhou, X. Tang, and X. Wang, "Coherent filtering: Detecting coherent motions from crowd clutters," in *Proc. 12th European Conf. Computer Vision*, Florence, Italy, Oct. 7-13, 2012, part II, pp. 857-871.
- [7] Y. Wu, Y. Ye, and C. Zhao, "Coherent motion detection with collective density clustering," in *Proc. 23rd Annu. ACM Conf. Multimedia*, Brisbane, Australia, Oct. 26-30, 2015, pp. 361-370.
- [8] W. Ge, R. T. Collins, and B. Ruback, "Vision-based analysis of small groups in pedestrian crowds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 5, pp. 1003-1016, 2012.
- [9] A. F. Aveni, "The not-so-lonely crowd: Friendship groups in collective behavior," *Sociometry*, vol. 40, no. 1, pp. 96-99, 1977.
- [10] F. Solera, S. Calderara, and R. Cucchiara, "Structured learning for detection of social groups in crowd," in *10th IEEE Int. Conf. Advanced Video and Signal-based Surveillance*, Krakow, Poland, August 27-30, 2013, pp. 7-12.
- [11] S. Yi, X. Wang, C. Lu, and J. Jia, "L0 regularized stationary time estimation for crowd group analysis," in *2014 IEEE Conf. Computer Vision and Pattern Recognition*, Columbus, OH, Jun. 23-28, 2014, pp. 2219-2226.
- [12] T. Lan, Y. Wang, W. Yang, S. N. Robinovitch, and G. Mori, "Discriminative latent models for recognizing contextual group activities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1549-1562, 2012.
- [13] L. Kratz and K. Nishino, "Going with the flow: Pedestrian efficiency in crowded scenes," in *Proc. 12th European Conf. Computer Vision*, Florence, Italy, Oct. 7-13, 2012, part IV, pp. 558-572.
- [14] C. Zhang, X. Yang, J. Zhu, and W. Lin, "Parsing collective behaviors by hierarchical model with varying structure," in *Proc. 20th ACM Multimedia Conf.*, Nara, Japan, Oct. 29-Nov. 2, 2012, pp. 1085-1088.
- [15] B. Zhou, X. Tang, H. Zhang, and X. Wang, "Measuring crowd collectiveness," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 8, 1586-1599, 2014.
- [16] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, and S. Yan, "Crowded scene analysis: A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 367-386, 2015.
- [17] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492-1496, 2014.
- [18] T. Brox and J. Malik, "Object segmentation by long term analysis of point trajectories," in *Proc. 11th European Conf. Computer Vision*, Heraklion, Greece, Sep. 5-11, 2010, part V, pp. 282-295.
- [19] M. Hu, S. Ali, and M. Shah, "Learning motion patterns in crowded scenes using motion flow field," in *19th Int. Conf. Pattern Recognition*, Tampa, FL, Dec. 8-11, 2008, pp. 1-5.
- [20] R. Mehran, B. E. Moore, and M. Shah, "A streakline representation of flow in crowded scenes," in *Proc. 11th European Conf. Computer Vision*, Heraklion, Greece, Sep. 5-11, 2010, part III, pp. 439-452.
- [21] W. Wang, W. Lin, Y. Chen, J. Wu, J. Wang, and B. Sheng, "Finding coherent motions and semantic regions in crowd scenes: A diffusion and clustering approach," in *Proc. 13th European Conf. Computer Vision*, Zurich, Switzerland, Sep. 6-12, 2014, part I, pp. 756-771.
- [22] A. B. Chan and N. Vasconcelos, "Modeling, clustering, and segmenting video with mixtures of dynamic textures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 5, pp. 909-926, 2008.
- [23] D. Kuettel, M. Breitenstein, L. Van Gool and V. Ferrari, "What's going on? Discovering spatio-temporal dependencies in dynamic scenes," in *IEEE Conf. Computer Vision and Pattern Recognition*, San Francisco, CA, Jun. 13-18, 2010, pp. 1951-1958.
- [24] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Miami, FL, Jun. 20-25, 2009, pp. 935-942.
- [25] W. Choi and S. Savarese, "A unified framework for multi-target tracking and collective activity recognition," in *Proc. 12th European Conf. Computer Vision*, Florence, Italy, Oct. 7-13, 2012, part IV pp. 215-230.
- [26] C. C. Loy, T. Xiang and S. Gong, "Detecting and discriminating behavioural anomalies," *Pattern Recognit.*, vol. 44, no. 1, pp. 117-132, 2011.
- [27] I. Couzin, "Collective minds," *Nature*, vol. 445, no. 7129, pp. 715-715, 2007.
- [28] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Phys. Rev. Lett.*, vol. 75, no. 6, pp. 1226, 1995.
- [29] M.-C. Chang, N. Krahnstoever, and W. Ge, "Probabilistic group-level motion analysis and scenario recognition," in *IEEE Int. Conf. Computer Vision*, Barcelona, Spain, Nov. 6-13, 2011, pp. 747-754.
- [30] C. Tomasi and T. Kanade, "Detection and tracking of point features," Carnegie Mellon Univ., Tech. Rep. CMU-CS-91-132, Apr. 1991.
- [31] B. Schölkopf, J. Platt, and T. Hofmann, *A Local Learning Approach for Clustering*, Cambridge: MIT Press, 2007 pp. 1529-1536.
- [32] C. C. Aggarwal, "A human-computer interactive method for projected clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 4, pp. 448-460, 2004.
- [33] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *J. Am. Stat. Assoc.*, vol. 66, no. 336, pp. 846-850, 1971.