IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 27, 2019, accepted July 9, 2019, date of publication July 16, 2019, date of current version August 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2929242

Direction, Velocity, Merging Probabilities and Shape Descriptors for Crowd Behavior Analysis

RAKHSHENDA JAVID[1], M. MOHSIN RIAZ[2], ABDUL GHAFOOR[1], AND NAVEED IQBAL RAO[1]
[1]Military College of Signals, National University of Sciences and Technology (NUST), Rawalpindi 46000, Pakistan
[2]The Center for Advanced Studies in Telecommunication (CAST), COMSATS University, Islamabad 45550, Pakistan

Corresponding author: Rakhshenda Javid (rjavaid.phdcs@students.mcs.edu.pk)

ABSTRACT Descriptors are important for quantifying crowd behavior. The existing descriptors generally provide information about crowd density, i.e., number of people/objects present in a defined spatial area. However, other properties of crowd like speed, direction, shape, and merging probabilities (of different crowds at group level) are also important for crowd analysis. In this paper, crowd descriptors (by mitigating the effects of outliers) are introduced which can be used for crowds having various densities. The simulations on various datasets show the applicability of the proposed descriptors.

INDEX TERMS Descriptors, crowd behavior analysis, outliers.

I. INTRODUCTION

The most prominent characteristic of moving objects (people, animals/birds, insects, micro-organism and vehicles etc.) is their coherent motion. Moving objects are influenced by their neighboring moving objects. Some examples of people coherent motion include human crowds at stations, shopping malls, zebra crossings etc. The shared destinations and synchronized movements of individuals forms a general crowd behavior [1]. Crowd direction and density descriptor can locate the source and sinks of different crowd groups in a scene. The motion of different crowd groups in a scene helps in estimating their speed which in turn helps to prevent any disastrous situation e.g. road accidents. The probability of one group merging with other group can help in crowd management and disaster control. Similarly to determine how uniform a crowd is, the shape of the crowd can help. Quantitative analysis of coherently moving groups is useful in different applications (like crowd scene understanding, video classification, group segmentation, visual surveillance, homeland security, disaster prevention and behavior analysis [2]).

Some of the existing quantitative descriptors include *collectiveness* [3]–[7], *uniformity and conflict* [4], [5], *stability* [4], [5], [7], [8], *excitement level* [9], *direction and velocity* [10]–[12], *abnormal behaviors* [13]–[15], *Curl and Divergence* [16], density [17] (any italic word in this document represents a crowd descriptor). Competitive studies on crowd analysis can be found in [18], [19], [19], [21]. The crowd *collectiveness* descriptor is estimated using local spatio-temporal relationships, velocity correlations and collective transition priors [3]–[5]. In [4], [5], *stability, uniformity and conflict descriptors* using collective transition prior is also presented. Another *collectiveness* descriptor has been proposed by [6] using global spatio-temporal information. However the use of velocity correlations effects the descriptor's results for static crowds. In [7] an online collectiveness descriptor is defined based on a spatio-temporal feature detector and Gaussian mixture models. They also calculated stability, conflict and density of detected crowd. Lagrangian coherent structures [8] segments the coherent crowd flows and analyzed *flow instability* for videos. However the method is not capable to segment incoherent motion and has significant time delays.

In Conigliaro *et al.* [22] measure the *overall excitement* of a crowd by detection and classification (set of predefined classes) of different patches in the frame. Co-occurrence and spatio-temporal dependency based model uses background subtraction at pixel level for direct event characterization and *behavior analysis* [23]. However the techniques do not work for dynamic backgrounds. In [10] and [11], *direction* and *velocity* are estimated by tracking of individual trajectories in non crowded scenarios. Computer vision techniques are used to create realism by simulating crowd. However, the simulation results are generated from a synthetic (rather a real) scenario.

The associate editor coordinating the review of this manuscript and approving it for publication was Yong Wang.

(a) $t = 32$ (b) $t = 33$ (c) $t = 34$

(d) $t = 35$ (e) $t = 36$ (f) $t = 37$

(g) $t = 38$ (h) $t = 39$ (i) $t = 40$

FIGURE 1. Clustering results on public pathway dataset.

Crowd anomaly detection based on scale invariant feature tracking suffers from computational complexity [24]. *Normal* and *abnormal* crowd behavior classification based on bag of words and social force model suffers from localization anomalies [13]. Dynamic textures and appearance model estimates crowd *normality* by detecting manually labeled activities [14]. Local *anomalies* detection based on probabilistic technique fails to detect global patterns [15]. *Curl and Divergence Trajectories descriptors* (CDT) [16] uses trajectory-based motion coding to measure the collective motion patterns. A new texture feature based on completed local binary pattern is used to estimate the crowd density [17]. The image in divided into blocks and each block is classified as low, medium or high density based on a trained support vector machine.

Overall the existing techniques use velocity correlations, background subtraction or synthetic data, whereas we focus on analyzing the crowd at group level.

Our contributions are outlined as

- The development of a new and effective approach, which combines well-founded literature techniques in a novel way to quantize the crowd at group level.
- The group properties *speed, direction, shape* and *merging probabilities* have been quantified in this paper.
- The proposed descriptors can deal with outliers efficiently.

The aim of this paper is to quantify the group properties *speed, direction, shape* and *merging probabilities* in the crowd. The proposed descriptors are capable of detecting and excluding outliers. Note that the presence of outliers can cause significant deviation and errors in the results. Simulation using various datasets demonstrate the effectiveness/applicability of proposed descriptors.

II. PROPOSED CROWD DESCRIPTORS

The crowd analysis use tracklets obtained by applying KLT tracker [3]. Let $I^{(t)}$ be the t^{th} frame in the video, the KLT tracker is,

$$I^{(t)}\Big|_{t=1}^{T} \xrightarrow[\text{Tracker}]{\text{KLT}} \Big(X, Y\Big) \tag{1}$$

where the coordinate matrices X and Y with respect to time ($t = 1, 2, \ldots, T$) and tracklet number ($p = 1, 2, \ldots, P$)

TABLE 1. Public pathway dataset.

Time	Cluster	Speed (Outlier)		Direction	Shape			Probabilities			
		Without	With		Circle	Oval	Ellipse	C1	C2	C3	C4
32	C1	2.07	2.07	-88.59	0.04	0.0152	0.0479	1	0.45	0.43	0.45
	C2	3.74	3.74	-104.15	0.005	0.002	0.0045	0.45	1	0.43	0.43
	C3	4.00	4.00	-105.72	0.005	0.0015	0.0066	0.43	0.43	1	0.44
	C4	1.62	1.62	81.95	0.019	0.0053	0.0166	0.45	0.43	0.44	1
33	C1	2.64	2.07	-84.53	0.04	0.0156	0.0484	1	0.45	0.43	0.45
	C2	3.52	3.74	-92.73	0.006	0.0031	0.0053	0.45	1	0.43	0.43
	C3	4.16	4.00	-104.07	0.005	0.0061	0.0068	0.43	0.43	1	0.44
	C4	1.48	1.48	51.89	0.019	0.0050	0.0160	0.45	0.43	0.44	1
34	C1	2.64	2.64	-83.50	0.04	0.0150	0.0492	1	0.45	0.43	0.45
	C2	4.06	3.52	-93.67	0.006	0.0032	0.0054	0.45	1	0.43	0.43
	C3	3.35	4.16	-105.76	0.003	0.0015	0.0063	0.43	0.43	1	0.44
	C4	1.62	1.62	66.32	0.018	0.0050	0.0153	0.45	0.43	0.44	1
35	C1	2.53	2.64	-83.82	0.03	0.0134	0.0480	1	0.45	0.43	0.45
	C2	4.19	4.06	-93.12	0.006	0.0034	0.0055	0.45	1	0.43	0.43
	C3	4.44	3.35	-106.42	0.007	0.0025	0.0077	0.43	0.43	1	0.44
	C4	1.33	1.33	80.44	0.017	0.0051	0.0148	0.45	0.43	0.44	1
36	C1	1.81	1.81	-86.18	0.0375	0.0124	0.0475	1	0.45	0.43	0.45
	C2	3.65	4.19	-89.16	0.006	0.0035	0.0055	0.45	1	0.43	0.43
	C3	4.22	4.44	-108.33	0.008	0.0029	0.0082	0.43	0.43	1	0.44
	C4	1.00	1.00	66.95	0.0174	0.0050	0.0144	0.45	0.43	0.44	1
37	C1	1.20	1.20	-91.92	0.032	0.0135	0.0426	1	0.45	0.43	0.45
	C2	3.66	3.65	-89.60	0.001	0.0011	0.0034	0.45	1	0.43	0.43
	C3	3.83	4.22	-105.62	0.009	0.0031	0.0085	0.43	0.43	1	0.44
	C4	1.09	1.09	68.54	0.016	0.0047	0.0141	0.45	0.43	0.44	1
38	C1	4.95	1.20	-91.00	0.060	0.0148	0.0698	1	0.58	0.60	0
	C2	58.68	3.66	-105.01	0.0097	0.0033	0.0086	0.58	1	0.59	0
	C3	42.79	1.30	68.58	0..016	0.0046	0.0135	0.60	0.59	1	0
	C4	0	0	0	0	0	0	0	0	0	0
39	C1	4.75	4.95	-89.52	0.05	0.0155	0.0692	1	0.58	0.60	0
	C2	58.49	3.78	-109.62	0.009	0.0033	0.0086	0.58	1	0.59	0
	C3	43.11	1.05	67.18	0.016	0.0045	0.0136	0.60	0.59	1	0
	C4	0	0	0	0	0	0	0	0	0	0
40	C1	5.19	4.75	-91.04	0.057	0.0161	0.0674	1	0.58	0.60	0
	C2	58.35	3.32	-104.04	0.006	0.0032	0.0072	0.58	1	0.59	0
	C3	44.21	1.36	82.29	0.015	0.0044	0.0128	0.60	0.59	1	0
	C4	0	0	0	0	0	0	0	0	0	0

are [3],

$$X = \begin{bmatrix} x_1(1) & x_1(2) & \dots & x_1(T) \\ x_2(1) & x_2(2) & \dots & x_2(T) \\ \vdots & \vdots & \ddots & \vdots \\ x_P(1) & x_P(2) & \dots & x_P(T) \end{bmatrix} \quad (2)$$

and,

$$Y = \begin{bmatrix} y_1(1) & y_1(2) & \dots & y_1(T) \\ y_2(1) & y_2(2) & \dots & y_2(T) \\ \vdots & \vdots & \ddots & \vdots \\ y_P(1) & y_P(2) & \dots & y_P(T) \end{bmatrix} \quad (3)$$

These tracklets are used to find initial clusters using the coherent filtering [25] scheme. which are then refined using group detection scheme [4] to obtain final cluster matrix. In coherent filtering the local spatiotemporal relationships of individuals in coherent motion is used to detect coherent motion patterns. Accurate crowd clustering is performed using voting based technique (which takes into account the clustering information of previous frames). Let the cluster matrix C be defined as,

$$C = \begin{bmatrix} c_1(1) & c_1(2) & \dots & c_1(T) \\ c_2(1) & c_2(2) & \dots & c_2(T) \\ \vdots & \vdots & \ddots & \vdots \\ c_P(1) & c_P(2) & \dots & c_P(T) \end{bmatrix}$$

where $c_p(t)_{\forall(p,t)} \in [1, N]$ and N represents total number of clusters.

which are then refined using group detection scheme [4] to obtain final cluster matrix. In coherent filtering the local spatiotemporal relationships of individuals in coherent motion is used to detect coherent motion patterns.

A. CROWD DIRECTION

The crowd in motion has some *direction* and where the people are moving coherently in different directions. This descriptor

(a) $t = 11$ (b) $t = 12$ (c) $t = 13$

(d) $t = 14$ (e) $t = 15$ (f) $t = 16$

(g) $t = 17$ (h) $t = 18$ (i) $t = 19$

FIGURE 2. Clustering results on shopping mall dataset.

helps to identify the direction/flow of the crowd. Consequently, the sink/source and anomalies (i.e. motion in wrong direction) in the crowd can be detected. The research closest to our work is proposed in [12] but it is not directly comparable to our work since they focus on the crowd as a whole and finding the dominant motions in the crowd where as we focus on the direction of different crowd groups in a crowded scene.

To find the *direction* of n^{th} group in a crowded scene, the horizontal $x_u^{n_1}(t)$, $x_u^{n_1}(t-1)$and vertical $y_u^{n_1}(t)$, $y_u^{n_1}(t-1)$ coordinates of each point u, present at time t and $t-1$ are considered (to eliminate the effect of outliers).

The direction $\theta_u^n(t)$ for each point in cluster n at time t is then given as

$$\theta_u^n(t) = \tan^{-1}\left(\frac{y_u^n(t) - y_u^n(t-1)}{x_u^n(t) - x_u^n(t-1)}\right) \tag{4}$$

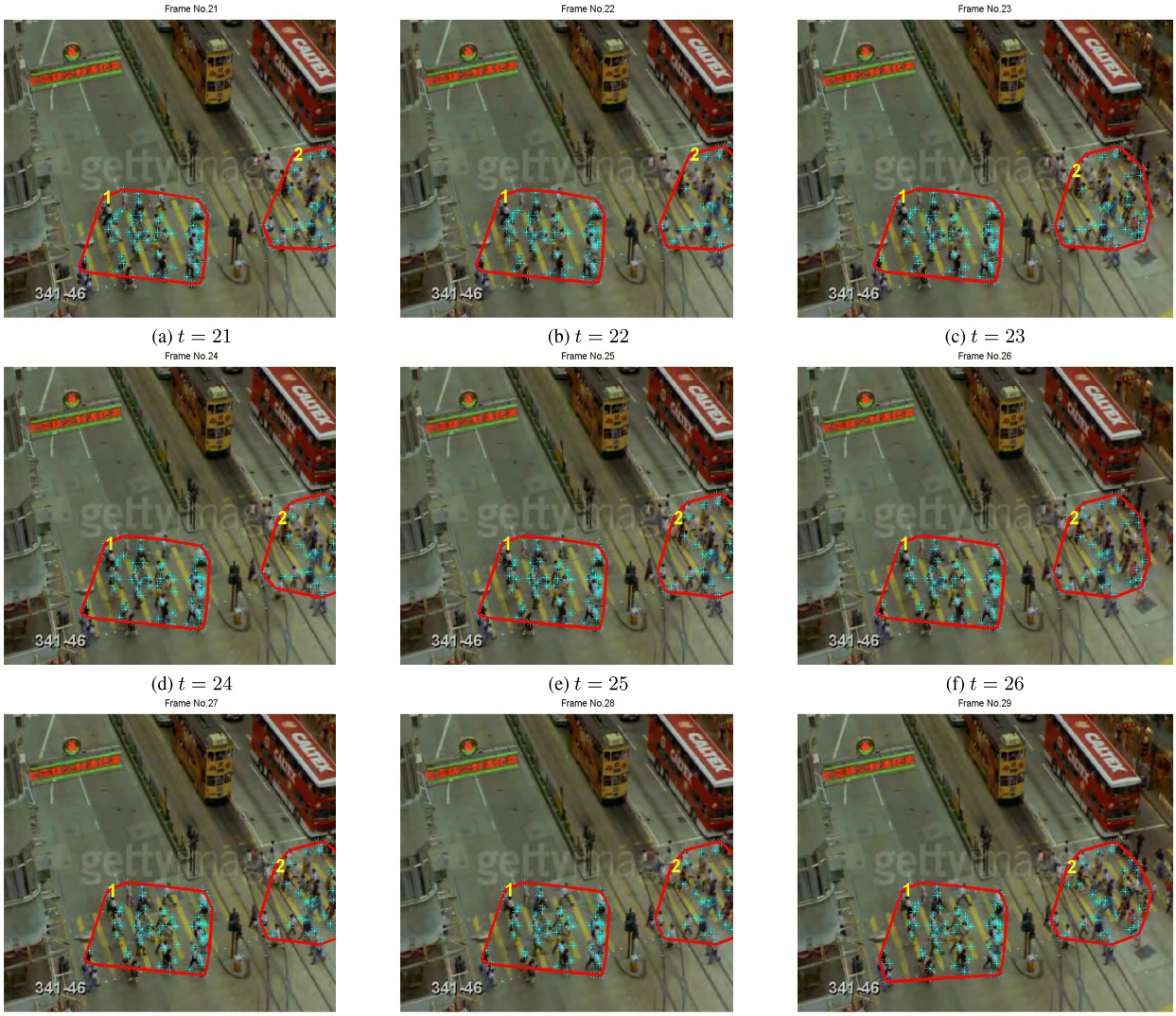
The final direction $\theta^n(t)$ is estimated by taking mean and mode of the directions computed for each point.

$$\theta^n(t) = \frac{\text{mean}\left(\theta_u^n(t)\right) + \text{mode}\left(\theta_u^n(t)\right)}{2} \tag{5}$$

B. CROWD SPEED/VELOCITY

The crowd moving in a direction will have some *speed/velocity* which provides the behavior of coherently moving objects. The *speed* descriptor is applicable in detection of traffic violations, pedestrian and bacterial abnormalities etc. The velocity of a crowd group at time t in a crowded scene is estimated by the distance covered in a specific amount of time δ i.e.,

$$D_n(t) = \frac{1}{U_n}\sum_{u=1}^{U_n}\sqrt{\left(x_u(t) - x_u(t-\delta)\right)^2 + \left(y_u(t) - y_u(t-\delta)\right)^2} \tag{6}$$



(a) $t = 21$ (b) $t = 22$ (c) $t = 23$

(d) $t = 24$ (e) $t = 25$ (f) $t = 26$

(g) $t = 27$ (h) $t = 28$ (i) $t = 29$

FIGURE 3. Clustering results on Zebra crossing dataset.

where U_n represents total number of points in the n^{th} cluster. The velocity $v_n(t)$ is defined as,

$$v_n(t) = \frac{D_n(t)}{\delta} \tag{7}$$

Note that the descriptors are computed over a complete group/cluster, therefore it is possible that some outliers can change the results drastically. To detect the outliers, the mean difference value of consecutive frames are compared with a pre-defined threshold. If any outlier is detected, the group is further divided into two sub-groups and the group with maximum points is selected for computation of different descriptors.

C. CROWD MERGING PROBABILITIES

The *merging probabilities* define how likely a cluster can merge with its neighboring clusters at specific time. The descriptor can be used to find any disastrous situations. A huge crowd or crowd accumulating at a specific area can be detected and managed before any disaster happens. To our knowledge no such descriptor has been described in previous studies. To calculate the merging probability, the distance between all the points of two crowd groups/clusters are computed. Let $\gamma_{n_1 n_2}(u, v, t)$ be the distance of u^{th} and v^{th} points present in n_1 and n_2 clusters respectively at time t such that,

$$\gamma_{n_1 n_2}(u, v, t) = \sqrt{\left(x_u^{n_1}(t) - x_v^{n_2}(t)\right)^2 + \left(y_u^{n_1}(t) - y_v^{n_2}(t)\right)^2} \tag{8}$$

$\forall(c_u(t) == n_1 \,\&\, c_v(t) == n_2)$. The distance matrix $\Gamma_{n_1 n_2}$ for cluster n_1 and n_2 is defined as,

$$\Gamma_{n_1 n_2}(t) = \begin{bmatrix} \gamma_{n_1 n_2}(1,1,t) & \gamma_{n_1 n_2}(1,2,t) & \dots & \gamma_{n_1 n_2}(1,V,t) \\ \gamma_{n_1 n_2}(2,1,t) & \gamma_{n_1 n_2}(2,2,t) & \dots & \gamma_{n_1 n_2}(1,V,t) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n_1 n_2}(U,1,t) & \gamma_{n_1 n_2}(U,2,t) & \dots & \gamma_{n_1 n_2}(U,V,t) \end{bmatrix}$$

TABLE 2. Shopping mall dataset.

Time	Cluster	Speed	Direction	Shape			Probabilities		
				Circle	Oval	Ellipse	C1	C2	C3
11	C1	1.11	45.8	0.0076	0.0039	0.0059	1	0.57	0.56
	C2	1.52	-124.9	0.10	0.0168	0.0836	0.57	1	0.61
	C3	0.54	38.61	0.013	0.0089	0.0212	0.56	0.61	1
12	C1	1.00	43.49	0.0077	0.0039	0.0058	1	0.57	0.56
	C2	2.32	-126.11	0.1015	0.0166	0.0838	0.57	1	0.60
	C3	0.81	44.09	0.0135	0.0089	0.0213	0.56	0.60	1
13	C1	1.07	44.12	0.0077	0.0038	0.0058	1	0.57	0.56
	C2	2.45	-123.55	0.0669	0.0233	0.0562	0.57	1	0.60
	C3	1.10	39.31	0.0132	0.0092	0.0209	0.56	0.60	1
14	C1	1.07	42.17	0.0078	0.0038	0.0057	1	0.57	0.56
	C2	2.56	-128.89	0.0672	0.0229	0.0563	0.57	1	0.60
	C3	1.36	40.32	0.0152	0.0108	0.0214	0.56	0.60	1
15	C1	1.52	43.92	0.0076	0.0040	0.0058	1	0.56	0.56
	C2	2.18	-124.15	0.0611	0.0204	0.0481	0.56	1	0.61
	C3	1.35	39.09	0.0152	0.0112	0.0220	0.56	0.61	1
16	C1	1.52	46.51	0.0077	0.0038	0.0058	1	0.57	0.56
	C2	1.44	-150.93	0.0775	0.0192	0.0589	0.57	1	0.61
	C3	1.32	38.50	0.153	0.0111	0.0229	0.56	0.61	1
17	C1	1.54	44.56	0.0077	0.0038	0.0058	1	0.56	0.57
	C2	1.41	-133.33	0.0642	0.0198	0.0511	0.56	1	0.61
	C3	1.62	37.58	0.0142	0.0109	0.0224	0.57	0.61	1
18	C1	1.53	44.51	0.0075	0.0041	0.0057	1	0.56	0.57
	C2	1.22	-136.48	0.0642	0.0199	0.0516	0.56	1	0.61
	C3	1.74	39.89	0.0136	0.0106	0.0222	0.57	0.61	1
19	C1	1.53	41.66	0.0075	0.0043	0.0059	1	0.56	0.57
	C2	1.11	-131.56	0.0616	0.0211	0.0492	0.56	1	0.61
	C3	1.13	39.69	0.0126	0.0097	0.0218	0.57	0.61	1

The mean distance is computed as,

$$\bar{\Gamma}_{n_1 n_2}(t) = \frac{1}{UV} \sum_{u=1}^{U} \sum_{v=1}^{V} \gamma_{n_1 n_2}(u, v, t) \tag{9}$$

and the merging probability between n_1 and n_2 cluster is,

$$\phi_{n_1 n_2}(t) = \frac{1 + e^{-\beta \bar{\Gamma}_{n_1 n_2}(t)}}{\alpha} \tag{10}$$

and for all clusters,

$$\Phi(t) = \begin{bmatrix} \phi_{11}(t) & \phi_{12}(t) & \dots & \phi_{1N}(t) \\ \phi_{21}(t) & \phi_{22}(t) & \dots & \phi_{2N}(t) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{N1}(t) & \phi_{N2}(t) & \dots & \phi_{NN}(t) \end{bmatrix}$$

Note that $\phi_{nn} = 1$ and $\phi_{n_1 n_2} = \phi_{n_2 n_1}$.

D. CROWD SHAPE

The groups in the crowd often have *circular shape* which can be used to measure how uniformly the crowd group/cluster is moving. Its application can be found in medical sciences when the shape of cancerous cells is different from that of neighboring cells. In human crowds it can be used in event management e.g. Olympics, where they group the performers in different shapes. To our knowledge no other research discusses the shape of the crowd groups in this context. Consequently the *shapes* of different groups are estimated using parametric equations of *circle, ellipse* and *oval*. The boundary of group is computed by joining the outermost points in group and the center of gravity (estimated mean value) is calculated. *Shapes* of different sizes with variable orientations and translations are placed on the cluster boundary. The mean absolute error is minimized and the best parameters (center, radius and orientation) define the shape of that specific cluster.

III. RESULTS AND DISCUSSION

We quantified four visual group properties in the form of descriptors defining *speed, direction, shape* and *merging probabilities*. The proposed descriptors are evaluated on CUHK dataset downloaded from [4], [5]. It includes crowd videos having variable crowd clusters and various densities. The 474 videos in the datasets are captured in 215 different scene environments like zebra crossings, malls, stations, and parks. A direct comparison of these descriptors with any other descriptors was not possible because of their novelty, quantification and group level estimations.

Three different cases are discussed in the paper. In case 1 we have a public pathway where people are moving in different *directions*. For the scene selected we have four clusters with their cluster numbers shown in figure 1. Group descriptors are evaluated for this case and are shown in table 1.

TABLE 3. Zebra crossing dataset.

Time	Cluster	Speed	Direction	Shape			Probabilities	
				Circle	Oval	Ellipse	C1	C2
21	C1	0.95	4.34	0.024	0.015	0.090	1	0.89
	C2	1.00	-155.25	0.011	0.011	0.032	0.89	1
22	C1	0.98	10.56	0.025	0.015	0.090	1	0.89
	C2	1.04	-156.95	0.012	0.011	0.032	0.89	1
23	C1	0.97	5.62	0.024	0.014	0.089	1	0.90
	C2	0.82	-133.75	0.009	0.009	0.070	0.90	1
24	C1	1.00	7.19	0.024	0.015	0.089	1	0.89
	C2	0.88	-165.88	0.008	0.009	0.068	0.89	1
25	C1	1.03	3.05	0.025	0.015	0.089	1	0.90
	C2	1.28	-141.58	0.009	0.009	0.069	0.90	1
26	C1	1.02	3.33	0.025	0.015	0.089	1	0.90
	C2	1.29	-146.25	0.008	0.008	0.031	0.90	1
27	C1	1.01	5.9	0.025	0.015	0.088	1	0.90
	C2	0.99	-151.73	0.007	0.007	0.032	0.90	1
28	C1	1.00	3.72	0.025	0.015	0.088	1	0.90
	C2	0.68	-154.12	0.007	0.007	0.032	0.90	1
29	C1	0.98	6.76	0.024	0.015	0.096	1	0.90
	C2	0.78	-150.74	0.007	0.007	0.032	0.90	1

In figure 1, cluster 1,2 and 3 are moving in the same *direction* while cluster 4 is moving in opposite *direction* which is verified by a minus sign shown in table 1. The *speed* of clusters remains almost constant from frame 32 to 37. But as the clusters 1 and 2 are merged, the *speed* descriptor value changes drastically which shows anomaly or ill defined descriptor due to outliers. To cater the problem of outliers an outlier correction is applied and the results are improved as shown in table 1. The *merging probabilities* are low because they are moving in the same *direction* and with a constant *speed*. None of the cluster qualifies as *circle, oval or elliptical* in *shape* so they have their error values. Lower error value means more closer to the shape. In this case lower error values are observed for *oval shape* at time 32 where all these clusters are nearly *oval* which can be seen in the figure 1 as well.

In case 2, we have a moving crowd (see figure 2) with three overlapping clusters numbered as 1, 2 and 3. Cluster 1 and 3 have motion in similar *direction* whereas cluster 2 has motion in the opposite *direction* which is shown by the minus sign in the results table 2. The clusters are moving with a constant *speed* as this scene is from a shopping mall. Similarly the *merging probability* of cluster 2 and 3 is higher because they are more closer to one another and opposite in *directions* as shown in the figure 2 as well as in table 2. None of the clusters have circular shape so we have higher error values.

Case 3 shows a scene from a zebra crossing where pedestrians are crossing the road from both sides thus having two clusters numbered as 1 and 2 (see figure 3). The direction of cluster 1 is along positive axis while the direction of cluster 2 is along negative axis. The results are shown in table 3. Since the distance between both the clusters is reducing with time and they are moving towards each other with constant *speed* so they have maximum *merging probabilities* as compared to previous cases. In this case it can be clearly seen that clusters are nearly *circular* or *oval* which is also verified in table 3.

IV. CONCLUSION

The quantification of crowded scenes can help in different disastrous situations. Most of the work in literature focuses on the crowd as a whole however the proposed technique focuses at group level and presents new quantitative descriptors to estimate *speed, direction, shape* and *merging probabilities* of different clusters moving in the crowd.

Simulations on various datasets demonstrate the applicability of proposed descriptors. Nonetheless, there is still much room for further improvements. The proposed descriptors can be extended for real time data. Similarly with time the size of the video data is increasing and may develop a need for faster algorithms in near future.

REFERENCES

[1] Z. Bolei, W. Xiaogang, and T. Xiaoou, "Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2871–2878.

[2] S. Saxena, F. Brémond, M. Thonnat, and R. Ma, "Crowd behavior recognition for video surveillance," in *Advanced Concepts for Intelligent Vision Systems*. Springer, 2008, pp. 970–981.

[3] Z. Bolei, T. Xiaoou, and W. Xiaogang, "Measuring crowd collectiveness," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3049–3056.

[4] J. Shao, C. C. Loy, and X. Wang, "Scene-independent group profiling in crowd," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2227–2234.

[5] J. Shao, C. C. Loy, and X. Wang, "Learning scene-independent group descriptors for crowd understanding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 6, pp. 1290–1303, Jun. 2017.

[6] M. Xu, C. Li, P. Lv, N. Lin, R. Hou, and B. Zhou, "An efficient method of crowd aggregation computation in public areas," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 10, pp. 2814–2825, Oct. 2018.

[7] Y. Xu, L. Lu, Z. Xu, J. He, J. Wang, J. Huang, and J. Lu, "Towards intelligent crowd behavior understanding through the STFD descriptor exploration," *Sens. Imag.*, vol. 19, p. 17, Dec. 2018.
[8] S. Ali and M. Shah, "A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–6.
[9] V. Rabaud and S. Belongie, "Counting crowded moving objects," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2006, pp. 705–711.
[10] 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.
[11] M. Paravisi, A. Werhli, J. J. Junior, R. Rodrigues, C. R. Jung, and S. R. Musse, "Continuum crowds with local control," *Comput. Graph. Int.*, vol. 25, no. 3, pp. 108–115, 2008.
[12] M. A. Cheriyadat, and R. J. Radke, "Detecting dominant motions in dense crowds," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 4, pp. 568–581, Aug. 2008.
[13] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 935–942.
[14] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, "Anomaly detection in crowded scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1975–1981.
[15] V. Saligrama and Z. Chen, "Video anomaly detection based on local statistical aggregates," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2112–2119.
[16] S. Wu, H. Yang, S. Zheng, H. Su, Y. Fan, and M.-H. Yang, "Crowd behavior analysis via curl and divergence of motion trajectories," *Int. J. Comput. Vis.*, vol. 123, no. 3, pp. 499–519, 2017.
[17] A. A. Alanazi and M. Bilal, "Crowd density estimation using novel feature descriptor," 2019, *arXiv:1905.05891*. [Online]. Available: https://arxiv.org/abs/1905.05891
[18] M. S. Zitouni, H. Bhaskar, J. Dias, and M. E. Al-Mualla, "Advances and trends in visual crowd analysis: A systematic survey and evaluation of crowd modelling techniques," *Neurocomputing*, vol. 186, pp. 139–159, Apr. 2016.
[19] H. Sharif, "An eigenvalue approach to detect flows and events in crowd videos," *J. Circuits, Syst. Comput.*, vol. 26, no. 7, 2017, Art. no. 1750110.
[20] J. M. Grant and P. J. Flynn, "Crowd scene understanding from video: A survey," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 13, no. 2, p. 19, 2017.
[21] A. A. Afiq, M. A. Zakariya, M. N. Saad, A. A. Nurfarzana, M. H. M. Khir, A. F. Fadzil, A. Jale, W. Gunawan, Z. A. A. Izuddin, and M. Faizari, "A review on classifying abnormal behavior in crowd scene," *J. Vis. Commun. Image Represent.*, vol. 58, pp. 285–303, Jan. 2019.
[22] D. Conigliaro, F. Setti, C. Bassetti, R. Ferrario, and M. Cristani, "ATTENTO: ATTENtion observed for automated spectator crowd analysis," in *Human Behavior Understanding*. Cham, Switzerland: Springer, 2013, pp. 102–111.
[23] Y. Benezeth, P.-M. Jodoin, V. Saligrama, and C. Rosenberger, "Abnormal events detection based on spatio-temporal co-occurences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2458–2465.
[24] P. Güler, A. Temizel, and T. T. Temizel, "Real-time global anomaly detection for crowd video surveillance using SIFT," in *Proc. 5th Int. Conf. Imag. Crime Detection Prevention*, Dec. 2013, pp. 1–12.
[25] B. Zhou, X. Tang, and X. Wang, "Coherent filtering: Detecting coherent motions from crowd clutters," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 857–871.

RAKHSHENDA JAVID received the M.S. degree from the National University of Sciences and Technology (NUST), Pakistan, where she is currently pursuing the Ph.D. degree. Her research interests include crowd behavior analysis and video processing.

M. MOHSIN RIAZ received the Ph.D. degree in electrical engineering from the National University of Sciences and Technology (NUST), Pakistan, in 2013. He is currently an Assistant Professor with the COMSATS Institute of Information Technology, Pakistan. His research interests include image processing and machine learning.

ABDUL GHAFOOR received the B.S. degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 1994, the M.S. degree in electrical engineering from the National University of Sciences and Technology (NUST), Islamabad, Pakistan, in 2003, and the Ph.D. degree from the University of Western Australia (UWA), in 2007. Since 2008, he has been with NUST, where he is currently an Associate Professor. His research interests include model/controller order reduction, image processing/matching, through wall imaging, and cognitive radio.

NAVEED IQBAL RAO received the M.S. degree in electrical engineering from the College of EME, National University of Sciences and Technology (NUST), Pakistan, in 2003, and the Ph.D. degree in computer software engineering from Tsinghua University, Beijing, China, in 2007. He is currently an Associate Professor with NUST. His research interests include image processing, computer vision, and machine learning.