A Crowd Behavior Identification Method Combining the Streakline with the High-Accurate Variational Optical Flow Model

XIAOFEI WANG¹,², ZISHU HE²,(Member, IEEE), RUI SUN¹, LEI YOU³, JUN HU¹ AND JIE ZHANG¹

¹ Department of Computer Science, Chengdu Normal University, Chengdu 610064, China
² School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
³ School of Information Science and Technology, Chengdu University, Chengdu 610106, China
⁴ Engineering Laboratory of Spatial Information Technology of Highway Geological Disaster Early Warning in Hunan Province, Changsha University of Science & Technology, Changsha 410114, China

Correspondence author: RUI SUN (151075@cdnu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61871085 and Grant 61671139, in part by the Scientific Research Foundation of Science and Technology Department of Sichuan Province of China under Grant 2018JY0129 and Grant 2018JY0202, in part by the Open Project Program of the State Key Lab of CAD&CG under Grant A1922, and in part by the Open Fund of Engineering Laboratory of Spatial Information Technology of Highway Geological Disaster Early Warning in Hunan Province under Grant kfj170602.

ABSTRACT A crowd behavior identification method is proposed by combining the streakline based on fluid mechanics with a high-accurate variational optical flow model in this paper. Firstly, calculated by the high-accurate variational optical flow model, the streaklines are used to acquire the crowd motion trajectory information. The angular histogram and the regions of interest in the scene are obtained by calculating and clustering the dasymetric dot maps of the starting and ending points of the trajectory, and then, combining the dasymetric dot map and angular histogram information to analyze whether there are specific crowd behaviors in the regions of interest, and thus to identify different types of crowd behavior in such scene. Finally, experimental comparison and analysis are made to verify the effectiveness and accuracy of the method proposed in this paper.

INDEX TERMS Streakline, High-accurate variational optical flow model, Clustering analysis, Angular histogram, Crowd behavior identification

I. INTRODUCTION Crowd behavior analysis has always been an important research topic in the field of computer vision, yet the occlusion between high-density crowds makes it difficult to identify and track individual targets in crowded scene. Therefore, it makes a lot of sense to understand the crowd behavior when the individual motion is unknown. Crowd behavior analysis is widely used in many application scenes, e.g. predicting road congestion to avoid traffic problems, discovering abnormal behavior or motion to avoid the occurrence of accidents, etc. Recently, some approaches, based on Social Force Model[1],[2], Temporal Context of Motion[3], Dynamic Texture[4], Stationary Crowd Groups[5],[6], Social Network Model[7], Sparse Reconstruction Cost (SRC)[8],[9], Structured Trajectory Learning (STL)[10], Deep Spatiotemporal Perspective[11], Top-Down Hierarchical Clustering Strategy[12], have been proposed to analyze crowd motion pattern and they indeed, work well in some specific scenes. However, these methods focus on a certain kind of crowd behavior and fail to pay attention to different types of crowd behavior. In many practical application cases, different types of crowd behavior may simultaneously occur in the same scene. Therefore, this paper focuses on how to detect and identify different types of crowd behavior. In existing references, five typical crowd behaviors, i.e. Lane, Bottleneck, Ring/Arc, Fountainhead and Block are defined in Ref. [13], and they are identified by calculating and analyzing the Jacobian matrix. Wu et al. [14] identified four crowd behaviors, i.e. Lane, Bottleneck, Arch and Fountainhead defined in Ref. [13] with Curl and Divergence of motion Trajectories (CDT) descriptors. Although these methods can achieve good identification effect in some scenes, they are based on traditional optical flow method (Lucas Kanade method). This means that when the video has very low resolution and jittering phenomenon, or when the crowds in the video move too fast or too slowly, the effect of traditional optical flow method is easily affected by these factors, which will lead to larger deviation in the calculating results of optical flow field, and thus result in a decline in the accuracy of these methods in identifying
different crowd behaviors. Therefore, in view of the situations above, this paper first adopts a high-accurate variational optical flow model [15] to improve the accuracy of optical flow detection, and then combines this variational optical flow model with the streakline [16] which has stronger anti-interference ability and can timely reflect the motion trend of the crowd targets in time and space to identify different crowd behaviors.

This paper mainly analyzes four specific crowd behaviors (similar to those defined in Ref. [13]), as shown in Fig. 1, including Lane, Bottleneck, Ring/Arch and Fountainhead, which have certain generality and universality and will appear in many scenes. Block, a crowd behavior defined in Ref. [13], can be decomposed into Lane representations in different directions by our algorithm, thus it is not discussed further in this paper. The rest of paper is organized as follows. The concept and definition of streakline combining a high-accurate variational optical flow model are mainly elaborated in Section II. The crowd behavior identification algorithm that we propose is introduced in Section III. In Section IV, we give the experimental comparison and analysis of crowd behavior identification. Finally, we give a conclusion to our paper in Section V.

![FIGURE 1. Four specific crowd behaviors.](image-url)

II. THE STREAKLINE COMBINING THE HIGH ACCURATE VARIATIONAL OPTICAL FLOW MODEL

This section involves a description of the streakline combined with the high accurate variational optical flow model in [15]. For completeness, a brief review of this variational optical flow model is required.

A. THE HIGH-ACCURATE VARIATIONAL OPTICAL FLOW MODEL

The basic thought of the variational optical flow calculation method is to convert the solution problem of optical flow into the minimization problem of global energy functional. The energy functional model of variational optical flow consists of data items and smoothing items, in which data items mainly include gray scale constancy assumption and gradient constancy assumption, while smoothing items mainly constrain the constancy assumptions to ensure a unique solution for the optical flow.

Let \( I_1, I_2 : (\Omega \subset \mathbb{R}^2) \rightarrow \mathbb{R} \) be the first and the second frame to be aligned. \( x := (x, y, t)^T \) denotes a point in the image domain \( \Omega \) and \( w := (u, v, 1)^T \) is the searched displacement vector between an image at time \( t \) and another image at time \( t + 1 \).

Energy functional model is expressed as follows:

\[
E(u, v) = E_{\text{gray}}(u, v) + \gamma E_{\text{gradient}}(u, v) + \mu E_{\text{smooth}}(u, v),
\]

where

\[
E_{\text{gray}}(u, v) = \int_{\Omega} \psi \left( I_2(x+w) - I_1(x) \right)^2 \, dx,
\]

\[
E_{\text{gradient}}(u, v) = \int_{\Omega} \psi \left( \nabla I_2(x+w) - \nabla I_1(x) \right)^2 \, dx,
\]

\[
E_{\text{smooth}}(u, v) = \int_{\Omega} \left( \nabla_x^2 u + \nabla_y^2 v \right) \, dx.
\]
I \geq \frac{\partial f(x+w)}{x}, \quad I = \frac{\partial f(x+w)}{y}
\begin{align}
I &= \frac{\partial f(x+w)}{x}, \quad I = \frac{\partial f(x+w)}{y} \\
I &= \frac{\partial f(x+w)}{x}, \quad I = \frac{\partial f(x+w)}{y} \\
I &= \frac{\partial f(x+w)}{x}, \quad I = \frac{\partial f(x+w)}{y}
\end{align}
(5)

Where, \( I_\alpha \) is the first-order partial derivative of \( I(x,y,t) \) in the \( x \) direction, \( I_t \) is the first-order partial derivative of \( I(x,y,t) \) in the time \( t \), and \( I_{x\alpha} \) is the second-order partial derivative of \( I(x,y,t) \) in the \( x \) direction. Other symbols can be expressed in the same manner.

According to the calculus of variations, a minimiser of (1) must fulfill the Euler-Lagrange equations
\[
\begin{align}
\psi'(I_{i+1} + \gamma (I_{i+1} + I_{i-1})) & \cdot (I_{i,1} + \gamma (I_{i,1} + I_{i-1} + I_{i+1} + I_{i+1})) \\
& - a \div(\nabla, u)^{\mathrm{T}} + \nabla, u \) = 0, \\
\psi'(I_{i+1} + \gamma (I_{i+1} + I_{i-1})) & \cdot (I_{i,1} + \gamma (I_{i,1} + I_{i-1} + I_{i+1} + I_{i+1})) \\
& - a \div(\nabla, v)^{\mathrm{T}} + \nabla, v = 0
\end{align}
(6)
(7)

The following can be obtained by solving (6) and (7) with iterative method
\[
\begin{align}
\psi'(I_{i+1} + \gamma (I_{i+1} + I_{i-1})) & \cdot (I_{i,1} + \gamma (I_{i,1} + I_{i-1} + I_{i+1} + I_{i+1})) \\
& - a \div(\nabla, u)^{\mathrm{T}} + \nabla, u \) = 0, \\
\psi'(I_{i+1} + \gamma (I_{i+1} + I_{i-1})) & \cdot (I_{i,1} + \gamma (I_{i,1} + I_{i-1} + I_{i+1} + I_{i+1})) \\
& - a \div(\nabla, v)^{\mathrm{T}} + \nabla, v \) = 0
\end{align}
(8)
(9)

where \( k \) is the number of iterations.

Assuming that the size of an image is \( m \times n \) and each pixel is calculated using (8) and (9), and a total of \( 2 \times m \times n \) large-scale equation sets \( AQ = b \) are constructed. The Gauss-Seidel iterative method or successive over-relaxation (SOR) iterative method can be used to solve this equation sets. In our algorithm, we can get the numerical solution of optical flow field of the entire image using successive over-relaxation iterative method. More implementation details are available in [15].

**B. IMPROVED STREAKLINE**

In fluid mechanics and flow visualization, streakline is well known as a tool for the measurement and the analysis of a flow. A streakline is a collection of particles initialized at a particular pixel and it is calculated as the follows.

Let \((x'_i(t), y'_i(t))\) be the position of a particle at time \(t\) and point \(p\) in frame \(i\) for \(i = 0,1,2,\ldots T\). Particle advection is achieved through the repetition of the initialization at \(p\) by
\[
\begin{align}
x'_i(t+1) &= x'_i(t) + u(x'_i(t), y'_i(t), t) \\
y'_i(t+1) &= y'_i(t) + v(x'_i(t), y'_i(t), t)
\end{align}
(10)

where \(u(x,y,t)\) and \(v(x,y,t)\) are the x and y components of the optical flow at position \((x,y)\) at time instant \(t\).

Particle advection for all \(i,t = 0,1,2,\ldots T\) using (10) can generate a large number of particle trajectories which may belong to a single sub-object, and particle trajectories distributed in one or more of such sub-object regions may be merged to give a representative trajectory that describes the motion therein. The trajectory of these particles is streakline. During trajectory extraction, samples are taken where every \(w\) pixel appears. The experiment has indicated that when \(w=5\), the corresponding result of trajectory can be found, respectively. The obtained streakline can recognize flow spatio-temporal changes more quickly than other methods [17]-[20], and what’s more, it can provide motion information of the flow for a period of time as well as capture crowd motions better in a dynamically changing flow. The accuracy and richness of the information obtained will have a great impact on the subsequent identification of crowd behaviors. More details concerning streakline can be found in Ref. [16],[21],[22]. Fig. 2 describes the corresponding streaklines obtained by the high-accurate variational optical flow model.

**III. PROPOSED ALGORITHM FOR CROWD BEHAVIOR IDENTIFICATION**

**A. DASYMETRIC DOT MAPS OF THE STARTING AND ENDING POINTS OF THE TRAJECTORY**

Using the streakline based on the high-accurate variational optical flow model described in Section 2, we can obtain a large number of crowd motion trajectories in the scene. In order to determine whether there are several specific crowd behaviors in such scene, we first need to confirm what regions
may contain several crowd behaviors in the video frame, which are namely, regions of interest. If there are multiple crowd behaviors in the video frame, multiple regions of interest can be acquired. It should be noted that crowd behaviors do not exist in all regions of interest, and subsequent processing of the regions of interest will eventually determine whether there are crowd behaviors in the regions of interest.

The dasymetric dot map of the starting and ending points of the trajectory is an image composed of the frequencies of interest. Obtained by conducting statistics on the position coordinates of the trajectory is an image composed of the frequencies of interest. Fig. 3 represents the dasymetric dot map of the starting and ending points of Lane, Bottleneck, Arch, and Fountainhead respectively. As can be seen from Fig. 3, for Lane and Ring/Arch, the density of the sampling points at the starting and ending points of the trajectory cluster is higher; for the motion that the crowds rush to a Bottleneck, the density of the sampling points at the ending point of the trajectory cluster is higher; for the motion that the crowds rush out from a Fountainhead, the density of the sampling points at the starting point of the trajectory cluster is higher. As the analysis above indicates that the calculation of the dasymetric dot maps of the starting and ending points of the trajectory cluster can be used as an important judgment basis in looking for the regions of interest.

![FIGURE 3. The dasymetric dot map of the starting and ending points of four crowd behaviors.](image)

### B. REGIONS OF INTEREST

By clustering the dasymetric dot maps, a certain range with the clustering center as its center is the region of interest. In this paper, DBSCAN clustering algorithm [23] is used for clustering analysis. Compared with other clustering algorithms (such as k-means, ISODATA algorithm, etc.), the advantages of this algorithm are, first, there is no need to know the number of classes in advance in the clustering operation; second, the cluster classes of any shapes can be found while the noise points can also be identified at the same time.

The density plots of the starting and ending points are respectively clustered by DBSCAN algorithm to remove some isolated sampling points, and get the number of cluster categories and the sampling points contained in each category. The position coordinates of the clustering center of each category are represented by the mean value of the position coordinates of the sampling point of the same category. The mathematical formula of the position coordinates of the clustering center is described as follows:

\[
C_{cx} = \frac{1}{N_m} \sum_{x \in C_m} x,
\]

\[
C_{cy} = \frac{1}{N_m} \sum_{y \in C_m} y,
\]

where \( N_m \) is the number of sampling points of the \( m \)th category, and the position coordinates of the clustering center of the \( m \)th category are \((C_{cx}, C_{cy})\).

Through the analysis of dasymetric dot maps, the clustering center of the dasymetric dot map of the starting point may be the starting point of the Ring/Arch and the Lane or the Fountainhead. The clustering center of the dasymetric dot map of the ending point may be the ending point of the Ring/Arch and the Lane or the Bottleneck.

### C. TRAJECTORY SCREENING

As there are a large number of trajectories being extracted, and that they may contain many fractured trajectories, the calculation amount of subsequent processing will see an increase as well as disturbance. Hence, after the calculation and clustering of dasymetric dot maps, the trajectories need to be effectively screened. The selected trajectory must meet the following conditions: the starting point of the trajectory is close to a clustering center of the dasymetric dot map of the starting point, or the ending point of the trajectory is close to a clustering center of the dasymetric dot map of the ending point. The trajectory selection process is as follows:

1. For each trajectory, we need to calculate the Euclidean distance from the ending point of the trajectory to each clustering center of the dasymetric dot map of the ending point to find the minimum distance. If the distance is less than the threshold \( \text{Traj\_Distance} \), indicating that the trajectory is close to the clustering center of the dasymetric dot map of an ending point, the trajectory will be saved. Otherwise, the trajectory will be removed.

2. For each trajectory, the Euclidean distance from the starting point of the trajectory to each clustering center of the dasymetric dot map of the starting point is calculated to find the minimum distance. If the distance is less than the
threshold \( \text{Traj\_Distance} \), indicating that the trajectory is close to the clustering center of the dasymetric dot map of a starting point, the trajectory will be saved. Otherwise, the trajectory will be removed.

(3) If there are few trajectories at the clustering center, indicating that there is an error in the clustering process, the information of the clustering center will be deleted.

(4) In order to make the clustering center more accurate, k-means method is adopted to re-calculate the clustering center of each category and thus, to get the corrected regions of interest.

D. ANGULAR HISTOGRAM OF TRAJECTORY CLUSTER

The absolute orientation information of a trajectory is also known as the orientation angle of the trajectory, which reflects the absolute motion orientation between front and rear sampling points of a trajectory. As shown in Fig. 4, \( P_{\tau-1}, P_{\tau}, \) and \( P_{\tau+1} \) are three continuous sampling points on the trajectory \( l = \{ P_{\tau} = (x_{\tau}, y_{\tau}), \tau = 1, 2, ..., num \} \), in which \( num \) is the number of sampling points on the trajectory, and the orientation angle at \( P_{\tau} \) is defined as follows:

\[
\alpha_{\tau} = \begin{cases} 
\arctan\left(\frac{y_{\tau+1} - y_{\tau}}{x_{\tau+1} - x_{\tau}}\right), & x_{\tau+1} - x_{\tau} > 0 \text{ and } y_{\tau+1} - y_{\tau} \neq 0 \\
\arctan\left(\frac{y_{\tau+1} - y_{\tau}}{x_{\tau+1} - x_{\tau}}\right) - \pi, & x_{\tau+1} - x_{\tau} < 0 \text{ and } y_{\tau+1} - y_{\tau} < 0 \\
\arctan\left(\frac{y_{\tau+1} - y_{\tau}}{x_{\tau+1} - x_{\tau}}\right) + \pi, & x_{\tau+1} - x_{\tau} < 0 \text{ and } y_{\tau+1} - y_{\tau} > 0 \\
-\pi/2, & x_{\tau+1} - x_{\tau} = 0 \text{ and } y_{\tau+1} - y_{\tau} < 0 \\
\pi/2, & x_{\tau+1} - x_{\tau} = 0 \text{ and } y_{\tau+1} - y_{\tau} > 0 \\
0, & y_{\tau+1} - y_{\tau} = 0
\end{cases}
\]  

(13)

For all trajectories of each clustering center, the orientation angles of the starting and ending points of the trajectory are calculated by (13), that is, the position coordinate \((x_{\tau}, y_{\tau})\) is the starting point of the trajectory, and the position coordinate \((x_{\tau+1}, y_{\tau+1})\) is the ending point of the trajectory. The angular histogram is counted according to the calculated orientation angle, in which the interval length is 10 degrees, and there are 36 intervals. Fig. 5- Fig. 8 shows the angular histograms of the starting and ending points of each corresponding trajectory in four crowd behavior trajectory clusters. It can be seen from these figures that the distribution of angular histogram varies with different crowd behaviors. The angular histogram of Lane/Arch trajectory cluster is mainly concentrated in one or two intervals, and the angular histogram of Bottleneck/Fountainhead trajectory cluster is distributed evenly. Lane/Arch and Bottleneck/Fountainhead can be distinguished by this feature.

FIGURE 5. Lane motion trajectory clusters and corresponding angular histograms of the starting and ending points.

FIGURE 6. Arch motion trajectory clusters and corresponding angular histograms of the starting and ending points.

FIGURE 7. Bottleneck motion trajectory clusters and corresponding angular histograms of the starting and ending points.
E. CROWD BEHAVIOR IDENTIFICATION

After screening the trajectories, we analyze the trajectories of each clustering center, namely, the regions of interest in this paper to judge different types of crowd behavior.

(1) Judgement of Lane/Arch and Bottleneck/Fountainhead: we need to find out the two intervals with the largest number of trajectories in the angular histogram. If the number of trajectories in these two intervals is greater than the threshold $th_N$, the regions of interest may be Lane or Arch, and the step (2) is proceeded. Otherwise, the regions of interest may be Bottleneck or Fountainhead, and the step (3) is proceeded.

(2) Judgement of Lane and Arch: for each trajectory of the clustering center, firstly, the orientation angle of a certain number of sampling points (the displacement between adjacent sampling points is too small) is calculated by (13), and then, the angular histogram is counted according to the calculated orientation angle, in which the interval length is 10 degrees, and there are 36 intervals. It can be seen from the analysis of Arch and Lane models that the angular histogram of each trajectory in the Arch is distributed evenly, and the angular histogram of each trajectory in the Lane is concentrated in a certain interval. Therefore, if the orientation angle of 90% single trajectory is concentrated in a certain interval of the angular histogram, the region of interest is judged to be Lane by this trajectory. Otherwise, it is judged to be Arch. Then, we conduct statistics on the judgment results of each trajectory in the clustering center, and take most of the judgment results as the final identification results.

(3) Judgement of Bottleneck and Fountainhead: Bottleneck and Fountainhead can be distinguished in a relatively easy way and judged by combining the dasymetric dot maps. If the region of interest is the clustering center of the dasymetric dot map of the starting point, the region of interest is the Fountainhead. If the region of interest is the clustering center of the dasymetric dot map of the ending point, the region of interest is the Bottleneck.

In order to understand the execution steps of the crowd behavior identification algorithm proposed in this paper in a better way, we have given the algorithm flowchart as shown in Fig. 9.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The experimental results and performance comparison analysis of crowd behavior identification are respectively presented to verify the accuracy and effectiveness of our algorithm. The videos used in the experiment are the crowd behavior data set of UCF with a total of 61 video sequences, and there are four crowd behaviors (143 in total) as shown in Fig. 1 in these videos. These video sequences come from the internet, PETS2009 database, etc., and involve crowd and traffic road videos in real scenes, covering different resolutions, angles of view and frame rates. In the experiment, $N_n = 0.7 \times N_{nud}$ ($N_{nud}$ is the total number of trajectories), the clustering radius is set as 20, and the number of the minimum data points in the neighborhood is set as $N_{nud}/20$.

A. CROWD BEHAVIOR IDENTIFICATION RESULTS

As the experimental results of four specific crowd behaviors are shown in Fig. 10, the first, second, third and fourth columns represent the identification processes of Fountainhead, Bottleneck, Lane and Arch respectively. From the first row to the fifth row, there are the original video frame,
extracted crowd motion trajectory, dasymetric dot maps of the starting and ending points of the trajectory, trajectory clustering result diagram and crowd behavior identification result in proper order. Bottleneck is denoted by red rectangle and red character “Bottleneck” and Fountainhead is denoted by those of yellow color. Ring/Arch is denoted by blue arch and blue character “Ring/Arch”, and the motion orientation is marked out while Lane is denoted by those of green color with motion orientation being marked out. Different colors represent different clusters in the trajectory clustering result diagram, and the black circle represents the clustering center of the ending point of the trajectory while the white circle represents the clustering center of the starting point of the trajectory. As can be seen from the experimental results in the first column of Fig. 10, there are two crowd behaviors in the scene, namely Fountainhead and Lane. In Fig. 10, the second, third and fourth columns are single crowd behavior, and the experimental results can reflect the real motion state of the crowd in the scene. Fig. 11 shows the identification results of some crowd behaviors in the video data set. In the first row of the second figure, “Block” motion type, as defined in the Ref. [13], is identified as “Lane” type with motion orientation by our algorithm.

FIGURE 10. The identification processes of four crowd behaviors.
FIGURE 11. Identification results of some crowd behaviors.

B. ANALYSIS OF CROWD BEHAVIOR IDENTIFICATION ALGORITHM

We have compared our algorithm with the Ref. [13] algorithm in terms of accuracy and time to verify the preciseness and effectiveness of the algorithm. The comparison between the two algorithms mainly concerns of accuracy rate (ACC), miss rate (MISS) and error rate (ERR) for crowd behavior identification. The mathematical formulas of ACC, MISS and ERR are defined as follows:

\[
ACC = \frac{D - F}{A} \tag{14}
\]

\[
MISS = \frac{M}{A} \tag{15}
\]

\[
ERR = \frac{F}{A} \tag{16}
\]

where \(D\) is the number of crowd behaviors identified (including the number of crowd behaviors that are correctly identified and identified as other motions), \(F\) is the number of crowd behaviors wrongly identified, \(A\) is the total number of true crowd behaviors, and \(M\) is the number of crowd behaviors that are not identified. The judgment of correct identification chiefly depends on the comparison of the identified crowd behaviors with the real motion regions manually marked. For Bottleneck and Fountainhead, the identification is judged to be correct as long as the Euclidean distance between the position identified by algorithm and the actual position is within the specified threshold. For Lane and Ring/Arch, in addition to satisfying the Euclidean distance condition, the accuracy of the marked crowd motion orientation and region must also be satisfied. Table I shows the comparison of the accuracy between our algorithm and the Ref. [13] algorithm.

<table>
<thead>
<tr>
<th>Category of Crowd Behavior</th>
<th>Algorithm</th>
<th>ACC (%)</th>
<th>MISS (%)</th>
<th>ERR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fountainhead</td>
<td>Our algorithm</td>
<td>82.76</td>
<td>10.34</td>
<td>6.90</td>
</tr>
<tr>
<td>Bottleneck</td>
<td>Our algorithm</td>
<td>80</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Ref. [13] algorithm</td>
<td>65</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Lane</td>
<td>Our algorithm</td>
<td>80.30</td>
<td>1.54</td>
<td>18.16</td>
</tr>
<tr>
<td></td>
<td>Ref. [13] algorithm</td>
<td>68.18</td>
<td>15.15</td>
<td>16.67</td>
</tr>
<tr>
<td>Ring/Arch</td>
<td>Our algorithm</td>
<td>82.14</td>
<td>7.14</td>
<td>10.72</td>
</tr>
<tr>
<td></td>
<td>Ref. [13] algorithm</td>
<td>60.71</td>
<td>17.86</td>
<td>21.43</td>
</tr>
</tbody>
</table>

TABLE II

TIME-CONSUMING COMPARISON BETWEEN OUR ALGORITHM AND REF. [13] ALGORITHM

<table>
<thead>
<tr>
<th>Video Name</th>
<th>Algorithm</th>
<th>Average time per frame for the extraction of the regions of interest /ms</th>
<th>Average Time per frame for the identification of crowd behavior /ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video 1</td>
<td>Our algorithm</td>
<td>98.36</td>
<td>12.84</td>
</tr>
<tr>
<td></td>
<td>Ref. [13] algorithm</td>
<td>80.01</td>
<td>108.55</td>
</tr>
<tr>
<td>Video 2</td>
<td>Our algorithm</td>
<td>94.87</td>
<td>13.12</td>
</tr>
</tbody>
</table>

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.
As can be seen from Table I, all indexes of our algorithm are significantly better than those in the Ref. [13] for Fountainhead, Bottleneck and Ring/Arch. This is because the motion information acquired by the streakline trajectory based on high-accurate variational optical flow model in our algorithm can well represent this kind of motion. For Lane, both ACC and MISS of our algorithm are better than those of the Ref. [13] algorithm, but the ERR is slightly higher than Ref. [13] algorithm for the reason that the initial position of the trajectory extracted by Arch is quite different from the hypothesis in this paper, so Arch is sometimes mistakenly identified as Lane.

Table II shows the time-consuming comparison of four crowd behavior videos as shown in Fig. 10 between our algorithm and Ref.[13] algorithm (four videos in Fig. 10 are named video 1, video 2, video 3 and video 4 in proper order from left to right) which can further verify the effectiveness of the algorithm. In order to show the time-consuming composition of the algorithm in a clearer way, the time consumed by the algorithm is divided into two parts, and they are, the time to extract the regions of interest and the time to identify crowd behaviors subsequently. Fig. 12 and Fig. 13 are the experimental results of two algorithms on 61 video sequences. They are the comparison of the average computing time per frame for extracting the regions of interest, and the average computing time per frame for crowd identification.

As can be seen from the time-consuming comparison in extraction of the regions of interest between our algorithm and Ref. [13] algorithm in Table II and Fig 12, our algorithm takes more time than Ref. [13] algorithm in this part. The main reason is that the trajectory extraction method adopted by our algorithm is the streakline flow based on high-accurate variational optical flow model, which has high computation complexity. However, the acquired motion trajectory information is richer and more accurate, which lays a foundation for the higher accuracy of subsequent crowd behavior identification. Ref. [13] algorithm adopts traditional Lucas-Lanade optical flow method, which is relatively simple to calculate and less time-consuming, but the crowd motion information acquired by this method is not rich, which will have a great negative impact on the accuracy of subsequent crowd behavior identification. At the same time, it can be seen from Fig. 12 that there isn’t much difference between the two methods in terms of the average computing time per frame of extracting the regions of interest for some video sequences. The main reason is that the optical flow in these videos belongs to the category of large displacement optical flow, and the variational optical flow method adopted in this paper has greater computational efficiency in the calculation of large displacement optical flow. It can be seen from the time-consuming comparison in the identification of crowd behaviors between the two algorithms.

As can be seen from the time-consuming comparison in the identification of crowd behaviors in Table II and Fig 13, our algorithm takes less time than Ref. [13] algorithm in this part for the reason that the regions of interest are identified with angular histograms of the trajectory and trajectory cluster in our algorithm. To calculate the Jacobi matrix, Ref. [13] algorithm needs to calculate the dense optical flow of the regions of interest to acquire the optical flow vector of every pixel, which is consequently, more time-consuming. To sum up, total time consumed by our algorithm is superior to Ref. [13] algorithm.
V. CONCLUSION
Combining the streakline with the high-accurate variational optical flow model, a crowd behavior identification method is proposed in this paper. We first collect the trajectory information that can represent the crowd motion trend by making the streaklines acquired by the variational optical flow model to perform particle advection. Then, the dasymetric dot maps of the starting and ending points of the trajectory are calculated and clustered to get the angular histograms and the regions of interest in the scenes. In the final step, we have identified different crowd behaviors by processing and analyzing the angular histograms and the regions of interest in the scenes. The experimental results show that the crowd behavior identification method proposed in this paper has a better identification rate. At the same time, for some special scenes that the video resolution is lower, or have jitters, the method in the video move too fast or too slowly, the method in this paper also proves to be more accurate and adaptable.

REFERENCES
XIAOFEI WANG received the Ph.D. degree from the College of Electronics and Information Engineering, Sichuan University, in 2014. He currently holds a postdoctoral position with the School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC), China. He is currently an Associate Professor at the Department of Computer Science, Chengdu Normal University, China. His main research interests include crowd behavior analysis and image processing.

ZISHU HE received Ph.D. degree in signal and information processing from UESTC, in 2000. He became a Member (M) of IEEE in 2011. Currently he is a professor in the School of Information and Communication Engineering of UESTC. His main research interests include array signal processing, digital beam forming, the theory on multiple-input multiple-output (MIMO) communication and MIMO radar, adaptive signal processing and interference cancellation.

RUI SUN received the Ph.D. degree from the College of Computer Science, Sichuan University, in 2014. He currently holds a postdoctoral position with Chengdu information technology co. LTD of Chinese Academy of Sciences. He is currently an Associate Professor at the Department of Computer Science, Chengdu Normal University, China. His main research interests include crowd complex networks and intelligent computing.

LEI YOU received the Ph.D. degree in nuclear technology and applications from Chengdu University of Technology, China, in 2012. Currently, he is associate professor at the Chengdu University, China. His main research interests include crowd behavior analysis and image processing.

JUN HU received the Ph.D. degree from the College of Computer Science, Sichuan University, in 2011. Currently, he is Professor at Chengdu Normal University. His main research interests include crowd evacuation analysis and image processing.

JIE ZHANG received Ph.D. degree from the College of Electronics and Information Engineering, Sichuan University, in 2015. His main research interests include machine learning, pattern recognition and image processing.