Translation domain segmentation model based on improved cosine similarity for crowd motion segmentation

Haibin Yu
Guoxiong Pan
Li Zhang
Zhu Li
Mian Pan

Translation domain segmentation model based on improved cosine similarity for crowd motion segmentation

Haibin Yu,^a Guoxiong Pan,^a Li Zhang,^b Zhu Li,^a and Mian Pan^{a,*}

^aHangzhou Dianzi University, College of Electronics and Information, Hangzhou, China

^bHangzhou Dianzi University, School of Computer Science and Technology, Hangzhou, China

Abstract. With the continuous growth of the global population, large-scale public gatherings have become more common, and crowd management at these gatherings has become an urgent problem for public safety management. Crowd motion analysis and early warning based on crowd motion segmentation in video surveillance systems has become an important research topic in computer vision. A translation domain segmentation (TDS) model based on improved cosine similarity (ICS) is proposed to segment moving crowds with different crowding levels and complex motion modes. The method reconstructs the objective function of the basic TDS model by ICS to simultaneously measure the difference between both magnitude and direction of two vectors; thus, under-segmentation due to the magnitude difference can be avoided. By switching between “localization” and “globalization” modes of the objective function, the algorithm can be applied to segment crowds with different densities and motion states. Moreover, by simultaneously introducing local motion magnitude and local frame difference magnitude thresholds, nonforeground regions can be excluded from the initial regions during region evolution. Experimental results show that the proposed method achieves superior performance and higher accuracy compared to existing flow field-based methods when applied to complex scenes containing moving crowds.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JEI.28.2.023011](https://doi.org/10.1117/1.JEI.28.2.023011)]

Keywords: crowd motion segmentation; translation domain segmentation; improved cosine similarity; motion field.

Paper 180821 received Sep. 22, 2018; accepted for publication Feb. 26, 2019; published online Mar. 13, 2019.

1 Introduction

With the continuous growth of the global population, large-scale public gatherings—such as religious pilgrimages, parades, celebrations, and sporting events—have become more frequent; therefore, crowd management at these gatherings has become an urgent and important problem to be addressed in the field of public safety management. Usually, people in these gatherings move in a limited space, such as urban streets, semiencllosed squares, and fully enclosed shopping malls, which are prone to catastrophic accidents, such as crowd congestion and trampling. To effectively supervise and guide crowds at such gatherings for avoiding accidents, video/image crowd monitoring technology has been widely used in video surveillance systems in many public places.^{1–8} Considering that the most common cause of various accidents is collision among crowds moving in different directions, dividing moving crowds into subgroups based on their movement direction is particularly important; this has been an important research topic in the field of image processing and computer vision.⁹

In recent years, researchers have proposed a variety of methods for motion segmentation of crowds for different congestion levels in complex scenes;^{10–12} these can be classified as traditional methods based on single person tracking^{13,14} and motion modeling methods that model the overall motion characteristics of the crowd.^{15–26} Among them, the segmentation method based on single person tracking relies on the complete trajectories of all individuals (or feature points and feature areas) in the crowd; therefore,

it is only suitable in the case when there is no occlusion or slight occlusion. Once the crowd increases, there is severe occlusion and the acquisition of complete trajectories becomes very difficult resulting in a significant drop in the performance of such methods.

Due to the difficulties encountered in tracking individuals (such as occlusion), considering the overall self-organization characteristics of a moving crowd, many researchers have started to focus on the entire crowd instead of individuals in the crowd so that an appropriate motion model can be adopted to model the overall motion of the crowd. At present, studies mainly focus on two kinds of methods for crowd motion modeling: flow-based model and nonflow-based model.⁹ Because a flow field such as the optical flow field is usually a good estimation of the motion field, a crowd motion segmentation method based on the flow field model was proposed as one of the earliest methods for crowd motion segmentation. It is the basis for a variety of subsequently proposed methods. In a crowd motion segmentation method based on the flow model, apart from the most commonly used optical flow fields,^{15,17–21} flow fields used for crowd motion modeling also include particle flow field^{22,23} and streak flow field.^{24–26} Most existing flow-based models only focus on a specific crowd density. For example, particle flow fields and streak flow fields are, generally, only suitable for a high-density slow-moving crowd. When this kind of model is applied to a low-density crowd, the possibility of oversegmentation is high. To make the motion model based on flow field applicable to crowds with any density, considering that different flow fields are essentially vector fields, some researchers introduced a vector domain segmentation method for specific vector distribution in the flow field model, such as local translation domain segmentation

*Address all correspondence to Mian Pan, E-mail: ai@hdu.edu.cn

(LTDS) proposed by Wu and Wong²⁰. Although LTDS can be applied for motion segmentation of crowds with different densities, it is prone to undersegmentation between the foreground and background because its objective function considers only the direction differences between vectors, regardless of the magnitude differences.

Among the nonflow-based methods, the most representative are dynamic texture-based methods^{27–30} and tracklet-based methods.^{31–34} Dynamic texture-based methods first model the moving crowd as a dynamic texture³⁵ with spatio-temporal statistical properties, and then use the matching among model parameters to perform crowd motion segmentation or abnormal behavior detection. As current dynamic texture models are relatively simple (such as linear dynamic system), crowd motion segmentation methods based on dynamic texture are currently only applicable to crowds with relatively simple movement modes and medium/low-densities. Tracklet-based methods first obtain the tracklets of the keypoints in the crowd using a tracker [such as Kanade–Lucas–Tomasi (KLT)], and then applies the similarity measure among the tracklets to complete the crowd motion segmentation. For example, Sharma and Guha³⁴ used the tracklets acquired by KLT to form a complete trajectory of keypoints, and then applied the trajectory clustering approach (TCA) to complete the segmentation of moving crowds. However, this method is not suitable for short-term motion segmentation owing to its requirement for long-term trajectories of the keypoints in the crowd. Zhou et al.^{31,32} and Fan et al.³³ utilized coherent filtering (CF) based on coherent neighbor invariance to perform short-term crowd motion segmentation via the local spatiotemporal relationships and motion correlations among tracklets. However, because coherent neighbor invariance primarily focuses on pairwise motion consistency and ignores the motion difference among all the keypoints in the local region of the center point, undersegmentation occurs easily in crowds with numerous motion patterns.

In this study, a TDS method based on improved cosine similarity (ICS) is proposed for crowd motion segmentation. This method adopts the optimized vector domain segmentation model in the flow field and can be applied for the segmentation of crowds with different crowding levels and complex motion modes. The main contributions of this study are as follows:

1. ICS is used to reconstruct the objective function of the TDS model. Compared with the similarity measure used in TDS, which can only measure the direction difference between vectors, ICS can measure the difference between both magnitude and direction of two vectors, which can effectively avoid undersegmentation caused by the magnitude difference; thus, it can greatly improve the adaptability of the TDS model.
2. The proposed method can be applied for motion segmentation of crowds with different densities and motion modes. When used for segmenting a crowd with high density and slow movement, the objective function can be localized so that the segmentation is completed by evaluating the motion consistency among local regions; when used for segmenting a moving crowd with medium- or low-density, the segmentation can be accomplished directly by evaluating

the motion consistency between all vectors and each initial region.

3. In the selection of initial regions, both the local motion magnitude threshold and local frame difference magnitude threshold are introduced in all candidate regions with the best local motion consistency to exclude the nonforeground regions from the initial regions as much as possible, which will inevitably lead to incorrect segmentation results.

2 Related Works

To overcome the problems of traditional methods based on individual tracking, researchers have performed crowd motion segmentation by establishing a series of crowd motion models according to the self-organization characteristics of a moving crowd. Currently, most crowd motion models are based on flow fields. Apart from the most commonly used optical flow field,^{15,17–21} the flow fields used for modeling moving crowd also include particle flow field^{22,23} and streak flow field.^{24–26}

2.1 Optical Flow Field

Research on modelling a moving crowd by optical flow is the most developed. For example, Hu et al.¹⁵ used a Gaussian adaptive resonance theory network to extract the prominent flow vectors from a dense optical flow field and then constructed a directed neighborhood graph based on the shortest path search for these flow vectors. Hierarchical agglomerative clustering algorithm was used to segment the directed neighborhood graph to obtain the final segmentation results of a moving crowd. Zhang et al.¹⁶ obtained trajectory chains of feature points using an orientation distribution function in a sparse optical flow field; then, they used a spectral clustering method (number of clusters determined by prior knowledge) to complete the classification of trajectories to perform slow-motion segmentation in densely crowded locations. This type of approach usually establishes a crowd motion model directly based on the similarity among optical flow vectors. Because the model is relatively simple, it is generally only applicable to a crowd with good motion consistency.

2.2 Particle Flow Field

Particle flow field is based on the Lagrangian fluid dynamics framework. When used for motion segmentation, the particle trajectory can be estimated by numerical integration through the moving particle grid in the optical flow field. For example, Ali and Shah²² first used the Lagrangian particle dynamic (LPD) algorithm to estimate the particle flow field based on the optical flow field, and then used the finite time Lyapunov exponent to extract the boundary among different flow fields to obtain the segmentation results for a crowd. Because a particle flow field ignores the spatial domain variation and the time delay is obvious, a crowd motion segmentation method based on particle flow field is only suitable for a high-density slow-moving crowd.

2.3 Streak Flow Field

To obtain a better flow-based model for a moving crowd, Mehran et al.²⁴ introduced streak line in fluid dynamics to

estimate the motion field of a crowd, which is called streak flow. Because streak flow can preserve the motion information of the flow for a period, the segmentation performance is better than particle flow when applied to crowds with obvious motion changes. However, the streak flow field is obtained on the basis of the optical flow field. The calculation process of streak flow requires very high accuracy optical flow; however, it is very difficult to meet accuracy requirements for optical flow using conventional methods.

2.4 Vector Domain Segmentation for Flow Field Model Application

To make the flow field model suitable for crowds with different densities and motion modes, some researchers have used the vector domain segmentation model^{36,37} in the vector field to establish a flow-based model of a moving crowd, such as the LTDS model proposed by Wu and Wong²⁰. Through the normalization of the vector magnitude and the addition of foreground area constraints, LTDS extends the TDS model proposed by Roy et al.³⁶ to the nonunit field; it can be applied to the segmentation of crowds with different densities by localization. However, because the objective function adopted by LTDS only considers the direction difference between vectors, undersegmentation between foreground and background can easily occur.

The proposed crowd motion segmentation method is based on a segmentation model in the optical flow field. The ICS measure is used to replace the similarity measure based on normalized vector inner product in the objective function of LTDS. The improved objective function can distinguish between both magnitude and direction of vectors to achieve better segmentation performance.

3 Method

The overall architecture of the crowd motion segmentation method proposed in this study is shown in Fig. 1. For applying the algorithm to a high-density slow-moving crowd: (1) the optical flow field between two consecutive frames is first obtained by means of an optical flow extraction algorithm (such as the Brox algorithm in Ref. 38); (2) the local ICS is used to obtain the local motion consistency map (LMCM); (3) the initial regions for level set evolution are extracted from the LMCM; (4) through the level set algorithm, the objective function of ICS-based LTDS is used to complete the region evolution in the LMCM, and regions with multiple motion directions are obtained; (5) the final crowd segmentation results are obtained by merging the regions with same direction. In contrast, when the algorithm is used for crowds with medium- or low-densities and complex motion modes, after the initial region extraction step, region evolution is not performed in the LMCM but in the global motion consistency map (GMCM) corresponding to each initial region. The GMCM is obtained by calculating the ICS between the average vector of each initial region and all the vectors.

3.1 TDS and LTDS

Given a unit vector field E , the translation domain (TD) is defined as the region in which all vectors have the same direction. The field lines in a TD are a set of parallel lines. Therefore, there must exist a unique normal vector perpendicular to the field lines, denoted as a , which can be used as a characteristic parameter to represent the TD. a is also known as the dominant translation parameter. Roy et al.³⁶ proved that a vector domain is a TD if and only if there is a unique vector $a(\Omega)$, for $\forall x \in \Omega$, that satisfies the following equation:

Fig. 1 Architecture of the proposed method.

$$a(\Omega) \cdot E(x) = 0. \quad (1)$$

According to Eq. (1), Roy et al.³⁶ proposed a TDS model for vector domain segmentation. The corresponding objective function of the model is as follows:

$$J(\tau) = \int_{\Omega(\tau)} [a(\tau) \cdot E(x)]^2 dx - \mu \int_{\Omega(\tau)} dx, \quad (2)$$

where Ω is a TD; τ is an evolution parameter defined to facilitate the use of the active contour algorithm; $[a(\tau) \cdot E(x)]^2$ is used to measure the motion consistency between all vectors in E and the TD determined by $a(\tau)$, which can be called the GMCM corresponding to $a(\tau)$; the integral term $\int_{\Omega(\tau)} [a(\tau) \cdot E(x)]^2 dx$ represents the sum of the squares of the errors in Ω in the GMCM, which reflects the motion consistency of Ω ; the integral term $\int_{\Omega} dx$ represents the area of Ω ; and μ is a positive constant.

On the basis of Eq. (2), given a deformation velocity V , the Gâteaux derivative in the direction of V can be given as follows:

$$\frac{dJ}{d\tau}(\tau, V) = - \int_{\partial\Omega(\tau)} [[a(\tau) \cdot E(x)]^2 - \mu] V(x) \cdot N_\tau(x) dl(x), \quad (3)$$

where $\partial\Omega$ is the boundary of Ω , $dl(x)$ is the integral variable of $\partial\Omega$, and N_τ is the unit normal vector of $\partial\Omega$ pointing in the direction of the interior of Ω . Then, the minimization of Eq. (2) is performed to use the active contour algorithm to minimize Eq. (3); the corresponding evolution equation of the active contour algorithm is as follows:

$$\begin{cases} \frac{\partial C}{\partial \tau}(\tau) = [[a(\tau) \cdot E]^2 - \mu] N_\tau, \\ C(\tau = 0) = C_0 \end{cases}, \quad (4)$$

where C is $\partial\Omega$, which is the contour of the TD for evolution, and C_0 is its initial value.

Meanwhile, Roy et al.³⁶ also proposed an estimation method for the dominant translation parameter $a(\tau)$, which takes the first half of the right side of Eq. (2) as the objective function and obtains the optimal estimate of $a(\tau)$ by minimizing the objective function, as follows:

$$\begin{aligned} \hat{a}(\tau) &= \arg \min_a \left[\int_{\Omega(\tau)} [a(\tau) \cdot E(x)]^2 dx \right] \\ &= \arg \min_a \left\{ a(\tau) \left[\int_{\Omega(\tau)} E(x) E(x)^T dx \right] a(\tau)^T \right\} \\ &= \arg \min_a \left[a(\tau) Q_\tau a(\tau)^T \right]. \end{aligned} \quad (5)$$

As Q_τ is a real symmetric matrix, the optimal estimation $\hat{a}(\tau)$ of $a(\tau)$ can be determined as the eigenvector corresponding to the smallest eigenvalue of Q_τ by the quadratic programming model.

Because the TDS model is proposed for the TD in the ideal unit vector field, it cannot be directly applied to the segmentation of an actual motion field. To make the TDS model applicable to crowd motion segmentation, Wu and Wong²⁰ established a LTDS model by the normalization

of the vector, addition of a foreground motion region area constraint, and localization according to Eq. (2). The objective function of LTDS is as follows:

$$\begin{aligned} J(A, \tau) &= \int_{\Omega(\tau)} \mathcal{F}[A(x), x] dx - \mu \int_{\Omega(\tau)} \mathcal{G}(\|U(x)\|) dx, \\ \mathcal{F}[A(x), x] &= \frac{1}{|\ell(x)|} \int_{\ell(x)} \left(A(x) \cdot \frac{U(\hat{x})}{\|U(\hat{x})\|} \right)^2 d\hat{x}, \\ \mathcal{G}(\|U(x)\|) &= \frac{1}{1 + \left(\frac{\gamma}{\|U(x)\|} \right)^2}, \end{aligned} \quad (6)$$

where Ω is a local TD; U is a motion field and its magnitude $\|U\|$ is mapped to the range $[0, 1]$; $\ell(x)$ is a neighborhood centered on x and $|\ell(x)|$ is the area of $\ell(x)$; $A(x)$ is the dominant translation parameter of $\ell(x)$ and can be estimated according to Eq. (5); and $\mathcal{F}[A(x), x]$ is the mean of the sum of the squared errors inside $\ell(x)$, reflecting the local motion consistency. Given the size of $\ell(x)$, we can calculate $\mathcal{F}[A(x), x]$ of each vector in U and use them to form an LMCM, which can better describe the consistent slow motion of high-density crowds than GMCM; γ is a constant; $\mathcal{G}(\|U(x)\|)$ is proportional to $\|U\|$, and is used to measure the extent to which x belongs to the foreground, so that $\mu \int_{\Omega(\tau)} \mathcal{G}(\|U(x)\|) dx$ can be considered as the foreground area constraint.

Similar to the evolution equation [Eq. (4)] of TDS, and considering that $\mathcal{F}[A(x), x]$ is an even function, which makes it impossible to distinguish two completely opposite vectors, the evolution equation of the active contour algorithm corresponding to the LTDS model is as follows:

$$\begin{cases} \mathcal{K}(x) = \begin{cases} \mathcal{F}[A(x), x] - \mu \mathcal{G}(\|U(x)\|), & \text{if } U(x) \cdot \bar{U}_{\ell(x) \cap \Omega} > 0 \\ \|\mathcal{F}[A(x), x] - \mu \mathcal{G}(\|U(x)\|)\|, & \text{if } U(x) \cdot \bar{U}_{\ell(x) \cap \Omega} \leq 0 \end{cases} \\ \frac{\partial C}{\partial \tau}(\tau) = \mathcal{K}(x) \cdot N_\tau, \\ C(\tau = 0) = C_0 \end{cases}, \quad (7)$$

where $\bar{U}_{\ell(x) \cap \Omega}$ is the average vector of the region $\ell(x) \cap \Omega$.

3.2 ICS

It is well known that the cosine similarity (CS) measure uses the cosine of the angle between two vectors in the vector space to measure the difference between these two vectors:

$$CS(x, y) = \cos \theta_{xy} = \frac{x \cdot y}{\|x\| \cdot \|y\|}, \quad (8)$$

where θ_{xy} is the angle between vectors x and y . The closer the cosine of the angle is to 1, which indicates that the angle will be close to 0, the more similar the two vectors will be. However, as the CS normalizes the magnitudes of the two vectors during calculation, it is impossible to measure the difference between the magnitudes of the two vectors.

To measure both magnitude and direction difference between the two vectors using CS, we first add a scale factor that can measure the difference between the magnitudes of the two vectors in the original CS, as shown in Eq. (8); this is defined as follows:

$$k_M(x, y) = \frac{\min(\|x\|, \|y\|)}{\max(\|x\|, \|y\|)}. \quad (9)$$

Considering that the range of $k_M(x, y)$ is $[0, 1]$ and that of CS is $[-1, 1]$, to unify the ranges of both values, by linear mapping, the CS is adjusted to

$$CS_1(x, y) = \frac{1}{2}(1 + \cos \theta_{xy}) = \frac{1}{2} \left(1 + \frac{x \cdot y}{\|x\| \cdot \|y\|} \right). \quad (10)$$

Based on Eqs. (9) and (10), the ICS is given as follows:

$$\begin{aligned} ICS(x, y) &= CS_1(x, y) \cdot k_M(x, y) \\ &= \frac{1}{2} \left(1 + \frac{x \cdot y}{\|x\| \cdot \|y\|} \right) \cdot \frac{\min(\|x\|, \|y\|)}{\max(\|x\|, \|y\|)} \\ &= \frac{\|x\| \cdot \|y\| + x \cdot y}{2 \cdot \max^2(\|x\|, \|y\|)}. \end{aligned} \quad (11)$$

The basic ICS defined in Eq. (11) cannot be directly used for TD segmentation because there is an imbalance when evaluating the magnitude and direction difference. For example, if the similarity is given as 0.5, when $k_M \rightarrow 1$ (i.e., there is almost no difference in magnitude between the two vectors) and $\cos \theta_{xy} \rightarrow 0$ can be obtained from Eqs. (10) and (11), then $\theta_{xy} \rightarrow \pi/2$; however, when $\theta_{xy} \rightarrow 0$ (i.e., there is almost no difference in direction between the two vectors), $CS_1(x, y) \rightarrow 1$, then $k_M \rightarrow 0.5$. This shows that the ICS between the two vectors with small difference in direction and half the difference in magnitude is the same as the ICS between the two vectors that are perpendicular to each other and have the same magnitude, as shown in Fig. 2(a). Considering that there are two vectors perpendicular to a given vector in the same plane, $ICS(x, y_1) = ICS(x, y_2) = ICS(x, y_3)$.

For the actual application of this study, $ICS(x, y_1)$, $ICS(x, y_2)$, and $ICS(x, y_3)$ should not be the same. Although the magnitude of y_1 is quite different from that of x , they can

Fig. 2 Imbalance of vectors with the same ICS defined in Eq. (11). (a) ICS between the two vectors with no difference in direction but half the difference in magnitude [i.e., $ICS(x, y_1)$] is the same as the ICS between the two vectors perpendicular to each other with same magnitude [i.e., $ICS(x, y_2)$ and $ICS(x, y_3)$]. (b) An optical flow field containing the optical flow vectors having the relationship described in panel (a) (the vector in red is the given reference vector x).

be considered to have good similarity because they have the same directional angle (i.e., they have exactly the same direction of motion). In contrast, both y_2 and y_3 are perpendicular to x , resulting in a large difference in their direction of motion. In this case, even if they have the same magnitude, the similarity between them should be very low. Otherwise, two vectors with vertical or even opposite directions may be segmented into the same region, i.e., all the vectors in Fig. 2(b) will be segmented into the same region. That is, when segmenting vectors, the importance of the similarity between the directional angles is significantly higher than the similarity between the magnitudes; for the case shown in Fig. 2, the condition $ICS(x, y_1) \gg ICS(x, y_2) = ICS(x, y_3)$ should be satisfied. To enhance the proportion of directional angle similarity in the basic ICS defined in Eq. (11), we add an exponential adjustment factor β to $CS_1(x, y)$ that measures the directional angle difference in ICS; then ICS is redefined as follows:

$$\begin{aligned} ICS(x, y, \beta) &= [CS_1(x, y)]^\beta \cdot k_M(x, y) \\ &= \left[\frac{1}{2} \left(1 + \frac{x \cdot y}{\|x\| \cdot \|y\|} \right) \right]^\beta \cdot \frac{\min(\|x\|, \|y\|)}{\max(\|x\|, \|y\|)}, \end{aligned} \quad (12)$$

where $\beta > 1$. The larger β is, the greater will be the proportion of the directional angle similarity in ICS. Meanwhile, if the minimum similarity $s_{\min}$ is given, i.e., $ICS(x, y, \beta) \geq s_{\min}$, the larger β is, the smaller will be the variation range (defined as $\mathcal{R}_\theta^\beta$) of the angle difference θ_{xy} , as shown in Fig. 3.

The value of β can be determined according to the distribution of the vectors in the optical flow field, especially the distribution of the vector direction by using the maximum allowable $\mathcal{R}_\theta^\beta$ at minimum similarity $s_{\min}$. For example, the minimum similarity $s_{\min}$ is given as 0.5 and $\mathcal{R}_\theta^\beta$ is given as $\pi/6$ at this similarity [as shown in Fig. 4(a)]; when $k_M = 1$, $[CS_1(x, y)]^\beta$ takes the minimum value of 0.5 and the corresponding direction angle θ_{xy} has the maximum value of $\pi/6$, so the following equation holds as follows:

Fig. 3 Relationship between $\mathcal{R}_\theta^\beta$ and β . When the minimum similarity $s_{\min}$ is given, i.e., $ICS(x, y, \beta) \geq s_{\min}$, the larger β is, the smaller will be $\mathcal{R}_\theta^\beta$.

Fig. 4 Determination of β and its effect given $s_{\min} = 0.5$ and $R_{\theta}^{\beta} = \pi/6$ corresponding to $s_{\min}$. (a) From $\text{ICS}(x, y_1, \beta) = \text{ICS}(x, y_2, \beta) = \text{ICS}(x, y_3, \beta)$, $\beta = 9.9969$ can be determined. (b) After β is determined, the values of k_M and θ_{xy} are both limited to the shaded area.

$$\begin{aligned} [\text{CS}_1(x, y)]^{\beta} &= \left[\frac{1}{2} (1 + \cos \theta_{xy}) \right]^{\beta} \\ &= \left[\frac{1}{2} \left(1 + \cos \frac{\pi}{6} \right) \right]^{\beta} = \frac{1}{2}. \end{aligned} \quad (13)$$

By solving Eq. (13), we can obtain $\beta = 9.9969$. Once β is determined, for a given minimum similarity $s_{\min} = 0.5$, θ_{xy} is limited to $\pi/6$ irrespective of the change in k_M , as shown in the shaded area in Fig. 4(b). In other words, when $\beta = 9.9969$, in the limit case, the ICS with $\theta_{xy} = 0$ and $k_M = 0.5$ is the same as the ICS with $\theta_{xy} = \pi/6$ and $k_M = 1$, i.e., $\text{ICS}(x, y_1, 9.9969) = \text{ICS}(x, y_2, 9.9969) = \text{ICS}(x, y_3, 9.9969)$ [as shown in Fig. 4(a)]; thus, this avoids extreme imbalance between the magnitude difference and the direction difference shown in Fig. 2. In this manner, when using the ICS to perform the segmentation of a TD, if the minimum similarity $s_{\min} = 0.5$ is given, the maximum directional angle difference among all vectors in the TD is restricted to $2R_{\theta}^{9.9969} = 2 \times (\pi/6) = \pi/3$, and the normalized maximum magnitude difference is limited to $[0.5, 1]$.

3.3 ICS-Based TDS and ICS-Based LTDS

It can be seen from Eqs. (2) and (6) that $\mathcal{F}[A(x), x]$ in the objective function of LTDS is similar to the term $\int_{\Omega(\tau)} [a(\tau) \cdot E(x)]^2 dx$ in the objective function of TDS; they essentially use the CS to measure whether the vector to be segmented is perpendicular to the dominant translation parameter $a(\Omega)$ of Ω . In the actual application process, when Ω is constantly evolving, the degree of its unit TD approximation becomes worse; the difference between the magnitudes of the vectors may also increase. In this case, if the magnitude difference is not equalized due to improper selection of μ , undersegmentation may occur, i.e., regions with vectors having small direction difference but large magnitude difference may be segmented into one region, as shown in Fig. 5. In an actual motion field, the difference between magnitude is often an important basis for distinguishing the foreground from the background; however, the error constraints in the objective function of both TDS and LTDS do not consider the magnitude difference between vectors, which causes the segmentation results to include more background in the motion regions.

Because ICS can measure the difference between both direction and magnitude of vectors, it is expected to improve

Fig. 5 Undersegmentation caused by the magnitude difference between foreground and background. In the left, the vectors of foreground (gray shaded region) and background (green shaded region) have no difference in direction, but they have obvious difference in magnitude. When the initial region Ω of the active contour algorithm is inside Ω_1 (as indicated by the red dashed line on the left), if the magnitude difference is not equalized due to improper selection of μ , the evolution result Ω' (the area within the red border on the right) will contain both the foreground and the background, i.e., $\Omega' = \Omega_1 \cup \Omega_2$.

the segmentation performance by improving the objective function of LTDS and TDS. The key is to use ICS to redefine the error constraint term in the objective function of LTDS and TDS.

Referring to Eq. (2), the objective function of the TDS redefined by ICS (ICS-TDS) is as follows:

$$\begin{aligned} J(\tau) &= \int_{\Omega(\tau)} (1 - \text{ICS}(\mathbf{b}(\tau), \mathbf{U}(\mathbf{x}), \beta))^2 d\mathbf{x} \\ &\quad - \mu \int_{\Omega(\tau)} \mathcal{G}(\|\mathbf{U}(\mathbf{x})\|) d\mathbf{x} \\ &= \int_{\Omega(\tau)} \left(1 - \left(\frac{1}{2} \left(1 + \frac{\mathbf{b}(\tau) \cdot \mathbf{U}(\mathbf{x})}{\|\mathbf{b}(\tau)\| \cdot \|\mathbf{U}(\mathbf{x})\|} \right) \right)^\beta \right. \\ &\quad \cdot \frac{\min(\|\mathbf{b}(\tau)\|, \|\mathbf{U}(\mathbf{x})\|)}{\max(\|\mathbf{b}(\tau)\|, \|\mathbf{U}(\mathbf{x})\|)} \left. \right)^2 d\mathbf{x} - \mu \int_{\Omega(\tau)} \mathcal{G}(\|\mathbf{U}(\mathbf{x})\|) d\mathbf{x}, \end{aligned} \quad (14)$$

where $\mathbf{b}(\Omega)$ is the same vector as the field line of the TD Ω . Similar to $\mathbf{a}(\Omega)$, $\mathbf{b}(\Omega)$ can also be used as the dominant translation parameter to represent Ω . However, unlike $\mathbf{a}(\Omega)$, because the magnitude of $\mathbf{b}(\Omega)$ is involved in the minimum and maximum calculation, its optimal estimation cannot be obtained by the quadratic programming algorithm. However, we can use the average vector $\bar{\mathbf{U}}_\Omega$ of all vectors in Ω to approximate $\mathbf{b}(\Omega)$ according to the definition of $\mathbf{b}(\Omega)$, i.e.,

$$\hat{\mathbf{b}}(\tau) = \bar{\mathbf{U}}_\Omega = \frac{1}{|\Omega(\tau)|} \int_{\Omega(\tau)} \mathbf{U}(\mathbf{x}) d\mathbf{x}. \quad (15)$$

Although $\mathbf{b}(\Omega)$ does not obtain its optimal estimation by Eq. (15), its acquisition process is greatly simplified

compared to $\mathbf{a}(\Omega)$; its time complexity is significantly reduced from $O(|\Omega|^3)$ to $O(|\Omega|)$.

After defining the objective function of ICS-TDS given by Eq. (14), referring to the localization method in Eq. (6), we can obtain the objective function of ICS-based LTDS (ICS-LTDS) by redefining $\mathcal{F}(\mathbf{A}(\mathbf{x}), \mathbf{x})$ with ICS:

$$\begin{aligned} J(\tau) &= \int_{\Omega(\tau)} \mathcal{F}_{\text{ICS}}[\mathbf{B}(\tau, \mathbf{x}), \mathbf{x}] d\mathbf{x} - \mu \int_{\Omega(\tau)} \mathcal{G}(\|\mathbf{U}(\mathbf{x})\|) d\mathbf{x}, \\ \mathcal{F}_{\text{ICS}}[\mathbf{B}(\tau, \mathbf{x}), \mathbf{x}] &= \frac{1}{|\ell(\mathbf{x})|} \int_{\ell(\mathbf{x})} (1 - \text{ICS}(\mathbf{B}(\tau, \mathbf{x}), \mathbf{U}(\hat{\mathbf{x}}), \beta))^2 d\hat{\mathbf{x}}, \\ \mathcal{G}(\|\mathbf{U}(\mathbf{x})\|) &= \frac{1}{1 + \left(\frac{\gamma}{\|\mathbf{U}(\mathbf{x})\|} \right)^2}, \end{aligned} \quad (16)$$

where $\mathcal{F}_{\text{ICS}}[\mathbf{B}(\tau, \mathbf{x}), \mathbf{x}]$ is the mean of the square of ICS-based errors in the neighborhood $\ell(\mathbf{x})$, which can be used to form the ICS-based LMCM; the definitions of the other parameters are the same as those in Eq. (6). It should be noted that the purpose of localization is to avoid the global parameter $\mathbf{a}(\Omega)$ or $\mathbf{b}(\Omega)$ from affecting the error estimation in the local range such that the estimation of $\mathbf{B}(\tau, \mathbf{x})$ in $\mathcal{F}_{\text{ICS}}[\mathbf{B}(\tau, \mathbf{x}), \mathbf{x}]$ should not be the average vector $\bar{\mathbf{U}}_{\Omega(\tau)}$ of $\Omega(\tau)$ defined by Eq. (15), but the average vector $\bar{\mathbf{U}}_{\ell(\mathbf{x})}$ of $\ell(\mathbf{x})$ as follows:

$$\hat{\mathbf{B}}(\tau, \mathbf{x}) = \bar{\mathbf{U}}_{\ell(\mathbf{x})} = \frac{1}{|\ell(\mathbf{x})|} \int_{\ell(\mathbf{x})} \mathbf{U}(\hat{\mathbf{x}}) d\hat{\mathbf{x}}. \quad (17)$$

Similar to TDS and LTDS, the minimization of the objective functions in Eqs. (14) and (16) can also be performed by the active contour algorithm. The evolution equation of ICS-TDS can be obtained by rewriting Eq. (4) as follows:

$$\begin{cases} \frac{\partial C}{\partial \tau}(\tau) = \left(\left(1 - \left(\frac{1}{2} \left(1 + \frac{\mathbf{b}(\tau) \cdot \mathbf{U}(\mathbf{x})}{\|\mathbf{b}(\tau)\| \cdot \|\mathbf{U}(\mathbf{x})\|} \right) \right)^\beta \cdot \frac{\min(\|\mathbf{b}(\tau)\|, \|\mathbf{U}(\mathbf{x})\|)}{\max(\|\mathbf{b}(\tau)\|, \|\mathbf{U}(\mathbf{x})\|)} \right)^2 - \mu \mathcal{G} \right) N_\tau \\ C(\tau = 0) = C_0 \end{cases} \quad (18)$$

As $\mathcal{F}_{\text{ICS}}[\mathbf{B}(\tau, \mathbf{x}), \mathbf{x}]$ is not an even function, opposite direction vectors can be well distinguished. Therefore, the evolution equation of the active contour algorithm corresponding to ICS-LTDS can be rewritten by referring to Eq. (7):

$$\begin{cases} \frac{\partial C}{\partial \tau}(\tau) = [\mathcal{F}_{\text{ICS}}[\mathbf{B}(\tau, \mathbf{x}), \mathbf{x}] - \mu \mathcal{G}(\|\mathbf{U}(\mathbf{x})\|)] \cdot N_\tau = [\mathcal{F}_{\text{ICS}}(\tau) - \mu \mathcal{G}] \cdot N_\tau \\ C(\tau = 0) = C_0 \end{cases} \quad (19)$$

3.4 Acquisition of Initial Regions

As we know, for the active contour algorithm, the selection of the initial regions has a significant impact on the final segmentation results. The segmentation results obtained by erroneous initial regions may be intertwined with the correct segmentation results; therefore, they cannot be eliminated. To obtain suitable initial regions of evolution, Wu and Wong²⁰ first selected the minimum point p in the LMCM and then calculated the mean $\bar{\mathcal{F}}_p$ of $\mathcal{F}$ and the mean optical flow magnitude $\bar{M}_p$ in the neighborhood of p . Only the neighborhood of a point with $\bar{\mathcal{F}}_p$ and $\bar{M}_p$ higher than given thresholds, which indicates that the motion consistency in its neighborhood is good enough and the possibility that its

neighborhood is the foreground is sufficiently high, can be considered as the initial region. However, when this initial region is used for medium- or low-density crowds, if the optical flow estimation error leads to some larger magnitudes of optical flow in the background, some erroneous initial regions may be selected in the background, as shown in Fig. 6(b). In Fig. 6(b), the initial regions in the red circle are all erroneous initial regions.

Although the magnitudes of optical flow in these erroneous regions are relatively large, the gray level changes in these background regions are actually small and can be easily estimated by the frame difference method. Therefore, in addition to $\bar{\mathcal{F}}_p$ and $\bar{M}_p$, we introduce the regional mean $\bar{D}_p$

Fig. 6 Example of the acquisition of initial regions. (a) The minimum (red points) extraction results in LMCM; (b) The initial regions extracted by $\bar{F}_p$ and $\bar{M}_p$, where the initial regions within the red circle are the erroneous initial regions in the background; and (c) The initial regions extracted after the introduction of $\bar{D}_p$, where all the erroneous initial regions in (b) are removed.

based on the frame difference for initial region determination. Only when $\bar{F}_p$, $\bar{M}_p$, and $\bar{D}_p$ are all higher than given thresholds, which are, respectively, called local error threshold $TH_{\bar{F}}$, local motion magnitude threshold $TH_{\bar{M}}$, and local frame difference magnitude threshold $TH_{\bar{D}}$, the neighborhood of the minimum point p can be confirmed as the initial region. The initial regions extracted after the introduction of $\bar{D}_p$ can be seen in Fig. 6(c). By comparing Figs. 6(b) and 6(c), it can be seen that most of the erroneous initial regions in the background are removed.

3.5 Algorithm Flow of the Proposed Method

As mentioned above, ICS-based LMCM is more suitable for segmenting high-density crowds with slow consistent motion because it is based on the local ICS error constraints that can well reflect the motion consistency between local regions. In contrast, ICS-based GMCM is based on the ICS error constraints between all vectors in the motion field and a given TD; therefore, it is more suitable for segmenting motion regions that are consistent with the given TD. If the initial regions contain all possible motion modes, the GMCMs corresponding to all the initial regions can be used to segment the moving crowd with medium- or low-density and complex motion modes.

In summary, the detailed flow of the proposed crowd motion segmentation algorithm based on ICS-LTDS/ICS-TDS is shown in Algorithm 1.

Algorithm 1 Proposed algorithm.

Input: Two consecutive frames in a video stream including moving crowd (I_1, I_2), crowding level (L , high or low), and thresholds of $\bar{F}_p$, $\bar{M}_p$ and $\bar{D}_p$ ($TH_{\bar{F}}$, $TH_{\bar{M}}$, $TH_{\bar{D}}$)

Output: Set of segments with directions ($\{\Omega_1, \Omega_2, \dots, \Omega_n\}$)

- (1) Calculating the optical flow field $\mathbf{U}$ between I_1 and I_2 by means of an optical flow extraction algorithm
- (2) Calculating the LMCM LM based on local ICS
- (3) Extracting the set of all the minimum points $P = \{p_1, p_2, \dots, p_k\}$ in LM

- (4) Calculating $\bar{F}_p$, $\bar{M}_p$, and $\bar{D}_p$ of each point in P
- (5) for $i = 1$ to k
- (6) if $\bar{F}_p(i) < TH_{\bar{F}} \parallel \bar{M}_p(i) < TH_{\bar{M}} \parallel \bar{D}_p(i) < TH_{\bar{D}}$
- (7) Delete p_i from P
- (8) end if
- (9) end for
- (10) Constructing a set of the initial regions $\{C_0^1, C_0^2, \dots, C_0^m\}$ by the undeleted points $\{p_1, p_2, \dots, p_m\}$ in P
- (11) if $L = \text{"high"}$ //crowd has high density and slow motion
- (12) for $j = 1$ to m
- (13) Region evolution through the level set algorithm based on ICS-LTDS-based evolution equation [Eq. (19)] with the initial region C_0^j in LM
- (14) end for
- (15) Regions with multiple motion directions $\{C^1, C^2, \dots, C^m\}$ are obtained
- (16) else //crowd has medium- or low-density and complex motion mode
- (17) for $j = 1$ to m
- (18) Calculating the GMCM GM_j corresponding to C_0^j based on ICS
- (19) Region evolution through the level set algorithm based on ICS-TDS based evolution equation [Eq. (18)] with the initial region C_0^j in GM_j
- (20) end for
- (21) Regions with multiple motion directions $\{C^1, C^2, \dots, C^m\}$ are obtained
- (22) end if
- (23) $\{\Omega_1, \Omega_2, \dots, \Omega_n\}$ is achieved by merging the regions with same direction in $\{C^1, C^2, \dots, C^m\}$.

4 Experiments and Discussion

In this study, two series of experiments were conducted to verify the performance of Algorithm 1. In the first series of experiments, the proposed ICS-LTDS algorithm was applied for the segmentation of high-density slow-moving crowds. In the second series of experiments, the proposed ICS-TDS algorithm was used to segment medium- or low-density crowds with complex motion modes. Most crowd video/image sequences used in the experiments were taken from two public datasets, i.e., UCF dataset²² and UCSD dataset,²⁸ and a small part from YouTube (<https://www.youtube.com/>). In addition, methods based on flow field models, specifically, the LTDS algorithm,²⁰ TDS algorithm,³⁶ LPD algorithm,²² streak flow-based algorithm,²⁴ and the prior knowledge-based trajectory tracking algorithm,¹⁶ and recently proposed typical nonflow-based methods, specifically, the CF algorithm³² and TCA,³⁴ were applied to the same dataset for performance comparison. During the experiments, all algorithms were run on the MATLAB 8.6 platform in the Windows 10 Pro environment. To evaluate the segmentation performance of the algorithm, the ground truth was set manually for some representative images. Meanwhile, on the basis of the ground truth, the numerical segmentation accuracy (SA) corresponding to different directions was given by means of the Jaccard similarity coefficient³⁹ among sets. The numerical SA is calculated as follows:

$$\begin{aligned} \text{SA}_i(\Omega, T) &= \frac{|\Omega_i \cap T_i|}{|\Omega_i \cup T_i|} \cdot 100\% \\ &= \frac{|\Omega_i \cap T_i|}{|\Omega_i| + |T_i| - |\Omega_i \cap T_i|} \cdot 100\%, \end{aligned} \quad (20)$$

where i represents the index of the regions corresponding to different motion directions, Ω_i is the segmentation result corresponding to i , and T_i is the ground truth region corresponding to i .

4.1 Segmentation of High-Density Slow-Moving Crowds

Algorithm 1 is applied for segmentation of high-density slow-moving crowds, as described in Sec. 3.5, in the LMCM by using the evolution equation [Eq. (19)] based on the objective function of ICS-LTDS ($\beta = 9.9969$). In all experiments, if the size of the entire image is mapped to $[0,1] \times [0,1]$, the size of the neighborhood $\ell(x)$ is 0.02×0.02 . The initial region of evolution is centered on the minimum of the LMCM and the size is 5 pixels $\times$ 5 pixels. The thresholds $\text{TH}_{\bar{\mathcal{F}}}$ and $\text{TH}_{\bar{\mathcal{M}}}$ are set to 0.06 and 2, respectively. When the threshold $\text{TH}_{\bar{\mathcal{D}}}$ is used for different image sequences, the set values are between 1.3 and 3.4. The window size for calculating $\bar{\mathcal{F}}_p$, $\bar{\mathcal{M}}_p$, and $\bar{\mathcal{D}}_p$ is 5 pixels $\times$ 5 pixels. The constant γ in $\mathcal{G}$ is set to 0.1. The constant μ is given by referring to LTDS.²⁰

$$\mu = \frac{\sin^2\left(\frac{\Delta\theta}{2}\right) - \bar{\mathcal{F}}}{\bar{\mathcal{G}}}, \quad (21)$$

where $\Delta\theta$ is the angle difference between the two regions to be segmented, which can be given first by $\mathcal{R}_\theta^\beta$ and adjusted according to the actual situation; $\bar{\mathcal{F}}$ and $\bar{\mathcal{G}}$ are the average values of $\mathcal{F}$ and $\mathcal{G}$ in the initial region. For the settings of

LTDS, parameters that are similar to those of ICS-LTDS, such as the size of $\ell(x)$, $\text{TH}_{\bar{\mathcal{F}}}$, $\text{TH}_{\bar{\mathcal{M}}}$, γ , and μ , are all set to the same values as in ICS-LTDS for performance comparison. The representative segmentation results of the flow field model-based methods (i.e., ICS-LTDS, LTDS, LPD, streak flow, and prior knowledge-based trajectory tracking) for four image sequences of high-density crowds with slow motion are shown in Fig. 7. The SA calculated according to Eq. (20) is shown in Table 1.

To compare the segmentation results of the proposed ICS-LTDS algorithm with those of the typical nonflow-based methods in scenarios with high-density slow-moving crowds, we selected pilgrims and marathon as test sequences, and applied the proposed ICS-LTDS, TCA,³⁴ and the CF-based algorithm³² to them. Representative segmentation results obtained from them are shown in Fig. 8 and the segmentation accuracies calculated according to Eq. (20) are shown in Table 2. For TCA, the following parameter settings were used: $p = 16$, $r = 40$ (10 for pilgrims), $\tau = 30$, $\Delta = 100$, $\delta = 1$, $\alpha = 0.1$, $\beta = 20$, and $\gamma = 150$; for the CF algorithm, $K = 20$, $z = 0.025$, the upper bound of Φ was taken as 1, and the threshold coefficient α was 0.5.

Similar to crowd flow in high-density slow-moving crowds, the proposed ICS-LTDS algorithm can also be applied to slow-moving traffic flow or mixed flow. We selected two image sequences for algorithm evaluation in this non-crowd scenario. In addition, flow-based ICS-LTDS, LTDS, and LPD, and nonflow-based TCA were selected for the evaluation and comparison in the same noncrowd scenario. Representative segmentation results for two image sequences of noncrowd scenes are shown in Fig. 9. The segmentation accuracies calculated according to Eq. (20) are shown in Table 3.

Further, we obtained the ground truth of 100 consecutive frames from the marathon sequence manually to perform evaluation on a sequence containing a slow-moving crowd. The segmentation results of the representative five consecutive frames of the three best-performing flow-based algorithms (i.e., the proposed ICS-LTDS, LTDS,²⁰ and LPD²²) and the nonflow-based CF algorithm³² are shown in Fig. 10. The highest, lowest, and average segmentation accuracies of all 100 frames in the entire sequence are shown in Table 4.

From the results in Figs. 7–10 and Tables 1–4, the following are clear: (1) As the principles of particle flow and streak flow are similar, both process motion information recorded for a period; therefore, the segmentation results of LPD and streak flow are similar. The proposed ICS-LTDS is an improvement of LTDS and so the segmentation results of these two methods are similar. (2) Although particle flow and streak flow are highly suitable for segmenting slow-moving targets, when the speed of the target has a downward trend and tends to stop, both methods will assume that the target has not moved for a period and thus abandon the extraction of the target. For example, in the highway traffic sequence (last row in Fig. 9), the traffic that decelerates and gradually stops because of congestion is not segmented by LPD. In contrast, as ICS-LTDS and LTDS are based on real-time optical flow, both segment these targets. (3) As the prior knowledge-based trajectory tracking method is used in the sparse optical flow field formed by feature points, it is only suitable for the motion segmentation of people in extremely crowded places. For a scene where the feature points are not

Fig. 7 Image sequences of high-density slow-moving crowds used for testing: (a) Mecca, (b) pilgrims, (c) marathon, and (d) kabba. Representative segmentation results (top to bottom): Ground truth images, segmentation results of the proposed ICS-LTDS, segmentation results of LTDS, segmentation results of LPD, segmentation results of streak flow, and segmentation results of prior knowledge-based trajectory tracking.

easily extracted or the crowd has looser distribution, missed segmentation is highly possible, as shown in the last row of Figs. 7(a)–7(d). (4) From the discussion in Sec. 3.3, as the objective function of LTDS only distinguishes the direction differences between vectors regardless of the magnitude differences, it is prone to produce undersegmentation results between foreground and background. This phenomenon is particularly evident in the segmentation results for the roundabout and highway traffic sequences in Fig. 9(c), and the marathon sequence in Fig. 10(c), which results in a lower SA than ICS-LTDS with improved objective function. (5) As TCA needs to record the complete trajectories of the key-points, it cannot cope with short-term changes of the moving regions. This makes it easy to produce undersegmentation

between the foreground and background, as shown in Figs. 8(c) and 9(e). (6) Except for the pilgrims and kabba sequence, the SA of ICS-LTDS is higher than other methods, indicating that the ICS-LTDS algorithm is more suitable for segmentation of high-density slow-moving crowds than existing flow-based and nonflow-based crowd segmentation algorithms.

4.2 Segmentation of Medium- or Low-Density Crowds with Complex Motion Modes

Algorithm 1 is applied to medium- or low-density crowds with complex motion modes, as described in Sec. 3.5, in the GMCMM corresponding to each initial region by using the evolution equation [Eq. (18)] based on the objective function

Table 1 Segmentation accuracies of the proposed ICS-LTDS, LTDS, LPD, streak flow, and prior knowledge-based trajectory tracking for the scenes in Fig. 7.

Test sequence	Proposed ICS-LTDS			LTDS ²⁰			LPD ²²			Streak flow ²⁴			Prior knowledge-based trajectory tracking ¹⁶		
	Red	Green	Average	Red	Green	Average	Red	Green	Average	Red	Green	Average	Red	Green	Average
Mecca	0.6104	—	0.6104	0.5922	—	0.5922	0.5846	—	0.5846	0.3623	—	0.3623	0.2376	—	0.2376
Pilgrims	0.4200	0.3693	0.3947	0.4086	0.3502	0.3794	0.4389	0.4229	0.4309	0.3978	0.4043	0.4011	0.3256	0.2489	0.2873
Marathon	0.5275	—	0.5275	0.4527	—	0.4527	0.4937	—	0.4937	0.4742	—	0.4742	0.4159	—	0.4159
Kabba	0.8535	0.6553	0.7544	0.8671	0.6322	0.7497	0.8903	0.7792	0.8348	0.4940	0.5769	0.5355	0.5881	0.3058	0.4470

Note: Bold values indicate the maximum of the average accuracy of the different methods for the same test sequence.

Fig. 8 Image sequences of high-density slow-moving crowds used for testing (top to bottom): pilgrims and marathon. Representative segmentation results: (a) ground truth images; (b) segmentation results of the proposed ICS-LTDS; (c) segmentation results of TCA; and (d) segmentation results of CF algorithm.

Table 2 Segmentation accuracies of the proposed ICS-LTDS, TCA, and CF algorithm for the scenes in Fig. 8.

Test sequence	Proposed ICS-LTDS			TCA ³⁴			CF algorithm ³²		
	Red	Green	Average	Red	Green	Average	Red	Green	Average
Pilgrims	0.4200	0.3693	0.3947	0.3029	0.2799	0.2914	0.3430	0.3216	0.3323
Marathon	0.5275	—	0.5275	0.4304	—	0.4304	0.4557	—	0.4557

Note: Bold values indicate the maximum of the average accuracy of the different methods for the same test sequence.

of ICS-TDS ($\beta = 9.9969$). In all experiments, the average vector of the initial region is taken as $\hat{b}(\tau)$ to calculate the GMCM. Other parameters, such as $TH_{\hat{F}}$, $TH_{\hat{M}}$, $TH_{\hat{D}}$, γ and μ , are all consistent with ICS-LTDS.

In LTDS, the evolution of the objective function is performed in LMCM, which is quite different from ICS-TDS. Therefore, in addition to LTDS, segmentation results of TDS, which also performs the evolution in the GMCM, are included for comparison. However, as the basic TDS proposed in Ref. 36 is only applicable to the unit vector field, it cannot be directly used in the actual scenes, as shown in Fig. 10. To make TDS available for crowd segmentation in the actual motion field, by referring to LTDS, we improve TDS by normalizing the motion vector and adding the foreground motion area constraint in the objective function of TDS:

$$J(\tau) = \int_{\Omega(\tau)} \left(\mathbf{a}(\tau) \cdot \frac{\mathbf{U}(\mathbf{x})}{\|\mathbf{U}(\mathbf{x})\|} \right)^2 d\mathbf{x} - \mu \int_{\Omega(\tau)} \mathcal{G}(\|\mathbf{U}(\mathbf{x})\|) d\mathbf{x}. \quad (22)$$

The TDS defined by Eq. (22) can be called the TDS with $\mathcal{G}$, abbreviated as TDS- $\mathcal{G}$. According to Eqs. (4) and (7), the evolution equation of TDS- $\mathcal{G}$ can be defined as follows:

$$\begin{cases} \frac{\partial C}{\partial \tau}(\tau) = \left(\left(\mathbf{a}(\tau) \cdot \frac{\mathbf{U}}{\|\mathbf{U}\|} \right)^2 - \mu \mathcal{G} \right) N_{\tau} \\ C(\tau = 0) = C_0 \end{cases} \quad (23)$$

In addition, among other methods based on flow field modeling, the particle flow-based LPD is selected for algorithm comparison. To facilitate the comparison among

Fig. 9 Image sequences of noncrowd scenes used for testing (top to bottom): roundabout and highway traffic. Representative segmentation results: (a) ground truth images; (b) segmentation results of the proposed ICS-LTDS; (c) segmentation results of LTDS; (d) segmentation results of LPD; and (e) segmentation results of TCA.

Table 3 SA of the proposed ICS-LTDS, LTDS, LPD, and TCA for the scenes in Fig. 8.

Test sequence	Proposed ICS-LTDS			LTDS ²⁰			LPD ²²			TCA ³⁴		
	Red	Green	Average	Red	Green	Average	Red	Green	Average	Red	Green	Average
Roundabout	0.3818	0.3846	0.3832	0.3072	0.3464	0.3268	0.3522	0.3948	0.3775	0.2628	0.3188	0.2908
Highway traffic	0.3831	0.4753	0.4292	0.2595	0.4090	0.3343	0.2938	0.3544	0.3241	0.1786	0.1553	0.1670

Note: Bold values indicate the maximum of the average accuracy of the different methods for the same test sequence.

Fig. 10 Segmentation results of the representative five consecutive frames (top to bottom) in the marathon sequence: (a) ground truth; (b) segmentation results of the proposed ICS-LTDS; (c) segmentation results of LTDS; (d) segmentation results of LPD; and (e) segmentation results of CF algorithm.

Table 4 Highest, lowest, and average segmentation accuracies of all 100 frames in the Marathon sequence.

Test sequence	Proposed ICS-LTDS			LTDS ²⁰			LPD ²²			CF algorithm ³²		
	Maximum	Minimum	Average	Maximum	Minimum	Average	Maximum	Minimum	Average	Maximum	Minimum	Average
Marathon	0.5864	0.5071	0.5462	0.5530	0.4459	0.4939	0.5603	0.4902	0.5226	0.4991	0.4501	0.4680

Note: Bold values indicate the maximum of the average accuracy of the different methods for the same test sequence.

Fig. 11 Representative segmentation result. Image sequences used for testing (top to bottom): crosswalk, UCSD pedestrians, Fudan pedestrians, and multidirectional pedestrians (MD pedestrians). (a) Ground truth images; (b) segmentation results of proposed ICS-TDS; (c) segmentation results of TDS-G; (d) segmentation results of LTDS; and (e) segmentation results of LPD.

algorithms, the common parameters of ICS-TDS, TDS-G, and LTDS are given the same values. The representative segmentation results of ICS-TDS, TDS-G, LTDS, and LPD applied to four image sequences are shown in Fig. 11, and the segmentation accuracies are shown in Table 5.

It should be noted that in a close-up scene containing multidirectional pedestrians as shown in the fourth row of Fig. 10, the term T_i in Eq. (20) is difficult to estimate due to the numerous moving directions of the pedestrians, mutual occlusion, and even differences in moving directions of the different parts of the same body, which makes it impossible to calculate the SA separately for different directions according to Eq. (20). Considering that the SA should simultaneously measure the similarity between both area and direction of the segment and the ground truths, on the basis of Eq. (20), we add the direction consistency measure between Ω_i and $\Omega_i \cap T$, so that the mean of the SA for different directions can be directly calculated without segmenting the ground truth T . The SA after adding the direction consistency measure is as follows:

$$A_{\text{avg}}(\Omega, T) = \frac{\sum_{\Omega_i \subset \Omega} |\Omega_i \cap T| \cdot \bar{S}_\varphi(\Omega_i, \Omega_i \cap T)}{|\Omega \cup T|} \cdot 100\% \\ = \frac{\sum_{\Omega_i \subset \Omega} |\Omega_i \cap T| \cdot \bar{S}_\varphi(\Omega_i, \Omega_i \cap T)}{|(\cup \Omega_i) \cup T|} \cdot 100\%, \quad (24)$$

where $\bar{S}_\varphi(\Omega_i, \Omega_i \cap T)$ is the direction consistency measure between Ω_i and $\Omega_i \cap T$, and its value should be mapped to $[0, 1]$ so that it can be calculated by means of CS as follows:

$$\bar{S}_\varphi(\Omega_i, \Omega_i \cap T) = \frac{1}{|\Omega_i \cap T|} \sum_{j \in (\Omega_i \cap T)} \frac{1}{2} [1 + \text{CS}[b_i, x_T(j)]], \quad (25)$$

where b_i is the average vector of the segmented region Ω_i and $x_T(j)$ is the vector of the ground truth T at location j . Substitute Eq. (25) into Eq. (24) to get the final form of A_{avg} , which is used to calculate the SA of the results in the fourth row in Fig. 10:

Table 5 SA of the proposed ICS-TDS, TDS- $\mathcal{G}$, LTDS, and LPD for the scenes in Fig. 11.

Test sequence	Proposed ICS-TDS			TDS- $\mathcal{G}$ ³⁶			LTDS ²⁰			LPD ²²		
	Red	Green	Average	Red	Green	Average	Red	Green	Average	Red	Green	Average
Crosswalk	0.3632	0.3710	0.3671	0.3016	0.3170	0.3093	0.2226	0.2825	0.2526	0.2081	0.2423	0.2252
UCSD pedestrians	0.4343	0.4717	0.4530	0.3360	0.4463	0.3912	0.3224	0.4676	0.3950	0.3639	0.5273	0.4456
Fudan pedestrians	0.4628	0.3191	0.3910	0.3573	0.2970	0.3272	0.3386	0.1422	0.2404	0.3884	0.3291	0.3588
MD pedestrians	—	—	0.5928	—	—	0.5554	—	—	0.4959	—	—	0.4049

Note: Bold values indicate the maximum of the average accuracy of the different methods for the same test sequence.

$$A_{\text{avg}}(\Omega, T) = \frac{\sum_{\Omega_i \in \Omega} \sum_{j \in (\Omega_i \cap T)} \frac{1}{2} [1 + \text{CS}[b_i, x_T(j)]]}{|(\cup \Omega_i) \cup T|} \cdot 100\%. \quad (26)$$

To compare the segmentation results of the proposed ICS-LTDS algorithm with those of the typical nonflow-based methods in crowds with medium- or low-density and complex motion modes, we selected the crosswalk and MD pedestrians as test sequences, and applied the proposed ICS-LTDS, TCA,³⁴ and the CF-based algorithm³² to them. Representative segmentation results are shown in Fig. 12 and the segmentation accuracies calculated according to Eq. (20) [Eq. (26) for MD pedestrians] are shown in Table 6. For

TCA, the parameter settings were as follows: $p = 16$, $r = 40$, $\tau = 30$, $\Delta = 100$, $\delta = 1$, $\alpha = 0.1$, $\beta = 20$ and $\gamma = 150$; for the CF algorithm, $K = 20$, $z = 0.025$, the upper bound of Φ was taken as 1, and the threshold coefficient α as 0.5.

Similar to Sec. 4.1, we obtained the ground truth of the 100 consecutive frames in the crosswalk sequence manually to evaluate sequences containing fast-moving crowds with complex motion modes. The segmentation results of the representative five consecutive frames of the three best-performing flow-based algorithms (i.e., the proposed ICS-TDS, TDS- $\mathcal{G}$,³⁶ and LPD²²) and the nonflow-based CF algorithm³² are shown in Fig. 13. The highest, lowest, and average segmentation accuracies of all 100 frames in the entire sequence are shown in Table 7.

(a) (b) (c) (d)

Fig. 12 Representative segmentation result. Image sequences used for testing (top to bottom): crosswalk and MD pedestrians. (a) Ground truth images; (b) segmentation results of proposed ICS-TDS; (c) segmentation results of TCA; and (d) segmentation results of CF algorithm.

Table 6 Segmentation accuracies of the proposed ICS-TDS, TCA, and CF algorithm for the scenes in Fig. 12.

Test sequence	Proposed ICS-TDS			TCA ³⁴			CF algorithm ³²		
	Red	Green	Average	Red	Green	Average	Red	Green	Average
Crosswalk	0.3632	0.3710	0.3671	0.1583	0.1706	0.1645	0.3072	0.3655	0.3364
MD pedestrians	—	—	0.5928	—	—	0.3924	—	—	0.3655

Note: Bold values indicate the maximum of the average accuracy of the different methods for the same test sequence.

Fig. 13 Segmentation results of the representative five consecutive frames (top to bottom) in the cross-walk sequence: (a) ground truth; (b) segmentation results of proposed ICS-TDS; (c) segmentation results of TDS-G; (d) segmentation results of LPD; and (e) segmentation results of CF algorithm.

From the results shown in Figs. 11–13 and Tables 5–7, the following is clear: (1) The particle flow-based LPD algorithm is not suitable for segmenting crowds with loose distribution, especially for segmenting crowds with complex motion modes. For example, in the segmentation results of crosswalk, as shown in the first row in Fig. 11(e) and the last row in Fig. 13(d), there are obvious missing or incorrect segmentations. Meanwhile, in the segmentation results of multidirectional pedestrians, several pedestrians with distinctly different motion directions are segmented in the same region, resulting in severe undersegmentation [as shown in Fig. 14, which is the zoomed-in version of the last row of Fig. 11(e)]. (2) As LMCM is based on the motion consistency between local regions, when pedestrians with little difference in motion direction occlude each other, their motion boundaries are not clear enough in LMCM, which makes LTDS prone to produce undersegmentation results for pedestrians that occlude each other, as shown in Fig. 15 [the zoomed-in version of the last row of Fig. 11(d)]. (3) Although TDS-G and LTDS have almost the same objective function, TDS-G evolves in the GMCM while LTDS evolves in the LMCM. Except for UCSD pedestrians, the

segmentation results of TDS-G are better than those of LTDS. It is indicated that the method based on the evolution in the GMCM is more suitable for segmenting medium- or low-density crowds with complex motion modes. As TDS-G has the same objective function as LTDS, which only distinguishes the vector direction but not the vector magnitude, the undersegmentation results between foreground and background caused by the objective function are still serious in the scene shown in Figs. 11 and 13; therefore, their SA is significantly lower than that of ICS-TDS. (4) As TCA cannot cope with short-term changes in the moving regions, when it is applied to crowd scenes containing complex motion patterns, the segmentation results still contain severe undersegmentation, as shown in Fig. 16(b) [zoomed-in version of the last row of Fig. 12(c)]. Meanwhile, as the coherent neighbor invariance mainly focuses on the pairwise motion consistency and ignores the motion difference among all the keypoints in the local region of the center point, when the CF algorithm is applied to crowd scenes containing complex motion patterns, undersegmentation results are easily produced, as shown in Fig. 16(c) [zoomed-in version of the last row of Fig. 12(d)]. (5) For the four scenes shown in

Table 7 Highest, lowest, and average segmentation accuracies of all 100 frames in the crosswalk sequence.

Test sequence	Proposed ICS-TDS						TDS- $\mathcal{G}^{36}$						LPD ²²						CF algorithm ³²					
	Maximum			Minimum			Maximum			Minimum			Maximum			Minimum			Maximum			Minimum		
	Red	Green	Average	Red	Green	Average	Red	Green	Average	Red	Green	Average	Red	Green	Average	Red	Green	Average	Red	Green	Average	Red	Green	Average
Crosswalk	0.374	0.373	0.345	0.313	0.319	0.319	0.336	0.320	0.205	0.251	0.272	0.286	0.183	0.155	0.249	0.335	0.373	0.296	0.283	0.283	0.319	0.283	0.283	0.319

Note: Bold values indicate the maximum of the average accuracy of the different methods for the same test sequence.

Fig. 14 Undersegmentation results produced by LPD: (a) ground truth and (b) segmentation results of LPD. In panels (a) and (b), the same ellipse corresponds to the same pedestrians. In the ground truth, the pedestrians in blue and yellow ellipses have at least two motion directions, but in the segmentation results of LPD, the pedestrians in each ellipse have the same motion direction.

Fig. 15 Undersegmentation results produced by LTDS: (a) Ground truth and (b) segmentation results of LTDS. In panels (a) and (b), the same ellipse corresponds to the same pedestrian(s). In the ground truth, the pedestrians in blue, yellow, and red ellipses have two motion directions, but in the segmentation results of LTDS, the pedestrians in each ellipse have the same motion direction.

Fig. 10, ICS-TDS is superior to other methods in terms of both segmentation effect and SA, indicating that the ICS-TDS algorithm is more suitable for segmenting medium- or low-density crowds with complex motion modes than existing flow-based and nonflow-based crowd motion segmentation algorithms.

5 Conclusion

In this study, a TDS model based on ICS was proposed; this model can be used to segment moving crowds with different crowding levels and complex motion modes. The method reconstructs the objective function of the TDS model by ICS so that the objective function can simultaneously measure the difference between both magnitude and direction of two vectors; thus, it can effectively avoid undersegmentation between the foreground and background due to the difference in magnitude. By switching between the “localization” and “globalization” modes of the objective function, i.e., performing the region evolution in LMCM or GMCM, the algorithm can be applied to segment crowds with different densities and motion states. In addition, by introducing the local motion magnitude threshold and local frame difference magnitude threshold simultaneously, the nonforeground regions can be excluded from the initial regions of evolution; thus, the undersegmentation and mis-segmentation results caused by erroneous initial regions are greatly reduced. The experimental results showed that for a variety of complex scenes containing moving crowds, the segmentation

Fig. 16 Undersegmentation results produced by TCA and CF algorithm. (a) Ground truth; (b) segmentation results of TCA; and (c) segmentation results of CF algorithm. In panels (a), (b), and (c), the same ellipse corresponds to the same pedestrian(s). In the ground truth, the pedestrians in blue, yellow, and purple ellipses have at least two motion directions, but in the segmentation results of TCA and CF algorithm, the pedestrians in each ellipse have the same motion direction.

performance and accuracy of the proposed ICS-LTDS/TDS-based crowd motion segmentation method are superior to existing flow field-based motion segmentation methods.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 61601156, 61701146, 61871164), and the Zhejiang Provincial Natural Science Foundation of China (Grant No. Q16F010019). The authors declare that they have no competing interests.

References

- H. Chaudhry et al., "Crowd region detection in outdoor scenes using color spaces," *Int. J. Model. Simul. Sci. Comput.* **9**(2), 1850012 (2018).
- D. Peleshko et al., "Design and implementation of visitors queue density analysis and registration method for retail video surveillance purposes," in *Proc. IEEE First Int. Conf. Data Stream Min. Process. (DSMP)*, Lviv, Ukraine, pp. 159–162 (2016).
- J. Luo et al., "Real-time people counting for indoor scenes," *Signal Process.* **124**, 27–35 (2016).
- M. Sajid, A. Hassan, and S. A. Khan, "Crowd counting using adaptive segmentation in a congregation," in *Proc. IEEE Int. Conf. Signal and Image Process. (ICSIP)*, Beijing, China, pp. 745–749 (2016).
- Z. Zhang, M. Wang, and X. Geng, "Crowd counting in public video surveillance by label distribution learning," *Neurocomputing* **166**, 151–163 (2015).
- X. Hu et al., "A novel approach for crowd video monitoring of subway platforms," *Optik* **124**(22), 5301–5306 (2013).
- B. Sirmacek and P. Reinartz, "Automatic crowd density and motion analysis in airborne image sequences based on a probabilistic framework," in *Proc. IEEE Int. Conf. Comput. Vision Workshops (ICCV Workshops)*, Barcelona, Spain, pp. 898–905 (2011).
- A. G. Abuarafeh, M. O. Khozum, and E. AbdRabou, "Real-time crowd monitoring using infrared thermal video sequences," *J. Am. Sci.* **8**(3), 133–140 (2012).
- T. Li et al., "Crowded scene analysis: a survey," *IEEE Trans. Circuits Syst. Video Technol.* **25**(3), 367–386 (2015).
- A. Dehghan et al., "Automatic detection and tracking of pedestrians in videos with various crowd densities," in *Pedestrian and Evacuation Dynamics*, U. Weidmann, U. Kirsch, and M. Schreckenberg, Eds., Springer, Zurich, Switzerland, pp. 3–19 (2012).
- H. Ullah et al., "Density independent hydrodynamics model for crowd coherency detection," *Neurocomputing* **242**, 28–39 (2017).
- D. Zhang et al., "High-density crowd behaviors segmentation based on dynamical systems," *Multimedia Syst.* **23**(5), 599–606 (2017).
- W. Ge, R. T. Collins, and R. B. Ruback, "Vision-based analysis of small groups in pedestrian crowds," *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(5), 1003–1016 (2012).
- L. Dong et al., "Fast crowd segmentation using shape indexing," in *Proc. 11th IEEE Int. Conf. Comput. Vision*, Rio de Janeiro, Brazil, pp. 1–8 (2007).
- M. Hu, S. Ali, and M. Shah, "Learning motion patterns in crowded scenes using motion flow field," in *Proc. Int. Conf. Pattern Recognit.*, Tampa, Florida, pp. 1–5 (2008).
- L. Zhang et al., "Crowd segmentation method based on trajectory tracking and prior knowledge learning," *Arabian J. Sci. Eng.* **43**, 7143–7152 (2018).
- D. Cremers and S. Soatto, "Motion competition: a variational approach to piecewise parametric motion segmentation," *Int. J. Comput. Vision* **62**(3), 249–265 (2005).
- T. Brox et al., "Colour, texture, and motion in level set based segmentation and tracking," *Image Vision Comput.* **28**(3), 376–390 (2010).
- A. S. Rao et al., "Crowd event detection on optical flow manifolds," *IEEE Trans. Cybern.* **46**(7), 1524–1537 (2016).
- S. Wu and H. S. Wong, "Crowd motion partitioning in a scattered motion field," *IEEE Trans. Syst. Man Cybern. Part B* **42**(5), 1443–1454 (2012).
- Y. Yuan, J. Fang, and Q. Wang, "Online anomaly detection in crowd scenes via structure analysis," *IEEE Trans. Cybern.* **45**(3), 548–561 (2015).
- S. Ali and M. Shah, "A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, Minneapolis, Minnesota, pp. 1–6 (2007).
- S. Mukherjee, D. Goswami, and S. Chatterjee, "A Lagrangian approach to modeling and analysis of a crowd dynamics," *IEEE Trans. Syst. Man Cybern.* **45**(6), 865–876 (2015).
- R. Mehran, B. E. Moore, and M. Shah, "A streakline representation of flow in crowded scenes," *Lect. Notes Comput. Sci.* **6313**, 439–452 (2010).
- X. Wang et al., "A high accuracy flow segmentation method in crowded scenes based on streakline," *Optik* **125**(3), 924–929 (2014).
- M. Gao et al., "Crowd motion segmentation and behavior recognition fusing streak flow and collectiveness," *Opt. Eng.* **57**(4), 043109 (2018).
- A. B. Chan and N. Vasconcelos, "Modeling, clustering, and segmenting video with mixtures of dynamic textures," *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(5), 909–926 (2008).
- V. Mahadevan et al., "Anomaly detection in crowded scenes," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, San Francisco, California, pp. 1975–1981 (2010).
- S. Duan, X. Wang, and X. Yu, "Crowded abnormal detection based on mixture of kernel dynamic texture," in *Proc. Int. Conf. Audio, Language and Image Process.*, Shanghai, China, pp. 931–936 (2014).
- Y. Ma, P. Cisar, and A. Kembhavi, "Motion segmentation and activity representation in crowds," *Int. J. Imaging Syst. Technol.* **19**(2), 80–90 (2009).
- B. Zhou, X. Tang, and X. Wang, "Coherent filtering: detecting coherent motions from crowd clutters," *Lect. Notes Comput. Sci.* **7573**, 857–871 (2012).
- B. Zhou et al., "Measuring crowd collectiveness," *IEEE Trans. Pattern Anal. Mach. Intell.* **36**(8), 1586–1599 (2014).
- Z. Fan et al., "Adaptive crowd segmentation based on coherent motion detection," *J. Signal Process. Syst.* **90**(12), 1651–1666 (2018).
- R. Sharma and T. Guha, "A trajectory clustering approach to crowd flow segmentation in videos," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, Arizona, pp. 1200–1204 (2016).
- G. Doretto et al., "Dynamic textures," *Int. J. Comput. Vision* **51**(2), 91–109 (2003).
- T. Roy et al., "Segmentation of a vector field: dominant parameter and shape optimization," *J. Math. Imaging Vision* **24**(2), 259–276 (2006).
- H. Li, W. Chen, and I.-F. Shen, "Segmentation of discrete vector fields," *IEEE Trans. Visual Comput. Graphics* **12**(3), 289–300 (2006).
- T. Brox et al., "High accuracy optical flow estimation based on a theory for warping," *Lect. Notes Comput. Sci.* **3024**, 25–36 (2004).
- P. Jaccard, "The distribution of the flora of the Alpine zone," *New Phytol.* **11**(2), 37–50 (1912).

Haibin Yu received his PhD degree in communication and information system from Zhejiang University in 2007. He is currently an associate professor in the College of Electronics and Information, Hangzhou Dianzi University. He was a visiting scholar at the Department of Neurosurgery, University of Pittsburgh, from November 2016 to November 2017. His research interests include computer vision, image processing, machine learning, and multisensor fusion.

Guoxiong Pan received his undergraduate degree in electronics and information engineering from Hangzhou Dianzi University in 2017. He is currently a graduate student in electronic science and technology at Hangzhou Dianzi University. His research interests include computer vision, image processing, and embedded systems.

Li Zhang received her PhD degree in communication and information system from Zhejiang University in 2008. She is currently a teacher in

the School of Computer Science and Technology, Hangzhou Dianzi University. Her research interests include computer vision, image processing, and deep learning.

Zhu Li received his PhD degree in Tokyo University of Agriculture and Technology. He is currently an associate professor in the College of Electronics and Information, Hangzhou Dianzi University. His research interests include computer vision, image processing, and machine learning.

Mian Pan received his PhD degree in electronic engineering from Xidian University in 2013. He is currently a lecturer in the College of Electronics and Information, Hangzhou Dianzi University. His main research interests include statistical signal processing, adaptive signal processing, computer vision, image processing, machine learning, and their applications in industry.