



# A Novel Intuitionistic Fuzzy Clustering Algorithm Based on Feature Selection for Multiple Object Tracking

Liang-qun Li<sup>1,2</sup> · Xiao-li Wang<sup>1,2</sup> · Zong-xiang Liu<sup>1,2</sup> · Wei-xin Xie<sup>1,2</sup>

Received: 24 October 2018/Revised: 11 March 2019/Accepted: 20 April 2019/Published online: 2 May 2019  
© Taiwan Fuzzy Systems Association 2019

**Abstract** In this paper, a novel intuitionistic fuzzy clustering algorithm based on feature selection (IFC-FS) for multiple object tracking is proposed. In the proposed algorithm, the neighborhood rough set is used to achieve the adaptive selection of the multiple object features of visual objects, which are applied to calculate the distance similarity measure between the objects and the observations. At the same time, in order to incorporate the local information of objects into the intuitionistic fuzzy clustering, the local information distances between objects and observations are estimated by using the optimal subpattern assignment metric based on the reference topology set, and a new intuitionistic fuzzy clustering based on maximum entropy principle is proposed by using the new similarity distance measure. Finally, the association probabilities among the objects and the observations are reconstructed by utilizing the intuitionistic fuzzy membership degrees. The experimental results show that the proposed algorithm can effectively improve the estimated accuracy and robustness of the association probabilities between the objects and the observations, and have the ability to track accurately multiple objects in the complex background and long-time occlusion environment.

**Keywords** Multiple object tracking · Intuitionistic fuzzy clustering · Feature selection · Neighborhood rough set

## 1 Introduction

Multiple object tracking (MOT), as one of the most important research topics in computer vision, is to estimate the states of objects, such as positions, sizes and identification (ID) of objects, etc. Nowadays, Multiple object tracking is widely applied in many fields, involving intelligent video surveillance [1, 2], virtual reality [3], human-computer interaction [4], traffic control [5], oceanography [6], intelligent robot [7], remote sensing [8], etc., and many multiple object tracking approaches have been proposed. However, there are many challenges unsolved for the multiple object tracking, for example, miss and false detections caused by occlusions, changing illumination conditions, the uncertainty of object motion, the deformation of pedestrians and so on. Recent years, with the significant improvement of object detection technique, especially, the object detection method based on the deep learning [9], the tracking-by-detection (TBD) framework has been widely applied for the visual multiple object tracking, which divides the multi-object tracking into two parts: data association and state estimate. Therefore, the detection based multi-object data association approaches have been widely studied [10, 11].

In order to solve the problem of data association of multiple object tracking, many algorithms have been proposed. According to the differences of the data association approaches, multiple object tracking approaches can be categorized into two classes: generative method [12–15] and discriminant method [16–19]. Tracking algorithm based on the generation method solves the data association problem by using the modeling of appearance, movement and other characteristics of the object and considers the similarity between the object and the observation in the tracking process. For example, Ross et al. [12] proposed a

---

✉ Liang-qun Li  
lqli@szu.edu.cn

<sup>1</sup> ATR Key Laboratory, Shenzhen University, Shenzhen 518060, People's Republic of China

<sup>2</sup> Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen University, Shenzhen 518060, People's Republic of China

tracking approach based on incrementally learned a low-dimensional subspace to adapt online to the changes in the appearance of the target. In their algorithm, based on incremental algorithms for principal component analysis, the model update is not only correctly updated the sample mean, but can also utilize a forgetting factor to fit older observations to improve overall tracking performance. Azab et al. [13] proposed an online object tracking-by-detection algorithm in particle filtering framework by using a stationary camera, which can utilize the fuzzy integral to incorporate the local binary pattern texture feature, the red-green-blue (RGB) color feature and the Sobel edge feature into the particle filtering to improve the performance of algorithm. Zhang et al. [14] proposed a particle filtering tracking algorithm based on sparse representation for the problems of appearance variations separately. In their approach, each object candidate defined by a particle, which can be linearly represented by the target and background templates with an additive representation error. Jang et al. [15] proposed a visual object tracking system which was tolerant to external imaging factors. Specifically, an integration of an online version of total-error-rate minimization-based projection network with an observation model of particle filter was proposed to effectively distinguish between the target object and the background. Mei et al. [16] proposed a robust visual tracking method by casting tracking as a sparse approximation problem in a particle filter framework. In this framework, occlusion, corruption and other challenging issues were addressed seamlessly through a set of trivial templates. In addition, the approaches based on correlation filter [17] and deep learning [18, 19] have shown excellent performance and attracted more attention.

Recently, with the increasing application of fuzzy set theory in multi-object tracking, it is shown that the intuitionistic fuzzy set (IFS) was very useful to describe vague and uncertain data [20]. For this reason, intuitionistic fuzzy set (IFS) has been widely applied to many research fields [21, 22], and many tracking algorithms based on intuitionistic fuzzy set are proposed [23–25]. In [22], by using intuitionistic fuzzy set, Chaira presented an intuitionistic fuzzy c-means clustering method through defining a new objective function based on intuitionistic fuzzy entropy. The proposed algorithm can incorporate the uncertainty information arising from the definition of the membership function by intuitionistic index. In [23], Li et al. proposed two intuitionistic fuzzy joint probabilistic data association filters (IF-JPDF1 and IF-JPDF2) were proposed for multiple target tracking. In their method, to incorporate uncertainty information of measurement, a new intuitionistic fuzzy clustering method is proposed based on intuitionistic fuzzy point operator, and the optimized intuitionistic fuzzy membership degrees were used to

reconstruct the association probabilities. In [24], Li et al. proposed an online video multi-object tracking algorithm based on intuitionistic fuzzy set, which uses the resulting intuitionistic fuzzy membership degree to reconstruct the association probability between object and observation. Li et al. [25] proposed a fuzzy logic data association approach by incorporating fuzzy logic rules for visual multi-object tracking, the association probabilities were allowed to be adjusted adaptively based on the conclusions of a set of fuzzy rules. Therefore, the proposed algorithm had the advantage that it did require no assumption of statistical models of measurement noise and of object dynamics.

In this paper, a novel intuitionistic fuzzy clustering algorithm based on feature selection and maximum entropy for multiple object tracking is proposed. The main contributions of this paper are summarized as follows: Firstly, since the statistical features of the object vary with the change of time and object position, to adaptively calculate the statistical feature similarity, the neighborhood rough set is employed to select object feature, eliminate the low confidence degree feature and improve the estimated accuracy of feature similarity. Secondly, to incorporate the local information of objects into the proposed algorithm, the optimal subpattern assignment (OSPA) distance based on the reference topology set is used to measure the local information distance between objects and observations. Thirdly, a novel intuitionistic fuzzy clustering method based on the maximum entropy principle and feature selection is proposed, and the optimized intuitionistic fuzzy membership degree is used to reconstruct the association probability matrix, which don't limit to the traditional probability statistics constraint and can real-time estimate the association probabilities with high accuracy.

The rest of this paper is organized as follows. Section 2 presents the proposed visual object tracking algorithm based maximum entropy intuitionistic fuzzy clustering. Section 3 describes the experiment results that compare the performances of all algorithms. Finally, Section 4 gives some conclusions.

## 2 The Proposed Multiple Object Tracking Algorithm

Fuzzy logic provides a framework and flexibility to couple human judgment with the standard mathematical, and can deal with engineering problems which are too complex or ill-defined to yield analytical solutions in a simple and robust way. In this section, a novel data association algorithm based on intuitionistic fuzzy clustering for visual multi-object tracking is proposed. The block diagram of the proposed data association algorithm is shown in Fig. 1. From Fig. 1, it shows that the proposed data association

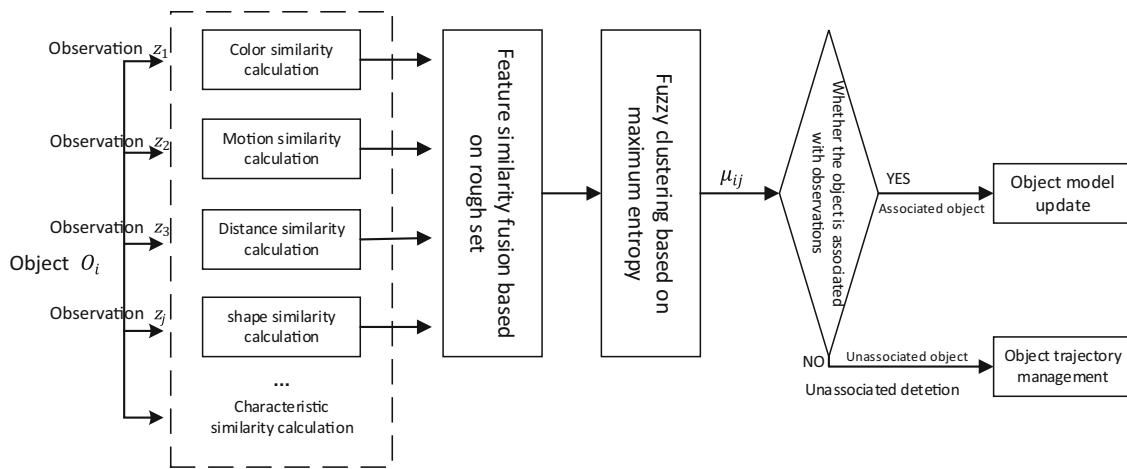


Fig. 1 Block diagram of the proposed data association algorithm

algorithm consists of the following steps: (1) Feature selection based on the neighbor rough set; (2) Fuzzy clustering based on maximum entropy principle; (3) Trajectory management; (4) Model update. In the proposed algorithm, we mainly focus on solving the problems of data association, while the track states are estimated using the Kalman filtering [2], and the appearance model update is similar to [24, 25].

### 2.1 Feature Selection Based on Rough Set

#### 2.1.1 Neighborhood Rough Set

For the visual object tracking, the feature information of the pedestrians often changes at different times and different locations. For example, the color information varies greatly due to the change of illumination conditions; the texture and shape information change with the swinging of hands, and so on. Generally, to visual multiple object tracking, the more number of the features we used, the higher performance we got. But the more number of object features are used, the computational load of the algorithm increases considerably. More importantly, it does not necessarily improve the data association accuracy. Actually, the use of more features may degrade the performance of the algorithm with the increasing of the computational time. Rough set theory plays an increasingly important role in dealing with uncertainty problem. Nowadays, rough set theory has been successfully applied to many fields. Rough set theory uses upper and lower approximation to represent the data uncertainty and can reduce the useless feature by the attribute reduction method. In the proposed algorithm, an attribute reduction method based on neighborhood rough set is proposed to select the object features and delete the redundant and useless features. The more robust attributes are retained to improve the accuracy and

efficiency of data association. The neighborhood rough set [26, 27] is defined as,

Suppose a finite non-empty set  $U = \{x_1, x_2, \dots, x_n\}$  in  $R$ ,  $D$  is the decision attribute, given arbitrary  $x_i \in U$ , the neighborhood  $\delta$  is defined as:

$$\delta(x_i) = \{x | x \in U, \Delta(x, x_i) \leq \delta\} \tag{1}$$

where  $\Delta(x, x_i)$  denotes a distance function, which is defined by

$$\Delta(x, x_i) = |x - x_i| \tag{2}$$

Suppose  $D_U$  denotes a partition set of the decision attribute  $D$ , and the upper and lower approximation of  $D_U$  are defined as:

$$D_U^{up} = \{x_i | \delta(x_i) \cap D_U \neq \emptyset, x_i \in U\} \tag{3}$$

$$D_U^{down} = \{x_i | \delta(x_i) \subseteq D_U, x_i \in U\} \tag{4}$$

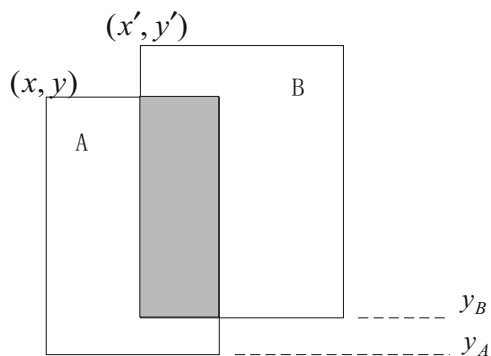
Given an information system  $IS = (U, D, V)$ , where  $U$  denotes the universe,  $V$  is the set of attribute values,  $D$  is the decision attribute, and  $D \subseteq V$ , the dependency degree of  $D$  to  $U$  is defined as:

$$\gamma_D = \frac{|D_U^{down}|}{|U|} \tag{5}$$

where  $|\cdot|$  is the cardinality of a set. The more dependent the decision attribute on the universe  $U$  is, the smaller the boundary domain of the decision attribute is; it shows that the decision attribute can be distinguished better, vice versa.

#### 2.1.2 Feature Selection

The mutual interference between adjacent objects is a key factor affecting the correct association rate of the objects, because these adjacent objects are very close in position and have similar appearance characteristics. So how to



**Fig. 2** Diagram of occlusion between object  $A$  and object  $B$

select the features of object is very important for visual multi-object tracking. In order to solve this problem, neighborhood rough sets are introduced to select the object features.

To construct the intuitionistic fuzzy clustering, multiple features are employed to estimate the similarity distance measure between object  $o_i, i = 1, 2, \dots, c$  and observation  $z_k, k = 1, 2, \dots, n$ , including the color feature  $f_1$ , edge feature  $f_2$ , texture feature  $f_3$ , shape feature  $f_4$ , distance feature  $f_5$ , motion direction feature  $f_6$  and overlap area  $f_7$ . Except the similarity measure of overlap area, the detail information of the similarity measures of other features can be found in [24]. The similarity of the overlap area between object  $o_i$  and observation  $z_k$  is defined as,

$$f_7(z_k, o_i) = \exp\left(-\frac{1 - w(z_k, o_i)}{2\sigma_w^2}\right) \quad (6)$$

where  $w(\cdot)$  denotes the overlap ratio of object and observation,  $\sigma_w^2$  denotes the variance of the overlap ratio of object and object. In order to estimate  $w(\cdot)$ , we suppose there have an object  $A$  and an observation  $B$ , and the overlap area between the rectangular box  $A$  and the rectangular box  $B$  is shown in Fig. 2.  $[x, y, w, d]$  and  $[x', y', w', h']$  denote the rectangular boxes  $A$  and  $B$ , respectively, where  $x, y$  denote the coordinates of the upper left corner of the rectangle, and  $w, d$  denote the width and the height of the rectangle. In Fig. 2, the overlapping rectangle area [24] can be denoted as  $[x_{o_i}, y_{o_i}, w_{o_i}, h_{o_i}]$ ,

$$\begin{aligned} x_{o_i} &= \max(x, x') \\ y_{o_i} &= \max(y, y') \\ w_{o_i} &= \min(x + w, x' + w') - x_{o_i} \\ h_{o_i} &= \min(y + h, y' + h') - y_{o_i} \end{aligned} \quad (7)$$

From Eq. (7), the overlapping rectangle area equals to  $w_{o_i} \cdot h_{o_i}$ . If  $w_{o_i} \leq 0$  or  $h_{o_i} \leq 0$ , it means two rectangles are not overlapped, the area of overlapping rectangles is 0.

In Fig. 2, it shows that the rectangular box of the object  $A$  is occluded by the rectangular box of the object  $B$ . The

overlapping shadow denotes the occlusion regions of the object  $A$  and the object  $B$ ; the overlap ratio  $w(A, B)$  between the object  $A$  and the object  $B$  is defined as:

$$w(A, B) = \frac{s(A \cap B)}{\min(s(A), s(B))} \quad (8)$$

Seven features are used to measure the similarity between object and observation. Since the statistical features of the object vary with the time, the importance of the feature to the proposed algorithm is different. For example, when an object is occluded, the direction of motion and the overlap area are more robust than other features, such as color, edge, etc. In this paper, to reduce the impact of invalid features, the attribute reduction method of neighborhood rough sets is used to select the object features and realize the adaptive selection of useful features. The feature selection method is summarized as follows:

1. Calculate the similarity measure  $f_l(z_k, o_i)$  between the object  $O = \{o_i\}_{i=1}^c$  and all observations  $Z = \{z_k\}_{k=1}^n$  for each feature  $l = 1, 2, \dots, 7$  as a domain  $U$ .
2. The lower approximation  $D_l^{down}$  of decision attributes  $D$  can be calculated as  $D_l^{down} = \{z_k | f_l(z_k, o_i) \geq \delta, k = 1, 2, \dots, n, i = 1, 2, \dots, c\}$ , then the dependence degree  $\gamma_D$  of the decision attribute  $D$  on  $U$  can be computed by Eq. (5).
3. If  $\gamma_D \geq \tau_{cf}$ , where  $\tau_{cf}$  is a constant, the feature  $i$  will be selected as a validate object feature to calculate the feature similarity. In this paper,  $\delta$  and  $\tau_{cf}$  are set to 0.1 and 0.7, respectively.

Finally, assume that there are  $L$  feature attributes that are selected by using the feature selection method, the similarity measure between the object  $o$  and the observation  $z$  can be estimated by using  $L$  feature similarities  $\{f_l(z, o)\}_{l=1}^L$ , which can be calculated as follow:

$$s(z, o) = \sum_{l=1}^L \alpha_l f_l(z, o), \alpha_l = f_l(z, o) / \sum_{l=1}^L f_l(z, o) \quad (9)$$

## 2.2 Maximum Entropy Intuitionistic Fuzzy Data Association

### 2.2.1 Construction of Intuitionistic Fuzzy Set

The intuitionistic fuzzy set (IFS) is an extension of Zadeh's fuzzy set, which was first proposed by Atanassov [20]. In intuitionistic fuzzy set, both the membership  $\mu_A$  and the non-membership  $\nu_A$  are introduced, and for all intuitionistic fuzzy set, Atanassov defined a new intuitionistic (or hesitation) index  $\pi_A$  to describe the uncertainty information of each element. Generally, an intuitionistic fuzzy set  $A$  in  $U$  is defined as follows:

$$A = \{ \langle u, \mu_A(u), v_A(u) \rangle \mid u \in U \} \tag{10}$$

where  $\mu_A(u) : U \rightarrow [0, 1]$  and  $v_A(u) : U \rightarrow [0, 1]$ ,  $\mu_A(u)$  and  $v_A(u)$  are called degree of membership and non-membership with the condition  $0 \leq \mu_A(u) + v_A(u) \leq 1$ , respectively.  $U$  being the referential set or universe, which in our case will always be non-empty and finite ( $Card(U) = n$ ).

For each IFS  $A$ , if

$$\pi_A(u) = 1 - \mu_A(u) - v_A(u) \tag{11}$$

where  $\pi_A(u)$  is the intuitionistic index of the element  $u$  to the intuitionistic fuzzy set  $A$ . Especially, if  $\pi_A(u) = 0$ , then IFS  $A$  is reduced to a fuzzy set. Due to the intuitionistic index, the membership values lie in the interval  $[\mu_A(u), \mu_A(u) + \pi_A(u)]$ .

According to Eq. (18), the sum of the membership, non-membership and intuitionistic index equals to one. Currently, the membership degree of IFS can be obtained by using the traditional fuzzy approaches, such as fuzzy clustering, fuzzy inference, etc., but there is no specific theoretical guidance for the definition of non-membership degree and intuitionistic index. In general, the intuitionistic fuzzy generator or the traditional fuzzy complement function can be used to define the intuitionistic index. Three fuzzy complement functions are often used to approximate the intuitionistic fuzzy complement, such as Yager function  $\gamma_1(x)$  [22], Sugeno function  $\gamma_2(x)$  [28] and Dubey function  $\gamma_3(x)$  [28], which can be written as

$$\gamma_1(x) = (1 - x^w)^{1/w}, w > 0 \tag{12}$$

$$\gamma_2(x) = \frac{1 - x}{1 + \lambda x}, \lambda > 0 \tag{13}$$

$$\gamma_3(x) = (1 - x) * (e^{-\frac{x}{\sigma}})^{1/\alpha}, \alpha > 0, \sigma > 0 \tag{14}$$

where  $\gamma_i(0) = 1, \gamma_i(1) = 0, i = 1, 2, 3$ . According to the conclusions in [28], to the image information processing, the Dubey fuzzy complement function is more suitable to be used to construct the intuitionistic index, because the Dubey intuitionistic fuzzy complement index is symmetric, while the other two approaches are not symmetric. In this paper, we chose the Dubey fuzzy complement function to approximate the intuitionistic index. Thus, with the help of Dubey' intuitionistic fuzzy complement, IFS becomes:

$$A = \left\{ \langle u, \mu_A(u), (1 - \mu_A(u)) * \left( e^{-\frac{\mu_A(u)}{\sigma\alpha}} \right)^{1/\alpha} \rangle \mid u \in U \right\} \tag{15}$$

and the intuitionistic index is:

$$\pi_A(u) = 1 - \mu_A(u) - (1 - \mu_A(u)) * \left( e^{-\frac{\mu_A(u)}{\sigma\alpha}} \right)^{1/\alpha} \tag{16}$$

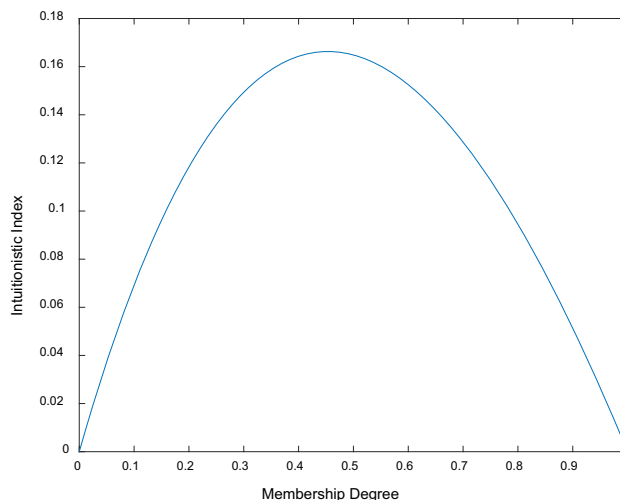


Fig. 3 Dubey' intuitionistic index

where  $\sigma$  is the standard deviation of membership values  $\mu_A(u)$ , which is in the range of 0.37–0.38.  $\alpha$  is positive constant. Figure 3 shows the relationship between the intuitionistic index and the membership degree  $\mu_A(u)$ .

### 2.2.2 Maximum Entropy Intuitionistic Fuzzy Clustering Based on Local Information

In the proposed intuitionistic fuzzy clustering algorithm, a modified objective function of maximum entropy fuzzy clustering [29] using IFS is constructed. Suppose dataset  $\mathbf{X} = \{x_k\}_{k=1}^n \in \mathbf{R}^s$  and cluster center  $\mathbf{V} = \{v_i\}_{i=1}^c$ , where  $v_i$  is  $i$ th cluster center. The objective function of maximum entropy fuzzy clustering [29] is defined as:

$$\begin{cases} \min_{\mu_{ik}} & \sum_{i=1}^c \sum_{k=1}^n \mu_{ik} d(x_k, v_i) \\ \max_{\mu_{ik}} & - \sum_{i=1}^c \sum_{k=1}^n \mu_{ik} \cdot \ln \mu_{ik} \\ & \sum_{i=1}^c \mu_{ik} = 1, \forall k \in 1, 2, \dots, n \end{cases} \tag{17}$$

where  $\mu_{ik}$  denotes the fuzzy membership degree of the data  $x_k$  belonging to the cluster center  $v_i$ ,  $d(x_k, v_i)$  denotes the distance between the data  $x_k$  and the cluster center  $v_i$ .

Using Lagrange multiplier method, the optimization objective function can be defined as:

$$\begin{aligned} J(\mathbf{U}, \mathbf{V}) = & - \sum_{k=1}^n \alpha_k \sum_{i=1}^c \mu_{ik} d(x_k, v_i) \\ & - \sum_{k=1}^n \sum_{i=1}^c \mu_{ik} \ln \mu_{ik} + \sum_{k=1}^n \lambda_k \left( \sum_{i=1}^c \mu_{ik} - 1 \right) \end{aligned} \tag{18}$$



Maximizing the objective function (18), the fuzzy membership degree can be obtained as follows:

$$\mu_{ik} = \frac{\exp(-\alpha_k(d(\mathbf{x}_k, \mathbf{v}_i)))}{\sum_{j=1}^c \exp(-\alpha_k(d(\mathbf{x}_k, \mathbf{v}_j)))} \tag{19}$$

where  $\alpha_k$  is called the discrimination factor, the designed details could be founded in [28].

In Eq. (19), only the feature similarity between the data samples is used to estimate the fuzzy membership degree and do not consider the local information of data sample, which may be very important for the visual object tracking. The main reason is that the precise of the membership degree will degrade because of the abrupt change of features caused by the occlusion, changes in lighting conditions, shadow effect, illumination variations, etc. On the contrary, there is no obvious change in the local information of object between two consecutive frames. So in order to incorporate the local information to improve the accuracy of fuzzy membership degree, a new intuitionistic maximum entropy fuzzy clustering is proposed.

Suppose that the object set is  $\mathbf{O}^{t-1} = [\mathbf{o}_1^{t-1}, \mathbf{o}_2^{t-1}, \dots, \mathbf{o}_c^{t-1}]$  at time  $t - 1$ ,  $\mathbf{O}^{t|t-1} = [\mathbf{o}_1^{t|t-1}, \mathbf{o}_2^{t|t-1}, \dots, \mathbf{o}_c^{t|t-1}]$  is the predicted object set at time  $t$ ,  $\mathbf{Z}^t = [z_1^t, z_2^t, \dots, z_n^t]$  is the observation set at time  $t$ . Considering the object predicted position as the cluster center, the objective function of fuzzy clustering is defined as.

$$\begin{cases} \min_{\mu_{ik}} & \sum_{i=1}^c \sum_{k=1}^n \mu_{ik} d(z_k^t, o_i^{t|t-1}) + \sum_{i=1}^c \sum_{k=1}^n \omega_{ik} \mu_{ik} L(z_k^t, o_i^{t|t-1}) \\ \max_{\mu_{ik}} & - \sum_{i=1}^c \sum_{k=1}^n \mu_{ik} \cdot \ln \mu_{ik} \\ & \sum_{i=1}^c \mu_{ik} = 1, \forall k \in 1, 2, \dots, n \end{cases} \tag{20}$$

where  $\mu_{ik}$  denotes the fuzzy membership degree between observation  $z_k^t$  and the clustering center  $o_i^{t|t-1}$ ,  $d(z_k^t, o_i^{t|t-1})$  denotes the distance between observation  $z_k^t$  and the cluster center  $o_i^{t|t-1}$ .  $L(z_k^t, o_i^{t|t-1})$  denotes the local information distance measure between observation  $z_k^t$  and the clustering center  $o_i^{t|t-1}$ .  $\omega_{ik}$  is the weight of local information. So the new optimization objective function can be given as

$$\begin{aligned} J(\mathbf{U}, \mathbf{V}) = & - \sum_{k=1}^n \alpha_k \sum_{i=1}^c \left( \mu_{ik} (d(z_k^t, o_i^{t|t-1}) + \omega_{ik} L(z_k^t, o_i^{t|t-1})) \right) \\ & - \sum_{k=1}^n \sum_{i=1}^c \mu_{ik} \ln \mu_{ik} + \sum_{k=1}^n \lambda_k \left( \sum_{i=1}^c \mu_{ik} - 1 \right) \end{aligned} \tag{21}$$

where  $\alpha_k$  denotes the discriminative factor, it usually is set to [0.4 0.6]. To identify  $\mu_{ik}$ , by using Lagrange multiplier method, the first derivative of (21) with respect to parameter  $\mu_{ik}$  equals zero, and we can obtain

$$\begin{aligned} \frac{\partial J}{\partial \mu_{ik}} = & -\alpha_k \left( d(z_k^t, o_i^{t|t-1}) + \omega_{ik} L(z_k^t, o_i^{t|t-1}) \right) \\ & - (1 + \ln \mu_{ik}) + \lambda_k \\ = & 0 \end{aligned} \tag{22}$$

which results in

$$\mu_{ik} = \exp \left( -\alpha_k \left( d(z_k^t, o_i^{t|t-1}) + \omega_{ik} L(z_k^t, o_i^{t|t-1}) \right) - 1 + \lambda_{ik} \right) \tag{23}$$

Substituting  $\mu_{ik}$  in (13) with (16), it follows that

$$\begin{aligned} \sum_{i=1}^c \mu_{ik} = & \sum_{i=1}^c \exp \left( -\alpha_k \left( d(z_k^t, o_i^{t|t-1}) + \omega_{ik} L(z_k^t, o_i^{t|t-1}) \right) \right. \\ & \left. - 1 + \lambda_{ik} \right) = 1 \end{aligned} \tag{24}$$

$$\exp(-1 + \lambda_{ik}) = \frac{1}{\sum_{j=1}^c \exp \left( -\alpha_k \left( d(z_k^t, o_j^{t|t-1}) + \omega_{jk} L(z_k^t, o_j^{t|t-1}) \right) \right)} \tag{25}$$

Then substituting (25) into (23), we can obtain

$$\mu_{ik} = \frac{\exp \left( -\alpha_k \left( d(z_k^t, o_i^{t|t-1}) + \omega_{ik} L(z_k^t, o_i^{t|t-1}) \right) \right)}{\sum_{j=1}^c \exp \left( -\alpha_k \left( d(z_k^t, o_j^{t|t-1}) + \omega_{jk} L(z_k^t, o_j^{t|t-1}) \right) \right)} \tag{26}$$

In order to incorporate the property of intuitionistic fuzzy set, the fuzzy membership degree is extended from the traditional fuzzy set to the intuitionistic fuzzy set by introducing the intuitionistic index [22], a new intuitionistic fuzzy membership degree is given

$$\mu_{ik}^* = \mu_{ik} + \varphi \pi_{ik} \tag{27}$$

where  $\mu_{ik}^*$  is the new intuitionistic fuzzy membership degree,  $\pi_{ik}$  denotes the intuitionistic index,  $\varphi \in [0, 1]$  denotes the constant scale factor. According to the definition in (16), the intuitionistic index  $\pi_{ik}$  is given by formula (28).

$$\pi_{ik} = 1 - \mu_{ik} - (1 - \mu_{ik}) \cdot \left( e^{\frac{-\mu_{ik}}{\sigma}} \right)^{1/\alpha} \tag{28}$$

where  $\alpha$  is a positive constant,  $\sigma$  is the standard deviation of fuzzy membership values  $\mu_{ik}$ , in this paper,  $\alpha$  and  $\sigma$  are set to 5, 0.36, respectively. Then, the intuitionistic fuzzy membership degree is normalized and obtained as follows:

$$\bar{\mu}_{ik}^* = \frac{\mu_{ik}^*}{\sum_{l=1}^c \mu_{lk}^*} \tag{29}$$

### 2.2.3 Approximate Calculation of Local Information Distance Measure

In the proposed fuzzy clustering algorithm, the key issue remains is how to calculate the similarity distance  $d(z_k^t, o_i^{t-1})$  and the local information distance measure  $L(z_k^t, o_i^{t-1})$ . According to the similarities defined above, the similarity distance  $d(z_k^t, o_i^{t-1})$  between the object  $o_i^{t-1}$  and the observation  $z_k^t$  can be defined as:

$$d(z_k^t, o_i^{t-1}) = 1 - s(z_k^t, o_i^{t-1}) \tag{30}$$

Moreover, to incorporate the local information of the object, the reference topology set (RET) [30] of the object  $o_i$  is given

$$\begin{aligned} \mathbf{RET}_i^o &= \{o_j | w(o_i, o_j) > \tau_2, j \neq i, j = 1, 2, \dots, n_i^o\} \\ &= \{t_{i,1}^o, t_{i,2}^o, \dots, t_{i,n_i^o}^o\} \end{aligned} \tag{31}$$

where  $w(o_i, o_j)$  is the ratio of overlap between the object  $o_i$  and the object  $o_j$ . The reference topology set of the object is a neighboring object set that is mutually occluded from the objects which the overlap ratio  $w(o_i, o_j)$  is greater than a certain threshold  $\tau_2$ , where  $\tau_2$  is called the topology radius. Similarly, the reference topology set (RET) of the object  $z_k$  can be defined as

$$\begin{aligned} \mathbf{RET}_k^z &= \{z_j | w(z_k, z_j) > \tau_2, j \neq k, j = 1, 2, \dots, n_k^z\} \\ &= \{t_{k,1}^z, t_{k,2}^z, \dots, t_{k,n_k^z}^z\} \end{aligned} \tag{32}$$

where  $n_i^o$  and  $n_k^z$  are the number of elements in  $\mathbf{RET}_i^o$  and  $\mathbf{RET}_k^z$ , respectively.

To measure the distance  $L(z_k^t, o_i^{t-1})$  of the object reference topology set  $\mathbf{RET}_i^o$  and the observation reference topology set  $\mathbf{RET}_k^z$ , the OSPA metric is utilized as follows:

$$\begin{aligned} L(z_k^t, o_i^{t-1}) &= L(\mathbf{RET}_k^z, \mathbf{RET}_i^o) = \frac{1}{n_i^o + n_k^z - q_h} \\ &\left( c_i^p (n_i^o + n_k^z - 2q_h) + \min \sum_{l=1}^{n_i^o} \sum_{r=1}^{n_k^z} [d(t_{i,l}^o, t_{k,r}^z)^p] \right)^{1/p} \end{aligned} \tag{33}$$

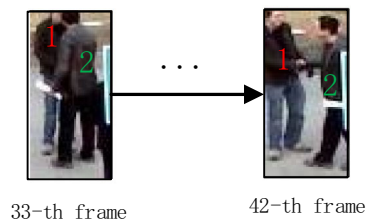


Fig. 4 Occlusion between the first object and the second object

where  $c_i$  is a penalty for unpaired items.  $q_h$  is the number of successful pair between  $\mathbf{RET}_i^o$  and  $\mathbf{RET}_k^z$ , which is given as:

$$q_h = \sum_{l=1}^{n_i^o} \sum_{j=1}^{n_k^z} h_{lj} \tag{34}$$

$$h_{lj} = \begin{cases} 1, & (t_{i,l}^o, t_{k,j}^z) \text{ successful pair} \\ 0, & \text{otherwise} \end{cases} \tag{35}$$

When the centroid distance between the elements  $t_{i,l}^o$  and  $t_{k,j}^z$  in two reference topological sets is the minimum and  $\|(x_o, y_o) - (x_z, y_z)\|_2 \leq T$ , it means that the elements of the two topology sets  $t_{i,l}^o$  and  $t_{k,j}^z$  are paired successfully.

Hence,  $n_i^o + n_k^z - 2q_h$  denotes the number of unsuccessful pairing,  $d(t_{i,l}^o, t_{k,j}^z)$  is distance between the centroid of elements  $(x_{i,l}^o, y_{i,l}^o)$  on the topology set.

$$\begin{aligned} d(t_{i,l}^o, t_{k,j}^z) &= \frac{1}{2} \left( \exp \left( - \frac{\|(x_{i,l}^o, y_{i,l}^o) - (x_{k,j}^z, y_{k,j}^z)\|_2^2}{2\sigma_D^2} \right) \right. \\ &\quad \left. + \exp \left( - \frac{1 - w(t_{i,l}^o, t_{k,j}^z)}{2\sigma_w^2} \right) \right) \end{aligned} \tag{36}$$

where  $(x_{i,l}^o, y_{i,l}^o)$  and  $(x_{k,r}^z, y_{k,r}^z)$  denote the centroid coordinates of the  $l$ -th element on the reference topological set  $\mathbf{RET}_i^o$  of the object  $o_i$  and the  $j$ -th element on the reference topological set  $\mathbf{RET}_k^z$  of observational  $z_k$ , respectively.  $\sigma_D^2$  denotes the distance variance.

As shown in Fig. 4, the reference topology set for the first object is the second object. At the frame 33, first object was severely occluded by second object, causing it to lose most of its appearance information, but the information of its reference topology set was relatively complete. There-

fore, the distance  $L(\mathbf{RET}_i^o, \mathbf{RET}_k^z)$  between the object and the observed reference topological set can be used to improve the association accuracy.

Considering the influence of the adjacent object to the object  $\mathbf{o}_i$ , the weight  $\omega_{ik}$  of local information is defined as:

$$\omega_{ik} = \begin{cases} \frac{\sum_{j=1}^{N_i} N_i(j)}{N_i} & N_i > 0, k = 1, 2, \dots, n_k^z \\ 0, & N_i = 0 \end{cases} \quad (37)$$

$$N_i(j) = \begin{cases} 1, & w(i, j) > \tau_2 \\ 0, & \text{otherwise} \end{cases} \quad (38)$$

where  $N_i$  denotes the number of adjacent objects of the object  $\mathbf{o}_i$ ,  $\tau_2$  is similar as Eq. (31, 32), in this paper, we will set  $\tau_2 = 0.6$ .

### 2.3 Model Update and Trajectory Management

Using the appearance model of the object at the current time and the object model of the previous moment is accumulated in a certain proportion as the object template for the subsequent frame. However, this method is only suitable for the case where the object deformation and appearance change are small. When the object is seriously occluded or deformed and the appearance changes, the object model tracked by the current frame may not be able to track the subsequent object very well. So the model update of object is very important, in this paper, the approach of model update is similar to the approach in [24].

Moreover, when the object enters the video scene, the corresponding tracking method needs to initialize the

object trajectory automatically. When the object leaves the video scene, the corresponding trajectory needs to be terminated automatically. Therefore, in order to better manage the object trajectory, the following two object trajectory management rules are given:

- (1) If the observation is not associated with temporary trajectories or existing trajectories after data association, the observation can be considered as a temporary trajectory that can be initialized by using the Kalman filter. At the same time, if the temporary trajectory is associated with  $\tau_{init}$  subsequent frames, the temporary trajectory is considered as a existing trajectory.
- (2) If the object trajectory is not associated with any observation and the centroid coordinates of the object trajectory are considered to be beyond the boundary, which mean the distance between the object and the video boundary is less than 3 pixel coordinates, the object trajectory is considered as the complete (ended) track. Moreover, if the object trajectory is not associated with any measurement more than  $\tau_{term}$  frames, it is also considered as the complete (ended) track.

According to the discussion in [25], in this paper, the threshold of starting  $\tau_{init}$  and terminating  $\tau_{term}$  is determined by the experience setting, which are set to  $\tau_{init} = 3$  and  $\tau_{term} = 8$ , respectively. Finally, the overall intuitionistic fuzzy clustering algorithm based on feature selection for multiple object tracking is described in Algorithm 1:



**Algorithm 1:** Intuitionistic fuzzy clustering algorithm based on feature selection for multiple object tracking (IFC-FS)

**Initialize:**  $\mathbf{O}^0 = [\mathbf{o}_1^0, \mathbf{o}_2^0, \dots, \mathbf{o}_c^0]$ ,  $\delta = 0.1, \tau_{cf} = 0.7, \tau_{ini} = 3, \tau_{term} = 8, \tau_a = 0.6$ , at  $t = 1$ .

**for**  $t \geq 1$

**Prediction**

**for**  $i = 1, 2, \dots, c$ , obtain the predicted value  $\mathbf{O}^{t-1} = [\mathbf{o}_1^{t-1}, \mathbf{o}_2^{t-1}, \dots, \mathbf{o}_c^{t-1}]$  by using Kalman filtering. **end**

**Feature selection and fusion**

**for**  $i = 1, 2, \dots, c, k = 1, 2, \dots, n, l = 1, 2, \dots, 7, L = 0$ .

- Calculate the similarity degrees of feature  $f_i(\mathbf{z}_k^t, \mathbf{o}_i^{t-1})$  by defined in Eq.(6) and Inference [9].
- Calculate the lower approximation  $D_l^{down} = \{z_k \mid f_i(\mathbf{z}_k^t, \mathbf{o}_i^{t-1}) \geq \delta, k = 1, 2, \dots, n, i = 1, 2, \dots, c\}$ .
- Compute the dependence degree  $\gamma_D$  by Eq.(5).
- If  $\gamma_D \geq \tau_{cf}$ , the feature  $i$  is valid, and  $L = L + 1$ .

**end**

**for**  $i = 1, 2, \dots, c, k = 1, 2, \dots, n, l = 1, 2, \dots, L$ , calculate the similarity measure  $S$  by Eq.(9), **end**

**Calculate data association matrix**

**for**  $k = 1, 2, \dots, n, i = 1, 2, \dots, c$ , calculate reference topology set distance  $L(\mathbf{z}_k^t, \mathbf{o}_i^{t-1})$  by Eq.(33). **end**

**for**  $k = 1, 2, \dots, n, i = 1, 2, \dots, c$ , estimate fuzzy membership degree  $\mu_{ik}^* = \mu_{ik} + \varphi\pi_{ik}$  by Eq.(34). **end**

**for**  $k = 1, 2, \dots, n, i = 1, 2, \dots, c$ , construct the data association matrix  $S = [s_{ik}^t]$

$$s_{ik}^t = \begin{cases} \bar{\mu}_{ik}^*, & \bar{\mu}_{ik}^* = \max_{s=1:n} \{\bar{\mu}_{is}^*\} \ \& \ \bar{\mu}_{ik}^* > \tau_a \\ 0, & \text{otherwise} \end{cases} \tag{39}$$

**end**

**Obtain the final assignment**

**while**  $s_{pq} \geq \tau_a$ , find by

$$s_{pq} = \max_{i \in \{1, 2, \dots, c_{t-1}\}} \left[ \max_{k \in \{1, 2, \dots, n\}} ([s_{ik}]_{c_{t-1} \times n}) \right] \tag{40}$$

- Associate the object  $p$  with the observation  $q$ , and set all elements of the row  $p$  and column  $q$  of matrix  $S$  to 0.

**Repeat until**  $s_{pq} < \tau_a$ .

**Object state update**

- Associated objects and observations: Using the Kalman filtering to predict and update the model and the state of the object based on the associated observations  $\{(\mathbf{o}_i, \mathbf{z}_k)\}$ .
- Unassociated observations: managed by using the trajectory management rule (1).
- Unassociated objects: managed by using the trajectory management rule (2).

**end**

Set  $t-1 \rightarrow t$ .

### 3 Experimental results

In order to evaluate the performance of the proposed algorithm, four algorithms are employed to compare with the proposed algorithm on the CLEAR MOT index, such as the TC\_ODAL algorithm [31], the RNN-LSTM algorithm [32], the ALEXTRAC algorithm [33] and the SiameseCNN algorithm [34]. In order to evaluate the performance of the proposed algorithm, in all experiments, five standard datasets are employed and all methods use the same detection results provided by the artificial channel feature (ACF) pedestrian detector [35].

#### 3.1 Datasets

In this section, the test dataset, PETS2009.S2L1, PETS2009.S2L2, AVG-TownCentre, TUD-crossing dataset and ETH-Sunnyday are employed. Where the PETS2009.S2L1 dataset and the PETS2009.S2L2 dataset are captured by cameras fixed outside with different angles of view in the same area. In the PETS09.S2L1 dataset, the number of pedestrian objects is less than that in the PETS09.S2L2 dataset. In the PETS09.S2L2 dataset, the pedestrians are crowded, the number of the objects is higher than that of the PETS09.S2L1 dataset, and there is a mutual occlusion between the objects for a long time. The AVG-TownCentre dataset is a video shot by a fixed camera in a busy shopping section. Its occlusion is more serious than other video dataset, and it is vulnerable to the influence of the clothing model in the window. The ETH-Sunnyday dataset is a video taken by a mobile camera. The background and object shape of ETH-Sunnyday video change strongly, which is easily affected by the change of light and shadow. Table 1 shows the detailed information of the used datasets.

For the quantitative evaluation, the CLEAR MOT metrics [24] are used, including:

- MOTA ( $\uparrow$ ): Multi-target tracking accuracy, it is defined as  $MOTA = 1 - (\sum_t (FP_t + FN_t + IDS_t)) / (\sum_t m_t)$ , where  $FP_t$ ,  $FN_t$ ,  $IDS_t$ , and  $m_t$  are the number of false positives, number of missed targets, number of

mismatches and number of total targets, respectively, at the time  $t$ .

- MOTP ( $\uparrow$ ): Multi-target tracking precision, average of the bounding box overlap over all tracked targets.
- IDS ( $\downarrow$ ): ID switch, the number of times that a tracked target changes its id.
- MT ( $\uparrow$ ): The ratio of mostly tracked trajectories, which are tracked for at least 80%.
- ML ( $\downarrow$ ): The ratio of mostly lost trajectories, which are tracked for less than 20%.
- FG ( $\downarrow$ ): Fragments, the number of times that a ground truth trajectory is interrupted.

Here, the symbol  $\uparrow$  means that higher scores indicate better results, and the symbol  $\downarrow$  means that lower scores indicate better results.

#### 3.2 Tracking Performance Comparison

##### 3.2.1 PETS.S2L1 Dataset

Figure 5 shows the tracking results of the IFC-FS algorithm on PETS.S2L1 dataset. It can be seen from Fig. 5 that in the frame 17, there is a serious occlusion between the object 1, 2, 3 and the traffic sign, but in frame 34, the object can still associate the three objects accurately. It shows that the IFC-FS can effectively solve the problem of mutual occlusion between objects and backgrounds. At frame 122, object 10 and object 13 move toward an opposite direction, and object 13 is heavily occluded by object 10, resulting that the color features, texture features and other appearance features are largely lost, but at frame 145, object 10 and object 13 can still be associated accurately. In mutual occlusion, the local information of the object plays a more important role and can effectively deal with the problem of long-term occlusion.

Table 2 gives the performance comparison of the proposed algorithm on the PETS2009.S2L1 dataset with that of the TC\_ODAL algorithm, the RRNN-LSTM algorithm and the ALEXTRAC algorithm. As can be seen, the IFC-FS algorithm is superior to the other three algorithms in MOTA, FN, FP, IDS and FG score. It shows that the IFC-FS algorithm can track objects correctly and effectively, and the intuitionistic fuzzy membership degree can effectively replace the association degree to associate the object with the observation and improve the association accuracy.

##### 3.2.2 PETS.S2L2 Dataset

Figure 6 shows the tracking results of the IFC-FS algorithm running on the dataset. It shows that object 12 is occluded by object 6, 41 and 48 at frame 67, because local information of the object is incorporated, which make the

**Table 1** Information of the used datasets

Name	FPS	Length	Resolution	Trajectories
PETS.S2L1	7	795	768 × 576	19
TUD-Stadtmitte	25	179	640 × 480	10
AVG-TownCentre	2.5	450	1920 × 1080	226
PETS.S2L2	7	436	786 × 576	42
TUD-crossing	25	201	640 × 480	13

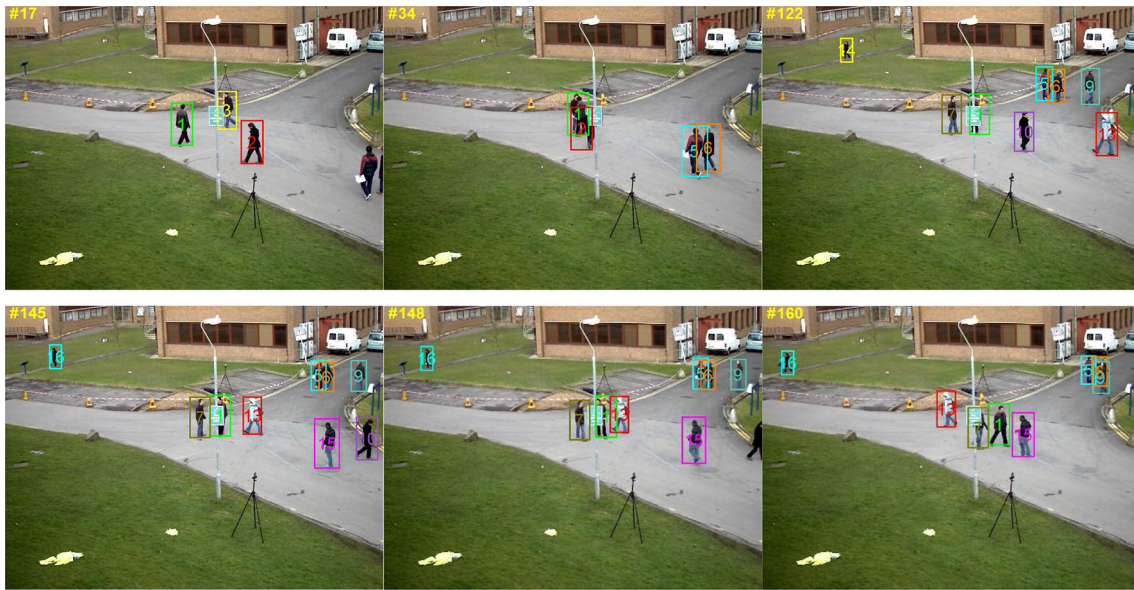


Fig. 5 Tracking results of the IFC-FS algorithm on PETS2009.S2L1 video sequence

Table 2 Performance comparison on PETS2009.S2L1

Method	MOTA (%↑)	MOTP (%↑)	MT (%↑)	ML (%↓)	FP (↓)	FN (↓)	IDS (↓)	FG (↓)
TC_ODAL	83.4	69.6	100	0.0	532	868	4	4
ALExTRAC	87.3	78.7	83.6	0.0	117	396	18	89
RNN-LSTM	67.7	71.4	78.9	0.0	648	603	193	152
IFC-FS	<b>88.6</b>	70.1	<b>94.3</b>	<b>0.0</b>	264	<b>291</b>	8	52

The bold values indicate the best performing tracker on each metric



Fig. 6 Tracking results of IFC-FS algorithm on PETS2009.S2L2



**Table 3** Performance comparison on PETS2009.S2L2 dataset

Method	MOTA (% $\uparrow$ )	MOTP (% $\uparrow$ )	MT (% $\uparrow$ )	ML (% $\downarrow$ )	FP ( $\downarrow$ )	FN ( $\downarrow$ )	IDS ( $\downarrow$ )	FG ( $\downarrow$ )
TC_ODAL	30.2	69.2	2.4	19.0	1074	5375	284	499
ALExTRAC	27.5	70.6	0.0	26.2	449	6153	385	359
RNN-LSTM	38.3	71.6	9.5	14.3	1016	4858	139	260
SiameseCNN	34.5	69.7	7.1	19.0	972	5364	282	424
IFC-FS	<b>55.1</b>	69.2	<b>17.9</b>	<b>10.5</b>	841	<b>3896</b>	168	321

The bold values indicate the best performing tracker on each metric

**Fig. 7** Tracking effect of IFC-FS method in AVG-TownCentre dataset**Table 4** Performance comparison on AVG-TownCentre dataset

Method	MOTA (% $\uparrow$ )	MOTP (% $\uparrow$ )	MT (% $\uparrow$ )	ML (% $\downarrow$ )	FP ( $\downarrow$ )	FN ( $\downarrow$ )	IDS ( $\downarrow$ )	FG ( $\downarrow$ )
TC_ODAL	1.6	69.0	0.0	71.7	899	6111	26	114
ALExTRAC	13.3	70.0	1.3	61.1	442	5637	118	152
RNN-LSTM	13.4	68.8	3.5	41.2	1206	4682	299	414
SiameseCNN	19.3	69.0	4.4	44.7	698	4927	142	289
IFC-FS	<b>36.8</b>	69.9	<b>17.9</b>	<b>38.4</b>	<b>353</b>	<b>4271</b>	85	219

The bold values indicate the best performing tracker on each metric

IFC-FS algorithm deal with the occlusion between the objects, so at frame 72, object 12 and its adjacent high-density pedestrians can still be associated correctly. In frame 162, object 32 is disturbed by object 82 until frame 171, no observations are updated for a long time, so the track of object 32 is deleted. By frame 185, when it reappeared, it had no previous track, so it was reinitialized to object 95.

To further verify the effectiveness of the proposed tracking algorithm, a comparative experiment is carried out on the PETS2009.S2L2 dataset with more dense objects and more severe occlusion. The experimental results are shown in Table 3. From Table 3, it can be seen that compared with the other three methods, the IFC-FS algorithm achieves the best performance in MOTA, FN, MT and ML scores. All of these show that the proposed algorithm can effectively implement the correct association

between the object and the observation. Moreover, the proposed algorithm achieves the second best performance in MOTP, FP and IDS scores.

### 3.2.3 AVG-TownCentre Dataset

Figure 7 shows an example of the tracking result of the IFC-FS algorithm in AVG-TownCentre video sequence. According to Fig. 7, the false observation 10 that appears in the window at frame 5, and in frame 9, the IFC-FS can delete the false observation in time according to the principle of track management. In frame 53, most of the objects in the scene can be tracked accurately by the IFC-FS algorithm. From frame 127 to 129, the IFC-FS algorithm can track most of the objects despite the mutual occlusion among the objects.



Fig. 8 Tracking results of the IFC-FS algorithm in TUD-crossing dataset

The quantitative results of the AVG-TownCentre dataset are shown in Table 4. From Table 4, we can see that the best MOTA, MT, ML, FP and FN score are realized by the IFC-FS algorithm, it means that the IFC-FS algorithm can more effectively and correctly associated the objects and better deal with the cases of mutual occlusion between the objects and the background. However, the MOTP score and the FG score are slightly worse than other algorithms. From the data of Table 2 to Table 4, it can be seen that IFC-FS algorithm achieved better performance on different datasets than other algorithms. The accuracy of the IFC-FS has

increased significantly in multi-object tracking, especially in the crowded and congested scenarios. It shows that the proposed algorithm by incorporating local information can better solve the association between the objects when they are occluded by each other.

### 3.2.4 Other Datasets

Figure 8 shows an example of the tracking result of IFC-FS algorithm in the TUD-crossing video sequence. TUD-crossing is a surveillance video of pedestrians crossing a



Fig. 9 Tracking results of the IFC-FS algorithm on ETH-Sunnyday dataset



highway. There are two types of people whose directions are completely opposite. When crossing the road, the intersection of different objects will affect the accuracy of association. From Fig. 8, it can be seen that object 1 and object 5 can be correctly associated with each other from frame 2 to frame 10, which shows that the IFC-FS algorithm can track the object correctly when occlusion occurs. In frame 46, object 3 and object 9 occluded by each other, which makes no observations associated with object 9. Finally, the object 9 is deleted from the object track array, but the object 3 can still be correlated accurately, which shows that the IFC-FS algorithm can effectively deal with the mutual occlusion between objects.

Figure 9 shows the tracking results of the IFC-FS algorithm in ETH-Sunnyday video sequences. Due to the camera movement and illumination, a false observation 14 appears at frame 60, which follows the camera and pedestrian movements, it is difficult to distinguish it from each other when associated, and it is not deleted until frame 89. In the process of moving the camera, the shape of the object changes at all times, but the IFC-FS algorithm can associate the object accurately when the shape of the object changes and the illumination changes. From frames 140 to 191, we can see that the object 22 becomes larger, and the width and height of the object box cannot fit the object itself. But in the IFC-FS algorithm, because the object model is updated, the updating rate of the object model can be controlled by increasing the updating coefficient, achieving a more accurate update of the object model.

## 4 Conclusion

In this paper, a novel maximum entropy intuitionistic fuzzy clustering algorithm based on feature selection is proposed for multiple object tracking in a single camera view. In the IFC-FS algorithm, in order to improve the estimated accuracy of the similarity of feature, the neighborhood rough set is used to adaptively select the suitable object feature; at the same time, in order to incorporate the local information of objects, the OSPA distance based on the reference topology feature is used to measure the local information distance between objects and observations. Finally, a novel objective function of the maximum entropy intuitionistic fuzzy clustering is constructed, and the optimize intuitionistic fuzzy membership degree is utilized to replace the association probability between the object and the observation. The experiment results show that the IFC-FS algorithm can effectively track multiple visual objects with high accuracy and robustness in the complex background and long-time occlusion environment, and the performance is better than that of existing

algorithms, such as the TC\_ODAL algorithm, the RRNN-LSTM algorithm and the ALExTRAC algorithm.

Although the proposed algorithm achieved a better results, there are also some challenges that may improve our proposed algorithm, such as the feature selection approach, how to make reasonable use of the local information of object, etc. So our future research will mainly focus on the feature selection by using the neighborhood rough set, improving the distance measure approach to incorporate the object local information and constructing suitable fuzzy optimize objective function to improve the performance of multiple object tracking.

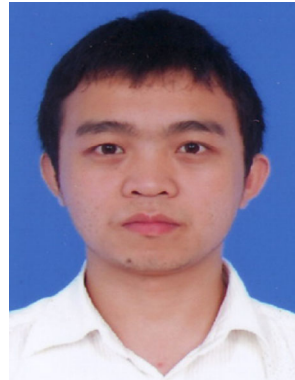
**Acknowledgements** This work was supported by the National Natural Science Foundation of China (61773267, 61301074), Science and Technology Program of Shenzhen (JCYJ20170302145519524, JCYJ20170818102503604).

## References

1. Kulbacki, M., Segen, J., Wojciechowski, S., et al.: Intelligent video monitoring system with the functionality of online recognition of people's behavior and interactions between people. In: Asian Conference on Intelligent Information and Database Systems, 2018, pp. 492–501
2. Cermeño, E., Pérez, A., Sigüenza, J.A.: Intelligent video surveillance beyond robust background modeling. *Expert Syst. Appl.* **91**(1), 138–149 (2018)
3. Han, D.T., Suhail, M., Ragan, E.D.: Evaluating remapped physical reach for hand interactions with passive haptics in virtual reality. *IEEE Trans. Vis. Comput. Graph.* **24**(4), 1467–1476 (2018)
4. Chen, C.: A study of the multi-feature gesture recognition tracking algorithm in VR-based human-computer interaction. *Softw. Eng.* **20**(12), 23–25 (2017)
5. Pop, S., Luculescu, M.C., Cristea, L., et al.: Improving communication between unmanned aerial vehicles and ground control station using antenna tracking systems. In: Auer, M., Zutin, D. (eds.) *Online engineering and internet of things*, vol. 22, pp. 532–539. Springer, Cham (2018)
6. Salucci, M., Robol, F., Anselmi, N., et al.: S-band spline-shaped aperture-stacked patch antenna for air traffic control applications. *IEEE Trans. Antennas Propag.* **66**(8), 4292–4297 (2018)
7. Cano, P., Ruiz-Del-Solar, J.: Robust tracking of soccer robots using random finite sets. *IEEE Intell. Syst.* **32**(6), 22–29 (2018)
8. Xu, X., Li, W., Ran, Q., et al.: Multisource remote sensing data classification based on convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* **56**, 1–13 (2018)
9. Han, J., Zhang, D., Cheng, G., et al.: Advanced deep-learning techniques for salient and category-specific object detection: a survey. *IEEE Signal Process. Mag.* **35**(1), 84–100 (2018)
10. Yang, H., Shao, L., Zheng, F., et al.: Recent advances and trends in visual tracking: a review. *Neurocomputing* **74**(18), 3823–3831 (2011)
11. Jang, S.I., Choi, K., Toh, K.A., et al.: Object tracking based on an online learning network with total error rate minimization. *Pattern Recogn.* **48**(1), 126–139 (2015)
12. Ross, D.A., Lim, J., Lin, R.S., et al.: Incremental learning for robust visual tracking. *Int. J. Comput. Vis.* **77**(1–3), 125–141 (2008)



13. Azab, M.M., Shedeed, H.A., Hussein, A.S.: New technique for online object tracking-by-detection in video. *IET Image Proc.* **8**(12), 794–803 (2014)
14. Zhang, S., Yao, H., Zhou, H., et al.: Robust visual tracking based on online learning sparse representation. *Neurocomputing* **100**(1), 31–40 (2013)
15. Jang, S.I., Choi, K., Toh, K.A., et al.: Object tracking based on an online learning network with total error rate minimization. *Pattern Recogn.* **48**(1), 126–139 (2015)
16. Mei, X., Ling, H.: Robust visual tracking using  $\ell_1$  minimization. In: *International Conference on Computer Vision*. IEEE, 2009, pp. 1436–1443
17. Gao, J., Zhang, T., Yang, X., et al.: Deep relative tracking. *IEEE Trans. Image Process.* **26**(4), 1845–1858 (2017)
18. Wang, L., Ouyang, W., Wang, X., et al.: Visual tracking with fully convolutional networks. In: *IEEE International Conference on Computer Vision*. IEEE Computer Society, 2015, pp. 3119–3127
19. Bertinetto, L., Valmadre, J., Henriques, J.F., et al.: Fully-convolutional siamese networks for object tracking. In: *European Conference on Computer Vision (ECCV 2016)*, 2016, pp. 850–865
20. Atanassov, K.: Intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **20**(1), 87–96 (1986)
21. Atanassov, K.: Intuitionistic fuzzy logics as tools for evaluation of data mining processes. *Knowl. Based Syst.* **80**, 122–130 (2015)
22. Chaira, T.: A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images. *Appl. Soft Comput.* **11**, 1711–1717 (2011)
23. Li, L.Q., Xie, W.X.: Intuitionistic fuzzy joint probabilistic data association filter and its application to multiobject tracking. *Signal Process.* **96**(3), 433–444 (2014)
24. Li, J., Xie, W.X., Li, L.Q.: Online visual multiple object tracking by intuitionistic fuzzy data association. *Int. J. Fuzzy Syst.* **19**(2), 1–12 (2016)
25. Li, L., Zhan, X., Liu, Z., Xie, W.: Fuzzy logic approach to visual multi-object tracking. *Neurocomputing* **281**(3), 139–151 (2018)
26. Pawlak, Z.: Rough sets. *Int. J. Comput. Inform. Sci.* **11**(5), 341–356 (1982)
27. Changzhong, W., Mingwen, S., Qiang, H., Yuhua, Q., Yali, Q.: Feature subset selection based on fuzzy neighborhood rough sets. *Knowl. Based Syst.* **111**(11), 173–179 (2016)
28. Todorova, L., Vassilev, P.: Algorithm for clustering data set represented by intuitionistic fuzzy estimates. *Int. J. Bioautom.* **14**(1), 61–68 (2010)
29. Liangqun, L., Hongbing, J., Xinbo, G.: Maximum entropy fuzzy clustering with application to real-time target tracking. *Signal Process.* **86**(11), 3432–3447 (2006)
30. Tian, W., Wang, Y., Shan, X., Yang, J.: Track-to-track association for biased data based on the reference topology feature. *IEEE Signal Process. Lett.* **21**(4), 449–453 (2014)
31. Bae, S., Yoon, K.: Robust online multi-object tracking based on tracklet confidence and online discriminative appearance learning. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 1218–1225
32. Son, J., Baek, M., Cho, M., Han, B.: Multi-object tracking with quadruplet convolutional neural networks. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5620–5629
33. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Read. Comput. Vis.* **24**(6), 726–740 (1987)
34. Milan, A., Rezatofghi, S., Dick, A., Reid, I., Schindler, K.: Online multi-target tracking using recurrent neural networks. In: *AAAI*, 2017, pp. 1–9
35. Dollár, P., Appel, R., Belongie, S.: Fast feature pyramids for object detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**(8), 1532–1545 (2014)



**Liang-qun Li** received the B.S. and M.S. degrees in signal and information processing from Xidian University, in 2002 and 2005, respectively, and Ph.D. degree in pattern recognition and intelligent system from the Xidian University in 2007, Xi'an, China. He was a Post-Doctoral researcher with Shenzhen University from 2007 to 2009. He visited the University of New Orleans (UNO) as a visiting research scientist from 2017 to 2018. He is currently a

professor of College of Information Engineering, Shenzhen University. His research interests are in signal processing, fuzzy sets, information fusion, target tracking. He has published 2 books and 70 papers in various journals and conference proceedings.



**Xiao-li Wang** is a Ph.D. student of school of information engineering, Shenzhen University. She research interests include multi-sensor information fusion, target tracking, etc.



**Zong-xiang Liu** received the Ph.D. degree in electronics information engineering from the Xidian University, Xi'an, China, in 1965. He is a professor of School of information engineering, Shenzhen University. His research interests include multi-sensor information fusion, multiple target tracking, and Random set theory, etc.



**Wei-xin Xie** is a professor and doctoral tutor of School of information engineering, Shenzhen University. His research interests include fuzzy information processing, multi-sensor information fusion, computer vision.