CrossMark

Online Visual Multiple Target Tracking by Intuitionistic Fuzzy Data Association

Li Jun[1] · Xie Wei-xin[1] · Li Liang-qun[1]

Abstract In this paper, a novel frame-by-frame data association algorithm based on intuitionistic fuzzy sets is proposed for online visual multiple target tracking. In the proposed algorithm, the association costs between targets and measurements are replaced by the intuitionistic fuzzy membership degrees which are obtained by a modified intuitionistic fuzzy c-means clustering algorithm. In addition, in order to mine useful information from the uncertain measurements, a new intuitionistic index is defined and the intuitionistic fuzzy point operator is applied to extract valuable information from the intuitionistic index. Experiments with challenging public datasets demonstrate that the proposed visual tracking algorithm improves tracking performance compared to other algorithms.

Keywords Data association · Visual multiple target tracking · Intuitionistic fuzzy point operator · Intuitionistic fuzzy clustering

1 Introduction

Visual multiple target tracking is to find the locations and sizes of multiple targets and maintaining the identities of targets in video sequences. Tracking of multiple targets is very important for many real applications and high-level analysis such as video surveillance, robotics, and activity analysis. Because of the significant improvements in object detection techniques in the recent years [1, 2], many researches in visual multiple target tracking have focused on the tracking-by-detection framework which transforms the multiple target tracking task into a data association problem. The targets are detected in individual frames and then the measurements (detection responses) are progressively associated into long trajectories.

According to the differences of the data association algorithms, multiple target tracking approaches can be categorized into two classes: offline multi-target tracking and online multi-target tracking. Offline multi-target tracking solves the data association problem over large temporal windows or even the whole video sequence by global optimization [3–8]. By exploiting information from future frames, these offline multiple target tracking algorithms can handle detection failures caused by occlusions, and show high accuracy and robustness. However, optimizing over large temporal windows leads to a significant temporal delay between the measurements and tracking results. Therefore, the offline multiple target tracking algorithms cannot be applied for real-time tracking.

Online multiple target tracking algorithms, which only consider information up to the current frame and locate the targets in the current frame, are more suitable for time critical online applications. Online multiple target tracking algorithms sequentially grow trajectories with online-provided measurements. The particle filter [9] and the joint probabilistic data association filter [10] are two classical approaches which are wildly used in online multiple target tracking systems. However, such methods often suffer from the exponentially increasing complexity in the number of tracked targets. Alternatively, a greedy scheme [11] or the Hungarian algorithm [12, 13] can be used to solve the data association problem more efficiently. Although these

✉ Li Liang-qun
lqli@szu.edu.cn

[1] ATR Key Lab of National Defense Technology, School of Information Engineering, Shenzhen University, 2-603 Box, Shenzhen 518060, Guangdong, People's Republic of China

Springer

online multiple target tracking approaches show improved performance, they tend to drift and to produce identity swap errors in complex scenes.

The main challenge for online visual multiple target tracking in unconstrained environments is that the measurements are often noisy and uncertain, such as missing detections caused by partial occlusions and false positive detections caused by changing illumination conditions. Thus, the resulting data association problem between the measurements and the targets in the online multiple target tracking systems is very difficult. Recently, some publications point out that utilizing fuzzy mathematics to model the uncertain information in the tracking system is helpful to improve the tracking performance [14–18],e.g., Fuzzy quadrature particle filter [14], Intuitionistic fuzzy joint probabilistic data association filter [15], fuzzy Kalman filter [16], and fuzzy coding histogram [18]. Inspired from [14], a novel online visual multi-target tracking method based on intuitionistic fuzzy data association is proposed in this paper. The main contributions of this paper are summarized as follows:

1) A novel online visual multi-target tracking method which integrates intuitionistic fuzzy data association with track management is proposed. Unlike [14], the proposed method does not use the algorithm structure of the joint probabilistic data association filter. In addition, automatically and correctly initialize or terminate a target track are enabled using track management.

2) The association costs between targets and measurements are replaced by the intuitionistic fuzzy membership degrees which can be obtained by a modified intuitionistic fuzzy clustering algorithm. The association costs in previous methods do not consider the uncertain information of the possible association pairs leading to the unsatisfying tracking results in complex scenes. However, by using the intuitionistic fuzzy membership degrees, the proposed method can deal with the uncertain and noisy measurements more effectively leading to the improvement in tracking accuracy. Furthermore, the modified intuitionistic fuzzy clustering algorithm is based on the multi-cue-based distance measure and a new intuitionistic index.

3) A new intuitionistic index is defined based on two different fuzzy c-means clustering. The defined intuitionistic index effectively models the uncertainty of the possible association pair. Combining with the intuitionistic fuzzy point operator which can extract useful information from the intuitionistic index, the uncertainty of the data association is effectively reduced leading to better tracking performance. In addition, the model of the intuitionistic index which is based on two different fuzzy c-means clustering has never been reported.

The rest of this paper is organized as follows. The proposed online visual multiple target tracking method based on intuitionistic fuzzy data association is presented in Sect. 2. Experiments are shown in Sect. 3. Conclusions are presented in Sect. 4.

2 Proposed Online Visual Multiple Target Tracking Method

In this section, a novel online visual multiple target tracking method based on intuitionistic fuzzy data association is proposed. The framework of the proposed tracking method is shown in Fig. 1.

In Fig. 1, an object detector based on aggregated channel features [2] is first applied to produce the set of measurements $\{z_i\}_{i=1,...,r}$. Each measurement z_i is represented by a bounding box around the target candidate region. Then, a frame-by-frame data association algorithm based on intuitionistic fuzzy sets is used to deal with the online-provided measurements which are imprecise and uncertain. The detailed derivation of the intuitionistic fuzzy sets-based data association algorithm is presented in Sect. 2.1. Through the intuitionistic fuzzy data association step, the set of correct association pairs $\{(o_n, z_k)\}_{n\in\{1,...,l\}; k\in\{1,...,r\}}$ is obtained. Then, the state of each existing target is updated and predicted by the Kalman filter. In the same time, track management is performed to enable automatically and correctly initialize or terminate an object track. The details of the track management are presented in Sect. 2.2. The overall online visual multi-target tracking algorithm is presented in Sect. 2.3.

2.1 Intuitionistic Fuzzy Data Association

Most previous online visual multi-target tracking approaches use Hungarian algorithm to solve the data association problem. In the Hungarian algorithm, the optimal association pairs are determined by minimizing the total association cost. However, the association costs in previous methods do not consider the uncertain information of the possible association pairs. In complex scenes, the measurements are often inaccurate, especially when many targets are close to each other. In this case, the association cost of an incorrect association pair might be lower than the association cost of the correct association pair. Therefore, the resulting association pairs are actually not the best association pairs. In order to deal with the noisy and uncertain measurements, intuitionistic fuzzy sets are used
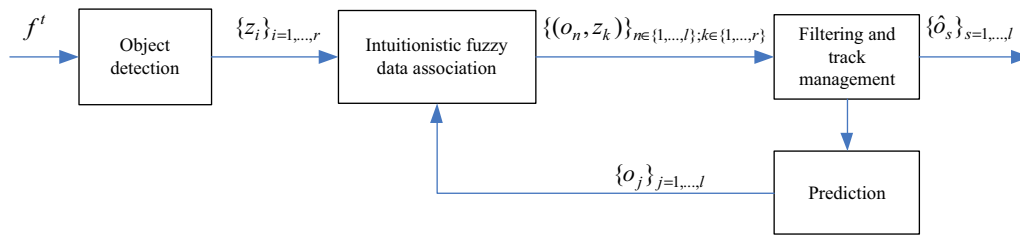
Fig. 1 The framework of the proposed online visual multi-target tracking method

to model the uncertainty of the possible association pairs, and the association costs are replaced by the intuitionistic fuzzy membership degrees.

2.1.1 Intuitionistic Fuzzy Point Operator

In order to represent the imprecise and imperfect information, Atanassov [19] extends fuzzy set to intuitionistic fuzzy set. An intuitionistic fuzzy set A in X is defined as follows:

$$A = \{ <x, \mu_A(x), v_A(x) > |x \in X\} \tag{1}$$

with the condition

$$0 \le \mu_A(x) + v_A(x) \le 1 \ \ \forall x \in X$$

where $\mu_A : X \to [0, 1]$ and $v_A : X \to [0, 1]$, μ_A represents the degree of membership, v_A represents the degree of non-membership of element $x \in X$ to the set A.

If

$$\pi_A(x) = 1 - \mu_A(x) - v_A(x) \tag{2}$$

then $\pi_A(x)$ is the degree of uncertainty of the membership of element $x \in X$ to the set A, which is called the intuitionistic index (degree of hesitancy). If $\pi_A(x) = 0$, then the intuitionistic fuzzy set is reduced to the ordinary fuzzy set.

In order to mine useful information from the intuitionistic index, the intuitionistic fuzzy point operator proposed in [20] is applied. It is defined as follows:

Definition 1 For each element $x \in X$, take $\alpha_x, \beta_x \in [0, 1]$, and $\alpha_x + \beta_x \le 1$, the intuitionistic fuzzy point operator $F_{\alpha_x, \beta_x} : IFS(X) \to IFS(X)$ is defined as follows:

$$F_{\alpha_x, \beta_x}(A) = \{ <x, \mu_A(x) + \alpha_x \pi_A(x), v_A(x) + \beta_x \pi_A(x) > |x \in X\}. \tag{3}$$

In [20], the author proves that the intuitionistic fuzzy point operator F_{α_x, β_x} has the following properties:

Theorem 1 Let $A, B \in IFS(X), x \in X, \alpha_x, \beta_x \in [0, 1]$, and $\alpha_x + \beta_x \le 1$,

(1) if $A \in FS(X)$,then $F_{\alpha_x, \beta_x}(A) = A$,

(2) $(F_{\alpha_x, \beta_x}(A^c))^c = F_{\alpha_x, \beta_x}(A)$,

(3) if $\mu_A(x) + v_A(x) \ne 0$ and $\alpha_x = \frac{\mu_A(x)}{\mu_A(x) + v_A(x)}$, $\beta_x = \frac{v_A(x)}{\mu_A(x) + v_A(x)}$ for each element $x \in X$, then $F_{\alpha_x, \beta_x}(A) = \{ <x, \alpha_x, \beta_x > |x \in X\}$ and $F_{\alpha_x, \beta_x}(A) \in FS(X)$.

Obviously, the intuitionistic fuzzy point operator F_{α_x, β_x} transforms the intuitionistic fuzzy set A to a new intuitionistic fuzzy set $F_{\alpha_x, \beta_x}(A)$ with intuitionistic index $\pi_{F_{\alpha_x, \beta_x}(A)}(x)$:

$$\pi_{F_{\alpha_x, \beta_x}(A)}(x) = 1 - (\mu_A(x) + \alpha_x \pi_A(x)) - (v_A(x) + \beta_x \pi_A(x))$$
$$= (1 - \alpha_x - \beta_x)\pi_A(x) \tag{4}$$

Therefore, for any $x \in X$, there is $\pi_{F_{\alpha_x, \beta_x}(A)}(x) \le \pi_A(x)$. If $A \in IFS(X)$, denotes $F^2_{\alpha_x, \beta_x}(A) = F_{\alpha_x, \beta_x}(F_{\alpha_x, \beta_x}(A))$, then

$$F^2_{\alpha_x, \beta_x}(A) = \{ <x, \mu_A(x) + \alpha_x \pi_A(x) + \alpha_x(1 - \alpha_x - \beta_x)\pi_A(x),$$
$$v_A(x) + \beta_x \pi_A(x) + \beta_x(1 - \alpha_x - \beta_x)\pi_A(x) > |x \in X\} \tag{5}$$

and

$$\pi_{F^2_{\alpha_x, \beta_x}(A)}(x) = 1 - [\mu_A(x) + \alpha_x \pi_A(x) + \alpha_x(1 - \alpha_x - \beta_x)\pi_A(x)]$$
$$- [v_A(x) + \beta_x \pi_A(x) + \beta_x(1 - \alpha_x - \beta_x)\pi_A(x)] .$$
$$= (1 - \alpha_x - \beta_x)^2 \pi_A(x) \tag{6}$$

Generally, for any positive integer n, there are

$$F^n_{\alpha_x, \beta_x}(A) = F_{\alpha_x, \beta_x}(F^{n-1}_{\alpha_x, \beta_x}(A))$$
$$= \{ <x, \mu_A(x) + \alpha_x \pi_A(x) \frac{1 - (1 - \alpha_x - \beta_x)^n}{\alpha_x + \beta_x},$$
$$= v_A(x) + \beta_x \pi_A(x) \frac{1 - (1 - \alpha_x - \beta_x)^n}{\alpha_x + \beta_x} > |x \in X\} \tag{7}$$

and

$$\pi_{F^n_{\alpha_x, \beta_x}(A)}(x) = (1 - \alpha_x - \beta_x)^n \pi_A(x) \tag{8}$$

where $\alpha_x + \beta_x \ne 0, \forall x \in X$.

If $x \in X, \alpha_x + \beta_x = 0$, namely $\alpha_x = 0$ and $\beta_x = 0$, then

$$\mu_{F^n_{\alpha_x,\beta_x}(A)}(x) = \mu_A(x)$$
$$v_{F^n_{\alpha_x,\beta_x}(A)}(x) = v_A(x) \tag{9}$$

Theorem 2 Let $A \in IFS(X)$, n is a positive integer. For each $x \in X$, $\alpha_x, \beta_x \in [0,1]$, and $\alpha_x + \beta_x \leq 1$, then, $\pi_{F^n_{\alpha_x,\beta_x}(A)}(x) \leq \pi_{F^{n-1}_{\alpha_x,\beta_x}(A)}(x) \leq \pi_A(x), \forall x \in X$.

Theorem 2 shows that the intuitionistic fuzzy point operator can reduce the degree of uncertainty of the membership. It indicates that the intuitionistic fuzzy point operator can mine useful information from the unknown information of the element.

2.1.2 Modified Intuitionistic Fuzzy Clustering

The modified intuitionistic fuzzy clustering algorithm combines the intuitionistic fuzzy point operator with a new intuitionistic index which is based on two different kinds of membership degrees. To calculate the membership degrees, fuzzy c-means clustering algorithm and its variants [21–23] can be used. Overview of the modified intuitionistic fuzzy clustering algorithm is shown in Fig. 2. In Fig. 2, the red boxes represent the targets denoted by $\{o_1, o_2, o_3\}$, and the yellow boxes represent the measurements denoted by $\{z_1, z_2, z_3, z_4, z_5\}$. By using the modified intuitionistic fuzzy clustering algorithm, the intuitionistic fuzzy membership degree matrix $[\eta]_{3\times5}$ can be obtained which reflects the uncertainty of all possible association pairs.

Suppose that there are the sets of states of the targets $O = \{o_1, \ldots, o_l\}$ predicted by Kalman filter and measurements $Z = \{z_1, \ldots, z_r\}$ produced by the object detector in the current frame. First, the following objective function is minimized,

$$\min_{\mathbf{U}} \left\{ J_m(\mathbf{U}) = \sum_{i=1}^{l} \sum_{k=1}^{r} u_{ik}^m g(o_i, z_k) \right\}, \quad o_i \in O, z_k \in Z$$
$$\sum_{i=1}^{l} u_{ik} = 1, \forall k, \quad m = 2 \tag{10}$$

It leads to

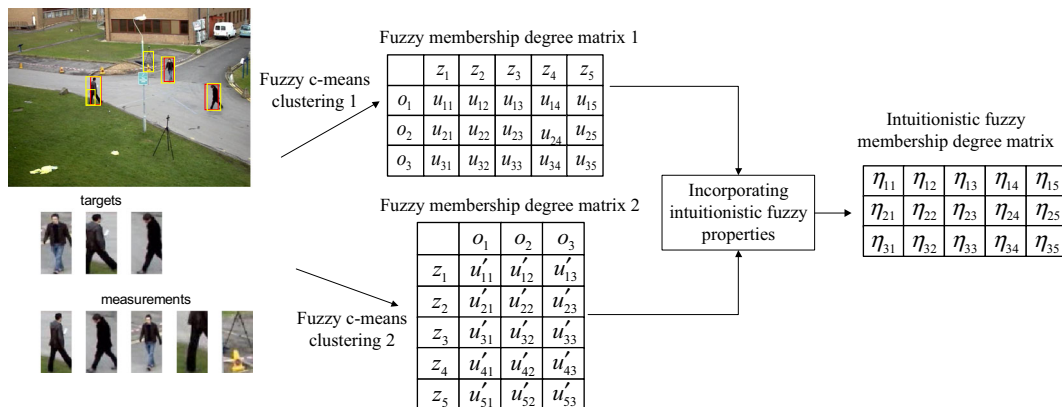$$u_{ik} = \left[\sum_{j=1}^{l} \left(\frac{g(o_i, z_k)}{g(o_j, z_k)} \right)^{\frac{2}{m-1}} \right]^{-1}$$
$$v_i = \frac{\sum_{k=1}^{r} (u_{ik})^m z_k}{\sum_{k=1}^{r} (u_{ik})^m}, i = 1, \ldots, l \tag{11}$$

where u_{ik} represents membership degree of the predicted state of the target o_i to the measurement z_k, v_i represents cluster center, and $g(\cdot, \cdot)$ represents the multi-cue-based distance measure defined by Eq. (21).

Then, the other objective function defined as follows is minimized,

$$\min_{\mathbf{U}'} \left\{ J_m(\mathbf{U}') = \sum_{i=1}^{l} \sum_{k=1}^{r} (u_{ki}')^m g(o_i, z_k) \right\}, \quad o_i \in O, z_k \in Z$$
$$\sum_{k=1}^{r} u_{ki}' = 1, \forall i, \quad m = 2 \tag{12}$$

It leads to

$$u_{ki}' = \left[\sum_{j=1}^{r} \left(\frac{g(o_i, z_k)}{g(o_i, z_j)} \right)^{\frac{2}{m-1}} \right]^{-1} \tag{13}$$

where u_{ki}' represents membership degree of the measurement z_k to the predicted state of the target o_i, and $g(\cdot, \cdot)$ represents the multi-cue-based distance measure defined by Eq. (21). Note that the constraints in Eq. (10) and Eq. (12) are different.

In order to incorporate intuitionistic fuzzy properties into the fuzzy membership degrees, the intuitionistic fuzzy point operator is applied. Given u_{ik} and u_{ki}' calculated with Eq. (11) and Eq. (13) respectively, a new intuitionistic fuzzy membership degree is defined as follows:

$$\eta_{ik} = u_{ik} \times u_{ki}' + \alpha_{z_k} \pi_{o_i}(z_k) \frac{1 - (1 - \alpha_{z_k} - \beta_{z_k})^n}{\alpha_{z_k} + \beta_{z_k}} \tag{14}$$

Fig. 2 Overview of the modified intuitionistic fuzzy clustering algorithm

where η_{ik} denotes the intuitionistic fuzzy membership degree of measurement z_k belonging to the target i, α_{z_k} and β_{z_k} denote the scale factor of the membership information and non-membership information in intuitionistic index $\pi_{o_i}(z_k)$, respectively.

The definition of the intuitionistic index $\pi_{o_i}(z_k)$ should reflect the uncertainty of the membership degree of the measurement z_k to the target i. Therefore, in the modified intuitionistic fuzzy clustering algorithm, the intuitionistic index is defined as follows:

$$\pi_{o_i}(z_k) = |u_{ik} - u'_{ki}| \tag{15}$$

The larger the difference between the membership degree of measurement z_k to target i and the membership degree of target i to measurement z_k, the more uncertain information the intuitionistic index contains. The choice of the parameters of α_{z_k} and β_{z_k} depends on the problem. In this paper, they are set to $\alpha_{z_k} = u_{ik} \times u'_{ki}$ and $\beta_{z_k} = 1 - \alpha_{z_k} - \pi_{o_i}(z_k)$, respectively. Equation (14) indicates that the intuitionistic fuzzy membership degree η_{ik} includes additional useful information extracted from the intuitionistic index, which also means that the uncertainty for assigning measurement z_k to target i is reduced.

2.1.3 Multi-cue-Based Distance Measure

In complex scenes, some targets are difficult to be discriminated by only relying on the appearance information because of their similar appearances. Therefore, in the modified intuitionistic fuzzy clustering algorithm, the information of multiple cues including appearance information, shape information, and motion information is exploited to construct the distance measure. First, the affinities between the states of the targets and the measurements based on five different cues which compensate each other to some extent are defined as follows.

1) The affinity between the predicted state of the target o and the measurement z in terms of spatial proximity is defined as follows:

$$f_1(o, z) = \exp\left(-\frac{\|(x_o, y_o) - (x_z, y_z)\|_2^2}{2\sigma_1^2 h_o}\right) \tag{16}$$

where x_o represents the x-coordinate of the center of the predicted state of the target o and y_o represents the y-coordinate of the center of the predicted state of the target o; x_z represents the x-coordinate of the center of the measurement z and y_z represents the y-coordinate of the center of the measurement z; h_o represents the height of the target o; parameter σ_1^2 represents the variance for the Euclidean distance between the two centers and is set to $\sigma_1^2 = 3$.

2) The affinity between the predicted state of the target o and the measurement z in terms of shape is defined as follows:

$$f_2(o, z) = \exp\left(-\frac{(h_o - h_z)^2}{2\sigma_2^2 \cdot (h_o)^2}\right) \tag{17}$$

where h_o represents the height of the target o and h_z represents the height of the measurement z, parameter σ_2^2 represents the variance for the height and is set to $\sigma_2^2 = 3$. Compared with the height information, the widths of the targets are less distinguishable. Therefore, the definition of affinity in terms of shape only relies on the height information.

3) It is assumed that the motion direction of the target does not change a lot between consecutive frames. The constant velocity motion model is used to model the motion of each target. Thus, the affinity between the state of the target o, which is predicted by the Kalman filter, and the measurement z in terms of velocity direction is defined as follows:

$$f_3(o, z) = \exp\left(\frac{|\arctan((y_z^t - y_o^{t-1})/(x_z^t - x_o^{t-1})) - \arctan(v_{y_o}^{t-1}/v_{x_o}^{t-1})|}{2\sigma_3^2}\right) \tag{18}$$

where y_z^t represents the y-coordinate of the center of the measurement z, and x_z^t represents the x-coordinate of the center of the measurement z in the current frame; y_o^{t-1} represents the y-coordinate of the center of the target o, and x_o^{t-1} represents the x-coordinate of the center of the target o estimated by the Kalman filter in the previous frame; $v_{y_o}^{t-1}$ represents the velocity of the target o along with the y-axis, and $v_{x_o}^{t-1}$ represents the velocity of the target o along with the x-axis estimated by the Kalman filter in the previous frame, $\arctan(v_{y_o}^{t-1}/v_{x_o}^{t-1})$ represents the motion direction of the target o in the image coordinate in the previous frame; parameter σ_3^2 represents the variance for the velocity direction and is set to $\sigma_3^2 = 1200$.

4) The affinity between the predicted state of the target o and the measurement z in terms of color information is defined as follows:

$$f_4(o, z) = \exp\left(-\frac{1 - \rho(H_c(o), H_c(z))}{2\sigma_4^2}\right) \tag{19}$$

where $H_c(\cdot)$ represents the RGB color histogram, ρ represents the Bhattacharyya coefficient for the color histogram of the target o and the color histogram of the measurement z, parameter σ_4^2 is set to $\sigma_4^2 = 3$.

5) The affinity between the predicted state of the target o and the measurement z in terms of the histogram of oriented gradient is defined as follows:

$$f_5(o, z) = \exp\left(-\frac{1 - \rho(H_g(o), H_g(z))}{2\sigma_5^2}\right) \qquad (20)$$

where $H_g(\cdot)$ represents the histogram of oriented gradient, ρ represents the Bhattacharyya coefficient for the histogram of oriented gradient of the target o and the histogram of oriented gradient of the measurement z, parameter σ_5^2 is set to $\sigma_5^2 = 3$.

Then, the multi-cue-based distance measure in the modified intuitionistic fuzzy clustering algorithm is defined as follows:

$$\begin{aligned} g(o, z) = 1 - f_1(o, z) \times f_2(o, z) \times f_3(o, z) \times f_4(o, z) \\ \times f_5(o, z) \end{aligned} \qquad (21)$$

where $g(o, z)$ denotes the multi-cue-based distance measure between the predicted state of the target o and the measurement z, f_1 is defined by Eq. (16), f_2 is defined by Eq. (17), f_3 is defined by Eq. (18), f_4 is defined by Eq. (19), and f_5 is defined by Eq. (20).

The fuzzy membership degrees are computed based on the multi-cue-based distance measure defined by Eq. (21), and the intuitionistic fuzzy membership degree matrix $[\eta_{ik}]_{l \times r}$ is constructed based on the obtained fuzzy membership degrees. Finally, the correct association pairs between the targets and the measurements are determined according to the principle of maximum intuitionistic fuzzy membership degree. The proposed data association algorithm based on intuitionistic fuzzy sets is summarized in Algorithm 1.

Algorithm 1 proposed data association algorithm based on intuitionistic fuzzy sets

Input: A set of predicted states of the targets $O = \{o_1, \ldots, o_l\}$ and a measurement set $Z = \{z_1, \ldots, z_r\}$ at frame t.

Output: The best association pairs
$\{(o_i, z_k)\}, i \in \{1, \ldots, l\}; k \in \{1, \ldots, r\}$

1. Compute the multi-cue-based distance measure between predicted state of each target in O and each measurement in Z by Eq. (16)–Eq. (21).

2. Construct the intuitionistic fuzzy membership degree matrix $[\eta_{ik}]_{l \times r}$ by Eq. (10)–Eq. (15). Mark every row and column of the current intuitionistic fuzzy membership degree matrix.

3. Find the maximum within all the marked elements of the intuitionistic fuzzy membership degree matrix, $\eta_{pq} = \max([\eta_{ik}]_{l \times r}), p \in \{1, \ldots, l\}, q \in \{1, \ldots, r\}$, then unmark every element in the p–th row and every element in the q–th column of the intuitionistic fuzzy membership degree matrix. If $\eta_{pq} \geq \{\eta_{pj}\}_{j=1, \ldots, r}$, $\eta_{pq} \geq \{\eta_{iq}\}_{i=1, \ldots, l}$ and $f_1 \geq \theta$, where f_1 is defined by Eq.(17) and the parameter θ is set to $\theta = 0.65$, then (o_p, z_q) is a correct association pair.

4. If there are both marked row and marked column in the intuitionistic fuzzy membership degree matrix, then return to step 3.

2.2 Track Management

When a new target enters the scene, the corresponding track needs to be initialized automatically. Also, when a tracked target leaves the scene, the corresponding track needs to be terminated automatically. Therefore, to enable automatically and correctly initialize or terminate a target track, several rules of track management are developed. These rules are described as follows:

i) The measurement which is not associated with any existing targets is initialized as a new track, and it is marked as a temporal target. At the same time, a Kalman filter is initialized for it.

ii) The temporal target which is consistently associated in λ_1 frames is marked as a real target, where the parameter λ_1 satisfies $\lambda_1 > 1$.

iii) The states of the temporal targets and of the real targets are updated and predicted by the corresponding Kalman filters.

iv) The tracks of the temporal targets which are not associated with any measurements in subsequent λ_2 frames are automatically terminated. The tracks of the real targets which are not associated with any measurements in subsequent λ_3 frames are automatically terminated. The parameters λ_2 and λ_3 satisfy $\lambda_3 \geq \lambda_2 \geq 1$.

2.3 Overall Online Visual Multiple Target Tracking Algorithm

The overall intuitionistic fuzzy data association-based online visual multiple target tracking method is described in Algorithm 2.

Algorithm 2 overall algorithm for the online visual multi-target tracking method

Input: t–th video frame f^t

Output: Estimated states of the targets $\{\hat{o}_k\}_{k=1, \ldots, l}$ with identities

1. Use the object detector based on aggregated channel features [2] to generate the measurements $\{z_i\}_{i=1, \ldots, r}$ at the current frame.

2. Find the best association pairs $\{(o_n, z_k)\}_{n \in \{1, \ldots, l\}; k \in \{1, \ldots, r\}}$ between the measurements $\{z_i\}_{i=1, \ldots, r}$ and the predicted states of the targets $\{o_j\}_{j=1, \ldots, l}$ by using the intuitionistic fuzzy sets-based data association algorithm which is described in **Algorithm 1**.

3. Update the states of the targets with the associated measurements according to rule (iii) of track management.

4. Initialize temporal target tracks for new targets and produce real target tracks according to rule (i) and rule (ii) of track management, respectively.

5. Terminate target tracks for invalid targets according to rule (iv) of track management.

6. Predict the states of existing targets at the next frame according to rule (iii) of track management.

3 Experiments

The proposed method is evaluated on three publicly available challenging video sequences, including both low-density and high-density sequences. The comparison results between the proposed method and other related methods are also reported. The detailed information of the used datasets is shown in Table 1.

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

- *MOTP*($\uparrow$): Multi-target tracking precision, average of the bounding box overlap over all tracked targets. It is defined as $\text{MOTP} = \left(\sum_t \sum_i \frac{\text{Area}(h_i^t \cap o_i^t)}{\text{Area}(h_i^t \cup o_i^t)}\right) \bigg/ \left(\sum_t n_t\right)$, where $\text{Area}(h_i^t \cap o_i^t)$ denotes the intersection area between bounding boxes of the ith ground truth and the ith tracking result at the time t, $\text{Area}(h_i^t \cup o_i^t)$ denotes the union area between bounding boxes of the ith ground truth, and the ith tracking result at the time t, n_t denotes the number of tracked targets at the time t.
- MOTA($\uparrow$): Multi-target tracking accuracy, it is defined as $\text{MOTA} = 1 - \left(\sum_t (FP_t + FN_t + IDS_t)\right) \bigg/ \left(\sum_t m_t\right)$, 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.
- 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.

For items with $\uparrow$, higher scores indicate better results, for those with $\downarrow$, lower scores indicate better results.

3.1 PETS.S2L1 Dataset

Quantitative results on the PETS.S2L1 dataset are shown in Table 2, while some visual tracking results by the

Table 1 The information of the used datasets

Name	FPS	Length	Resolution	Trajectories
PETS.S2L1	7	795	768 × 576	19
PETS.S2L2	7	436	786 × 576	42
Town Centre	2.5	450	1920 × 1080	226

proposed method are shown in Fig. 3. As can be seen, the proposed method achieves competitive results to other multi-target tracking approaches. The proposed method achieves the best performance in terms of MOTA score. This means that the proposed method can handle the false positive detections effectively. At the same time, the MOTP score of the proposed method is a little lower, and also is the FG score. This means that the output target trajectories of the proposed method are not properly aligned with the ground truth, leading to more fragments. This is because that most targets in PETS.S2L1 do not satisfy the simple constant velocity assumption. Therefore, the precision of the target state estimated by the Kalman filter is low. In the PETS.S2L1 sequence, people walk close together causing frequent occlusion, and there also exist complete occlusions caused by the traffic sign. These challenges make data association more difficult. Figure 3 shows some tracking examples of the proposed method. As can be seen, the proposed method is able to track the targets correctly under severe occlusions to some extent, such as person 2 and person 3 at frame 25; person 6 and person 9 at frame 155. However, the identities of person 2 and person 3 are changed to 6 and 8, respectively, at frame 65, because both persons make a sharp direction change after long-term occlusions.

3.2 Town Centre Dataset

Quantitative results on the Town Centre dataset are shown in Table 3, while some visual tracking results by the proposed method are shown in Fig. 4. As can be seen, the proposed method achieves better performance on most scores. The proposed method achieves the best performance in terms of MOTA, MT, and ML scores. This means that the proposed method correctly tracks more persons, although the IDS score and the FG score of the proposed method are a slightly higher. In the Town Centre sequence, there are more targets than in the PETS.S2L1 sequence. On average, 16 people are visible at any time, resulting in frequent dynamic occlusions. Many people are not detected because of the complete or partial occlusions. Tracking is difficult in this semi-crowded environment, especially at the low frame rate. Figure 4 shows some tracking examples by the proposed method. As can be seen, the proposed method is able to track multiple targets robustly under frequent occlusions to some extent, such as person 4 at frame 19; person 17, person 21, person 22, person 23, person 24, and person 26 at frame 57 and frame 63; person 39, person 45, and person 46 at frame 134. Some persons are lost due to the missed detections, such as the persons pushing a bicycle at frame 19 and frame 134, respectively.

3.3 PETS.S2L2 Dataset

Quantitative results on the PETS.S2L2 dataset are shown in Table 4, while some visual tracking results by the proposed method are shown in Fig. 5. As can be seen, the proposed method achieves the best performance in terms of MOTA, IDS, MT, ML, and FG scores, while keeping the MOTP score as high as 67.3 %. The obvious

performance improvement indicates the effectiveness of the proposed method. The PETS.S2L2 sequence is more difficult than PETS.S2L1 since the crowded density of PETS.S2L2 is much higher and missing detections often occur due to frequent occlusions. Moreover, target appearance changes heavily, caused by different illumination conditions in different image regions. However, by using data association algorithm based on intuitionistic

Table 2 Comparison of tracking results on PETS.S2L1

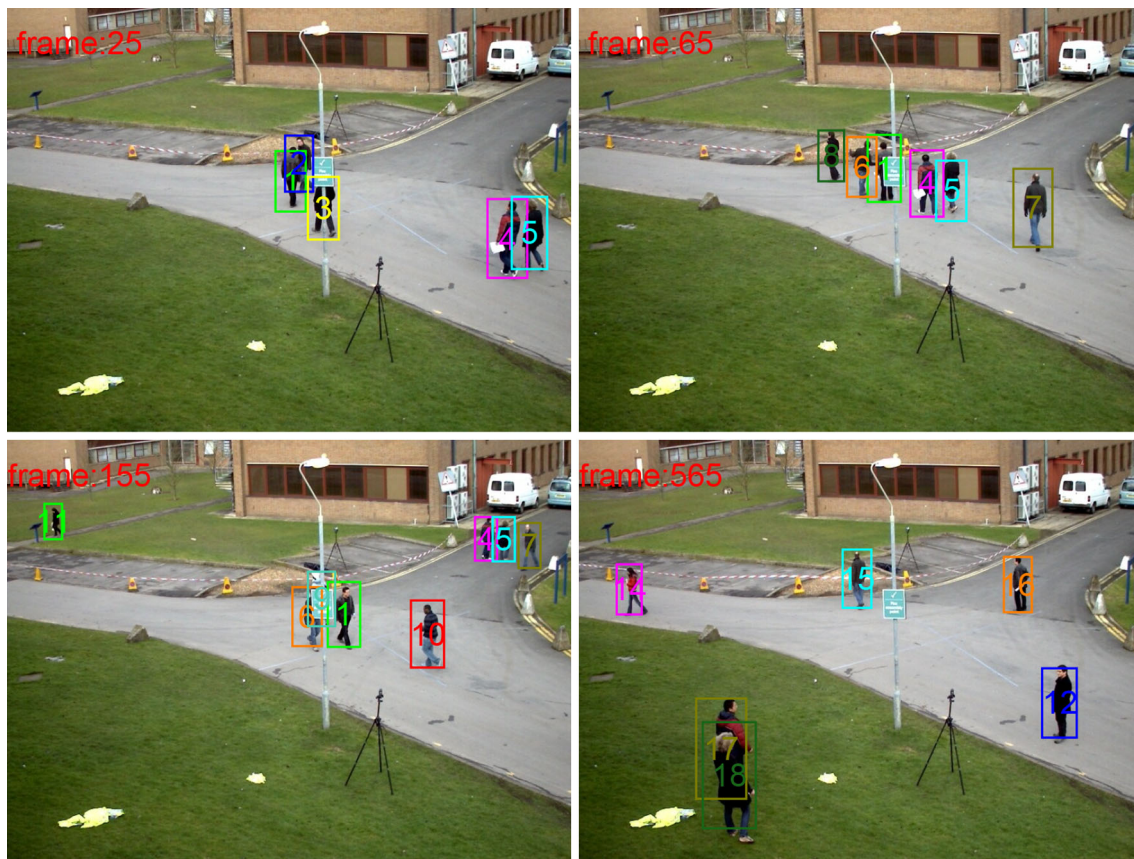| Method	MOTA (%)	MOTP (%)	IDS	MT (%)	ML (%)	FG
Berclazet al. [4]	80.3	72.0	13	89.5	10.5	22
Bae et al. [13]	83.0	69.6	4	100	0	4
Proposed method	88.7	65.7	4	94.7	0	52

Fig. 3 Tracking examples on PETS.S2L1

Table 3 Comparison of tracking results on Town Centre

Method	MOTA (%)	MOTP (%)	IDS	MT (%)	ML (%)	FG
Bae et al. [13]	1.6	69.0	26	0.0	71.7	114
Dicle et al. [8]	15.0	70.0	76	2.2	58.4	217
Proposed method	34.6	69.4	49	10.0	44.7	197

Fig. 4 Tracking examples on Town Centre

Table 4 Comparison of tracking results on PETS.S2L2

Method	MOTA (%)	MOTP (%)	IDS	MT (%)	ML (%)	FG
Bae et al. [13]	30.2	69.2	284	2.4	19.0	499
Dicle et al. [8]	34.4	70.0	251	0.0	23.8	514
Proposed method	51.2	67.3	171	16.3	7.0	318

Table 5 Performance comparison results

Algorithm	Sequence	MOTA (%)	MOTP (%)	IDS	MT (%)	ML (%)	FG
Intuitionistic fuzzy data association	PETS.S2L1	88.7	65.7	4	94.7	0	52
	Town Centre	34.6	69.4	49	10.0	44.7	197
	PETS.S2L2	51.2	67.3	171	16.3	7.0	318
Data association based on association cost	PETS.S2L1	88.3	65.7	8	94.7	0	53
	Town Centre	34.9	69.3	96	9.0	39.5	239
	PETS.S2L2	51.1	67.3	265	16.3	7.0	431

Table 6 Speed comparison results

Algorithm	Berclaz et al. [4]	Dicle et al. [8]	Bae et al. [13].	Proposed algorithm
Frame/second	3.0	5.4	3.4	1.4

fuzzy sets, the proposed method successfully tracks more targets. Figure 5 shows some tracking examples by the proposed method. As can be seen, the proposed method is able to track multiple targets robustly under severe and partial occlusions to some extent, such as person 20, person 27, and person 28 at frame 40 and frame 80. The proposed method also can deal with illumination and appearance variation to some extent, such as person 33 at frame 80 and frame 120; person 21 at frame 40, frame 80, frame 120, and frame 160.
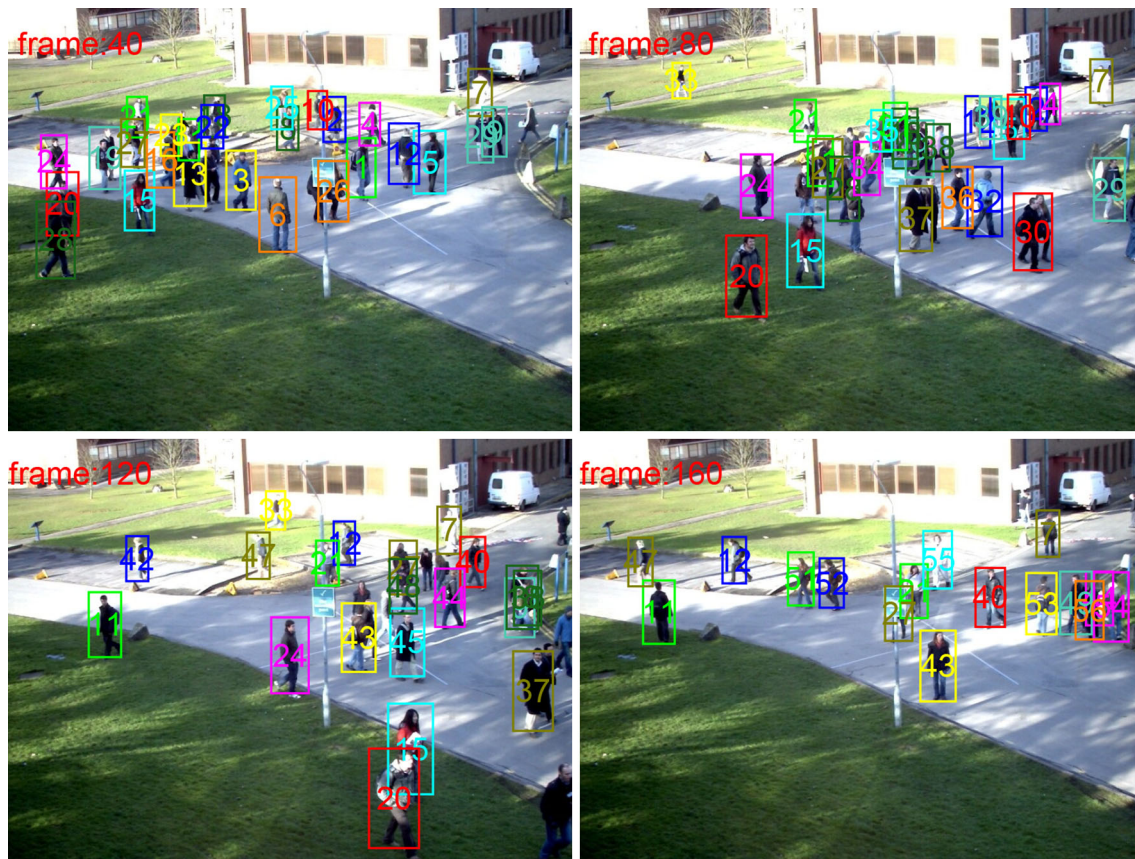
Fig. 5 Tracking examples on PETS.S2L2

3.4 Performance Evaluation of the Intuitionistic Fuzzy Data Association

To verify the effectiveness of the proposed intuitionistic fuzzy sets-based data association algorithm, the Hungarian algorithm is applied using Eq. (16) to compute the association cost instead of the intuitionistic fuzzy membership degrees, and its performance is compared with the proposed method. Quantitative results of the comparison on all datasets are shown in Table 5. As can be seen, although the performance differences in terms of MOTA score and MOTP score between the proposed method and the association cost-based method are not much, the proposed method achieves significant improvements in terms of IDS score and FG score. Compared with the association cost-based method, the proposed method reduces the ID switches and fragments by about 49 and 52 on average, respectively. The obvious improvement in IDS and FG scores indicates that the proposed method can better track targets in complex scenes due to the correct association. It verifies the effectiveness of the intuitionistic fuzzy data association.

3.5 Speed of the Proposed Online Visual Multiple Target Tracking Algorithm

The proposed algorithm is implemented using the MATLAB on a PC with 3.6 GHz × 2cores CPU and 8 GB RAM. The complexity of the proposed algorithm depends on the number of detections and targets in a sequence. The average speeds of both the proposed algorithm and the compared methods are given in Table 6. Note that the reported numbers in Table 6 do not include the detection time. As presented in Table 6, although the proposed algorithm achieves significant improvement in terms of tracking accuracy, the speed of the proposed algorithm is the lowest. It means that the proposed algorithm sacrifices a small amount of computational time in exchange for better tracking accuracy. However, the differences of the speeds are not much. The most expensive processing in the proposed algorithm is spent in calculating the multi-cue-based distance measure. To improve the computational complexity of the proposed algorithm, the parallel programming or GPU-based processing can be used. The real-

time implementation of the proposed algorithm is a topic for future research.

3.6 Discussion

In the low-density sequence PETS.S2L1, the proposed method achieves competitive results to other multi-target tracking approaches. Compared with the offline multi-target tracking approach [4], the proposed method achieves better performance in terms of tracking accuracy mainly due to the utilization of the information of multiple cues including appearance, shape, and motion. By using the multi-cue-based distance measure, the proposed method performs well in discriminating different targets, especially when two targets walk closely. Compared with the online visual multi-target tracking approach [13], the proposed method also achieves better performance in terms of tracking accuracy. Because the rules of track management in the proposed method are able to correctly initialize most target tracks, thus most false positives are filtered out. However, the proposed method produces more fragments compared with the above two approaches. This is mainly because the Kalman filter cannot effectively handle the complex movements of the targets. A possible solution to this problem is to replace the current Kalman filter with non-linear filter which is able to model the motion of the target more accurately.

In the more crowded scene Town Centre and the high-density sequence PETS.S2L2, the appearances of the targets are quite similar and the occlusion happens frequently among the targets, which make data association more difficult. Compared with both offline multi-target tracking approach [8] and online visual multi-target tracking approach [13], the proposed method achieves significant improvement in most evaluation scores mainly due to the utilization of intuitionistic fuzzy data association. The intuitionistic index effectively models the uncertainty of the possible association pair, and the intuitionistic fuzzy point operator can extract useful information from the intuitionistic index. Therefore, compared with the other methods, the proposed intuitionistic fuzzy data association algorithm can better deal with the uncertain and noisy measurements, which is further confirmed by the experiment in Sect. 3.4. The drawback of the proposed method is that the higher accuracy is obtained by sacrificing the computational complexity. A possible solution to this problem is to use parallel programming or GPU-based processing. In addition, although the proposed method focuses on the visual multi-target tracking, the idea of intuitionistic fuzzy data association can also be used in other related fields [25–31].

4 Conclusion

In this paper, a novel intuitionistic fuzzy data association-based online visual multi-target tracking method is proposed. Different from previous tracking methods, the proposed method does not use the algorithm structure of the joint probabilistic data association filter and instead replaces the association costs between the measurements and the targets with the intuitionistic fuzzy membership degrees which are obtained by the modified intuitionistic fuzzy clustering algorithm. Furthermore, a novel intuitionistic index is defined based on two different fuzzy c-means clustering in the proposed method. In addition, a new track management part is developed to enable automatically and correctly initialize or terminate the target track. Experimental results using challenging sequences show the improved performance of the proposed online visual multi-target tracking method compared with other approaches.

References

1. Felzenszwalb, P.F., Girshick, R.B., McAllester, D., et al.: Object detection with discriminatively trained part-based models. IEEE Trans. Pattern Anal. Mach. Intell. **32**(9), 1627–1645 (2010)
2. Dollár, P., Appel, R., Belongie, S., et al.: Fast feature pyramids for object detection. IEEE Trans. Pattern Anal. Mach. Intell. **36**(8), 1532–1545 (2014)
3. Butt, A.A., Collins, R.T.: Multi-target tracking by lagrangian relaxation to min-cost network flow. In: Proceedings of Computer Vision and Pattern Recognition, pp. 1846–1853 (2013)
4. Berclaz, J., Fleuret, F., Türetken, E., et al.: Multiple object tracking using K-shortest paths optimization. IEEE Trans. Pattern Anal. Mach. Intell. **33**(9), 1806–1819 (2011)
5. Yang, B., Nevatia, R.: Multi-target tracking by online learning a CRF model of appearance and motion patterns. Int. J. Comput. Vision. **107**(2), 203–217 (2014)
6. Milan, A., Schindler, K., Roth, S.: Detection- and trajectory-level exclusion in multiple object tracking. In: Proceedings of Computer Vision and Pattern Recognition, pp. 3682–3689 (2013)
7. Milan, A., Roth, S., Schindler, K.: Continuous energy minimization for multi-target tracking. IEEE Trans. Pattern Anal. Mach. Intell. **36**(1), 58–72 (2014)
8. Dicle, C., Camps, O., Sznaier, M.: The way they move: tracking multiple targets with similar appearance. In: Proceedings of Computer Vision and Pattern Recognition, pp. 2304–2311 (2013)
9. Okuma, K., Taleghani, A., DeFreitas, N. et al.: A boosted particle filter: multitarget detection and tracking. In: Proceedings of ECCV, pp. 28–39 (2004)

10. Fortmann, T., Bar Shalom, Y., Scheffe, M.: Sonar tracking of multiple targets using joint probabilistic data association. IEEE J. Ocean. Eng. **8**(3), 173–184 (1983)

11. Breitenstein, M.D., Reichlin, F., Leibe, B., et al.: Online multi-person tracking-by-detection from a single, uncalibrated camera. IEEE Trans. Pattern Anal. Mach. Intell. **33**(9), 1820–1833 (2011)

12. Possegger, H., Mauthner, T., Roth, P.M., et al.: Occlusion geodesics for online multi-object tracking. In: Proceedings of Computer Vision and Pattern Recognition, pp. 1306–1313 (2014)

13. Bae, S., Yoon, K.: Robust online multi-object tracking based on tracklet confidence and online discriminative appearance learning. In: Proceedings of Computer Vision and Pattern Recognition, pp. 1218–1225 (2014)

14. Li, L.Q., Li, C.L., Cao, W.M., Liu, Z.X.: Fuzzy quadrature particle filter for maneuvering target tracking. Int. J. Fuzzy Syst. (2015). doi:10.1007/s40815-015-0105-5

15. Li, L.Q., Xie, W.X.: Intuitionistic fuzzy joint probabilistic data association filter and its application to multitarget tracking. Sig. Process. **96**(3), 433–444 (2014)

16. Pai, N.S., Li, T.H.S.: Tracking 3D moving targets with an integrated fuzzy Kalman filter. Int. J. Fuzzy Syst. **5**(4), 201–211 (2003)

17. Chao, C.H., Hsueh, B.Y., Hsiao, M.Y., Tsai, S.H., Li, T.H.S.: Fuzzy target tracking and obstacle avoidance of mobile robots with a stereo vision system. Int. J. Fuzzy Syst. **11**(3), 183–191 (2009)

18. Liu, H.P., Wang, L.Y., Sun, F.C.: Mean-shift tracking using fuzzy coding histogram. Int. J. Fuzzy Syst. **16**(4), 457–467 (2014)

19. Atanassov, K.: Intuitionistic fuzzy set. Fuzzy Sets Syst. **20**(1), 87–96 (1986)

20. Liu, Huawen.: Multi-criteria fuzzy decision making and fuzzy reasoning algorithm based on intuitionistic fuzzy sets and interval-valued fuzzy sets (Ph.D. thesis). Shandong University of China (2005)

21. Miyamoto, S.: Different objective functions in fuzzy c-means algorithms and kernel-based clustering. Int. J. Fuzzy Syst. **13**(2), 89–97 (2011)

22. Huang, C.-M., Ghafoor, Y., Huang, Y.-P., Liu, S.-I.: A dolphin herding inspired fuzzy data clustering model and its applications. Int. J. Fuzzy Syst. (2015). doi:10.1007/s40815-015-0093-5

23. Son, L.-H., Hai, P.-V.: A novel multiple fuzzy clustering method based on internal clustering validation measures with gradient descent. Int. J. Fuzzy Syst. (2015). doi:10.1007/s40815-015-0117-1

24. Bernardin, K., Stiefelhagen, R.: Evaluating multiple object tracking performance: the CLEAR MOT metrics. EURASIP J. Image Video Process. **2008**(1), 1–10 (2008)

25. Gao, H.-J., Ding, C.-X., Song, C.-W., Mei, J.-Y.: Automated inspection of E-shaped magnetic core elements using K-tSL-center clustering and active shape models. IEEE Trans. Indus. Inform. **9**(3), 1782–1789 (2013)

26. Wang, X., Gao, H., Kaynak, O., Sun, W.: On-line deflection estimation of X-axis beam on positioning machine. IEEE/ASME Trans. Mechatron. (2015). doi:10.1109/TMECH.2015.2456241

27. Zhao, Y., Cachard, C., Liebgott, H.: Automatic needle detection and tracking in 3D ultrasound using an ROI-based RANSAC and Kalman method. Ultrason. Imaging **35**(4), 283–306 (2013)

28. Liu, J., Laghrouche, S., Wack, M.: Observer-based higher order sliding mode control of power factor in three-phase AC/DC converter for hybrid electric vehicle. Int. J. Control. **87**(6), 1117–1130 (2014)

29. Li, H., Gao, Y., Shi, P., Lam, H.: Observer-based fault detection for nonlinear systems with sensor fault and limited communication capacity. IEEE Trans. Autom. (2015). Control. doi:10.1109/TAC.2015.2503566

30. Li, H., Wu, C., Yin, S., Lam, H.-K.: Observer-based fuzzy control for nonlinear networked systems under unmeasurable premise variables. IEEE Trans. Fuzzy Syst. (2015). doi:10.1109/TFUZZ.2015.2505331

31. Li, H., Pan, Y., Shi, P., Shi, Y.: Switched fuzzy output feedback control and its application to Mass-spring-damping system. IEEE Trans. Fuzzy Syst. (2015). doi:10.1109/TFUZZ.2015.2505332

Li Jun is a Ph.D. student in the School of Information Engineering, Shenzhen University, Shenzhen, Guangdong, China. His research interests include object tracking and pattern recognition.

Xie Wei-xin was born in Guangzhou (Guangdong), China, on December 10, 1941. He received the bachelor degree in 1965 from Xidian University, Xi'an, China. Currently, he is with the College of Information Engineering, Shenzhen University, Shenzhen, China and works as a Ph.D. supervisor. His fields of interests include Intelligent information processing, Fuzzy information processing, Image processing, Pattern recognition, etc.

Li Liangqun was born in 1979. He is an associate professor of information engineering, Shenzhen University.