

A New Algorithm for Improving Basic Model Based Foreground Detection Using Neutrosophic Similarity Score

Keli Hu* and En Fan

Department of Computer Science and Engineering, Shaoxing University, Shaoxing
312000, Zhejiang, China
*ancimoon@gmail.com

Abstract

Foreground detection is a task for detecting the moving objects in the scene like in video surveillance. Several basic background models are often used due to their high efficiency. However, their results are not good when there exists noisy information generated by the bad weather, camera jitter, etc. Neutrosophic set (NS) is as a new branch of philosophy dealing with the origin, nature and scope of neutralities. It has an inherent ability to handle the indeterminant information like the noise included in images and video sequences. Till now, NS has been successfully applied into many computer vision research fields, such as image segmentation and skeleton extraction. In this paper, we utilize the NS to improve the robustness of the basic background model. First, a sequence of background subtraction images is calculated. Second, such a sequence is represented in the NS domain via three membership subsets T , I , and F . Finally, a neutrosophic similarity score (NSS) is defined to highlight the moving objects on the image plane. Then the Otsu's method is employed to determine an optimized value for extracting the area of the moving objects. Experiments are conducted on a variety of real-world video sequences, and the experimental results demonstrate that the proposed approach can extract the moving objects robustly.

Keywords: foreground detection, background model, neutrosophic set, neutrosophic similarity score

1. Introduction

Analysis of video sequences is an active research field. Detecting moving objects is the first step for the applications in this research area [1, 2], like video surveillance, optical motion capture. The purpose of such a task is to separate the moving objects called “foreground” from the static information called the “background”.

The method of background subtraction is mainly used for foreground separation, and lots of efforts have been made in recent years. A mixed Gaussians model for moving objects detection was proposed in [3], each pixel was modeled as a mixture of Gaussians (MOG) and an online approximation was used to update the model. Due to the performance of the mixed Gaussians in background subtraction, many works were proposed to improve the original MOG [2, 4, 5]. For video surveillance, imprecision and uncertainties are always generated in the whole process of background subtraction. Thus, fuzzy concepts are introduced into the task of foreground detection [6]. A fuzzy model was formulated in [7] to deal with decision problems arising in the foreground detection process. In [8], the authors proposed a novel background modeling method for motion detection in dynamic scenes, based on Type-2 Fuzzy Gaussian Mixture Model and Markov Random Field. The algorithm performs well when there exist dynamic

Received (July 2, 2017), Review Result (August 10, 2017), Accepted (August 30, 2017)

backgrounds as waving trees and water rippling. A comprehensive survey on fuzzy concepts used in background subtraction steps were given in [6].

Neutrosophic set (NS) [9] is as a new branch of philosophy to deal with the origin, nature and scope of neutralities. It has an inherent ability to handle the indeterminate information like the noise included in images [10-13] and video sequences. Till now, NS has been successfully applied into many computer vision research fields, such as image segmentation [10-13] and skeleton extraction [14]. Algorithms for clustering also employ the NS theory. For instance, to improve the performance of the method of c-means clustering, indeterminate factors were considered by using a NS model [15]. For the applications of the NS theory in image segmentation domain, a specific neutrosophic image was usually computed [10-13] to conquer the noise interference. In addition, many kinds of NS criteria for handling the noise disturbance are proposed [10-13, 16]. In [16], multiple criteria were proposed for solving the thresholding problem on a static image, and a neutrosophic image was built before calculating the optimal threshold. It has been proved that such a method has greatly improved the performance of image thresholding when confronting the noise problem. However, to our own knowledge, though the NS theory has shown its advantage for solving the noise problem in applications like image segmentation and skeleton extraction, information between frames of a video sequence has not been considered in the NS domain. It is meaningful to introduce the NS theory into the background modelling problem to handle those imprecision and uncertainties when determine the foreground pixel through the information from several consecutive frames.

In this work, we propose a robust foreground detection method based on the basic background model and the NS theory. First, three kinds of NS based criteria are proposed. Then a neutrosophic background subtraction image is calculated using a weighted similarity degree measurement in NS domain [17]. Finally, the method for determining a more reliable threshold is proposed to separate the foreground from the background.

The remainder of this paper is organized as follows. In Section 2, main steps of our algorithm are first given, and then the details of the proposed algorithm are illustrated in the following subsections. Experimental evaluations and discussions are presented in Section 3, and Section 4 is the conclusion.

2. Our Method

The main steps of our method for foreground extraction are summarized in Figure 1. The third step plays as the core of the method. Generally speaking, the major task of the method is to build a neutrosophic image based on the basic foreground separation model, and then determine a proper threshold to separate the moving objects from the background. Details of each step are given in the following subsections.

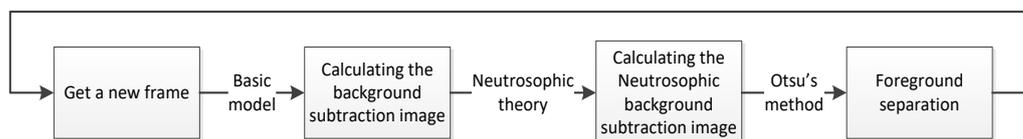


Figure 1. Main Steps of the Proposed Algorithm

2.1. Calculating Background Subtraction Image

The basic model for foreground separation is applied for generating the background subtraction image. Suppose I_t is the frame of one video sequence at time t , B_t is the background model calculated at time t , where

$$B_t = (1 - \alpha)B_{t-1} + \alpha I_t \quad (1)$$

where B_{t-1} is the previous background model at time $t-1$, $\alpha \in (0,1)$, it determines the update rate of the background model. Simply, the foreground area is calculated as

$$F_t = \begin{cases} 1 & \text{if } |I_t(x) - B_t(x)| > T \\ 0 & \text{else} \end{cases} \quad (2)$$

where $I_t(x)$ and $B_t(x)$ is the intensity at the pixel location x , corresponding to the frame and the background model respectively. If the difference between $I_t(x)$ and $B_t(x)$ is larger than the threshold T , the pixel x is then judged as an element of the foreground objects.

$S_t = |I_t - B_t|$ is called the background subtraction image here. As seen in Equation (2), the element in S_t with a higher value is more likely to be judged as foreground.

The performance of such a foreground extraction method greatly depends on the choice of the parameter T . In this work, instead of finding the proper thresholding value through the original background subtraction image, a neutrosophic subtraction image is calculated and employed, to deal with the challenges of sequence noise, bad weather, camera jitter, changing illumination, etc.

2.2. Neutrosophic Similarity Score (NSS)

A neutrosophic set with multiple criteria can be expressed as follows:

Let $A = \{A_1, A_2, \dots, A_m\}$ be a set of alternatives and $C = \{C_1, C_2, \dots, C_n\}$ be a set of criteria. Then the character of the alternative A_i ($i = 1, 2 \dots m$) can be represented by the following information:

$$A_i = \left\{ \left\langle C_j, T_{C_j}(A_i), I_{C_j}(A_i), F_{C_j}(A_i) \right\rangle \middle| C_j \in C \right\} i=1, 2 \dots m, j=1, 2 \dots n \quad (3)$$

where $T_{C_j}(A_i), I_{C_j}(A_i), F_{C_j}(A_i) \in [0,1]$. $T_{C_j}(A_i)$ denotes the degree to which the alternative A_i satisfies the criterion C_j , $I_{C_j}(A_i)$ indicates the indeterminacy degree to which the alternative A_i satisfies or does not satisfy the criterion C_j , $F_{C_j}(A_i)$ indicates the degree to which the alternative A_i does not satisfy the criterion C_j .

A method for Multicriteria decision-making based on the correlation coefficient under single-valued neutrosophic environment is proposed in [17]. The similarity degree between two elements A_i and A_j is defined as:

$$S_{C_k}(A_i, A_j) = \frac{T_{C_k}(A_i)T_{C_k}(A_j) + I_{C_k}(A_i)I_{C_k}(A_j) + F_{C_k}(A_i)F_{C_k}(A_j)}{\sqrt{T_{C_k}^2(A_i) + I_{C_k}^2(A_i) + F_{C_k}^2(A_i)} \sqrt{T_{C_k}^2(A_j) + I_{C_k}^2(A_j) + F_{C_k}^2(A_j)}} \quad (4)$$

Assume the ideal alternative

$$A^* = \left\{ \left\langle C_j, T_{C_j}(A^*), I_{C_j}(A^*), F_{C_j}(A^*) \right\rangle \middle| C_j \in C \right\} i=1, 2 \dots m, j=1, 2 \dots n$$

Then the similarity degree between any alternative A_i and the ideal alternative A^* can be calculated as

$$S_{C_k}(A_i, A^*) = \frac{T_{C_k}(A_i)T_{C_k}(A^*) + I_{C_k}(A_i)I_{C_k}(A^*) + F_{C_k}(A_i)F_{C_k}(A^*)}{\sqrt{T_{C_k}^2(A_i) + I_{C_k}^2(A_i) + F_{C_k}^2(A_i)} \sqrt{T_{C_k}^2(A^*) + I_{C_k}^2(A^*) + F_{C_k}^2(A^*)}} \quad (5)$$

Suppose $w_k \in [0,1]$ is the weight of each criteria C_k , then the weighted correlation coefficient between an alternative A_i and the ideal alternative A^* is defined by

$$W(A_i, A^*) = \sum_{k=1}^n w_k \frac{T_{C_k}(A_i)T_{C_k}(A^*) + I_{C_k}(A_i)I_{C_k}(A^*) + F_{C_k}(A_i)F_{C_k}(A^*)}{\sqrt{T_{C_k}^2(A_i) + I_{C_k}^2(A_i) + F_{C_k}^2(A_i)} \sqrt{T_{C_k}^2(A^*) + I_{C_k}^2(A^*) + F_{C_k}^2(A^*)}} \quad (6)$$

2.3. Calculating the Neutrosophic Background Subtraction Image

In this work, we try to translate the original background subtraction image S_t into the neutrosophic set domain. Let S_{NS_t} be the corresponding neutrosophic image. Suppose U is a universe, FP is a foreground pixel set in U . A pixel in S_{NS_t} is denoted as $P_{NS}(T, I, F)$, $T, I, F \in [0, 1]$ which means it is T true belonging to the FP set, I indeterminate, and F false.

According to the above definition, a pixel in S_{NS_t} is interpreted as

$$P_{NS_t}(x) = (T_t(x), I_t(x), F_t(x))$$

$T_t(x), I_t(x), F_t(x)$ represent memberships belonging to the foreground pixel set FP , indeterminate set and background pixel set, respectively. When considering the criteria, a pixel in S_{NS_t} can interpreted as

$$P_{NS_t}(x) = (C_g, T_t(x), I_t(x), F_t(x))$$

Three criteria are proposed. At the intensity criteria C_g , the corresponding $T_t(x), I_t(x), F_t(x)$ are defined as

$$T_{C_g}^t(x) = \frac{S_t(x) - S_{\min}^t}{S_{\max}^t - S_{\min}^t} \quad (7)$$

$$I_{C_g}^t(x) = \frac{SD(x) - SD_{\min}}{SD_{\max} - SD_{\min}} \quad (8)$$

$$F_{C_g}^t(x) = 1 - T_{C_g}^t(x) \quad (9)$$

where $S_t(x)$ is the intensity value of the pixel x in the original background subtraction image, S_{\max}^t and S_{\min}^t is the maximal and minimal intensity value among all of the pixels in S_t , respectively. $SD(x)$ is the standard deviation of the corresponding intensity values in the past N background subtraction images, which can be represented as

$$SV(x) = \text{stddev}(S_{t-N+1}(x), S_{t-N+2}(x), S_t(x)) \quad (10)$$

For Equation (8), SD_{\max} is the maximal standard deviation value, and SD_{\min} is the minimal one. The ideal alternative A^* is set to $(C_g, 1, 0, 0)$ under this criteria. From the Equation (5), the similarity degree between P_{NS_t} and the ideal alternative under the intensity criteria is calculated as

$$S_{C_g}(P(x), A^*) = \frac{T_{C_g}(A_i)T_{C_g}(A^*)}{\sqrt{T_{C_g}^2(A_i) + I_{C_g}^2(A_i) + F_{C_g}^2(A_i)}} \quad (11)$$

Mean filtering is known as a simple but effective procedure for image smoothing and noise reduction. It is employed as the second criteria for judging the similarity degree, we call this criteria mean criteria. Under this criteria, the corresponding $T_t(x), I_t(x), F_t(x)$ are defined as

$$T_{C_m}^t(x) = \frac{Sm_t(x) - Sm_{\min}^t}{Sm_{\max}^t - Sm_{\min}^t}, Sm_t(x) = \frac{1}{w \times w} \sum_{m=-w/2}^{w/2} \sum_{n=-w/2}^{w/2} S_t(x+m, y+n) \quad (12)$$

$$I_{C_m}^t(x) = \frac{SDm(x) - SDm_{\min}}{SDm_{\max} - SDm_{\min}} \quad (13)$$

$$F_{C_m}^t(x) = 1 - T_{C_m}^t(x) \quad (14)$$

where w in Equation (12) is the width of the mean filter window, Sm_{\max}^t and Sm_{\min}^t is the maximal and minimal intensity value after smoothing, respectively. $SD(x)$ is the corresponding standard deviation after smoothing.

According to the structure of any object, pixels nearby the pixel which is considered to be an element of the foreground with a high probability are also more likely to be parts of the foreground. Due to this, regional maxima is applied as the third criteria, then the corresponding $T_t(x), I_t(x), F_t(x)$ are represented as

$$T_{C_{ma}}^t(x) = \frac{Sma_t(x) - Sma_{\min}^t}{Sma_{\max}^t - Sma_{\min}^t} \quad (15)$$

$$Sma_t(x) = \max \{S_t(x+m, y+n) \mid m, n \in (-w/2, w/2)\}$$

$$I_{C_{ma}}^t(x) = \frac{SDma(x) - SDma_{\min}}{SDma_{\max} - SDma_{\min}} \quad (16)$$

$$F_{C_{ma}}^t(x) = 1 - T_{C_{ma}}^t(x) \quad (17)$$

where $SDma(x)$ is the corresponding standard deviation based on regional max.

As the ideal alternative defined for the intensity criteria, the ideal alternative at the mean criteria and the regional maxima criteria are defined as $(C_m, 1, 0, 0)$ and $(C_{ma}, 1, 0, 0)$, respectively. Thus, by applying Equation (6), the weighted similarity degree of a pixel x can be generated as

$$\begin{aligned} W(P(x), A^*) &= w_g \frac{T_{C_g}(A_i)T_{C_g}(A^*)}{\sqrt{T_{C_g}^2(A_i) + I_{C_g}^2(A_i) + F_{C_g}^2(A_i)}} + w_m \frac{T_{C_m}(A_i)T_{C_m}(A^*)}{\sqrt{T_{C_m}^2(A_i) + I_{C_m}^2(A_i) + F_{C_m}^2(A_i)}} \\ &+ w_{ma} \frac{T_{C_{ma}}(A_i)T_{C_{ma}}(A^*)}{\sqrt{T_{C_{ma}}^2(A_i) + I_{C_{ma}}^2(A_i) + F_{C_{ma}}^2(A_i)}} \quad (18) \\ &= w_g S_{C_g}(P(x), A^*) + w_m S_{C_m}(P(x), A^*) + w_{ma} S_{C_{ma}}(P(x), A^*) \end{aligned}$$

where $w_g + w_m + w_{ma} = 1$. Then the weighted similarity for a pixel x is employed as the value of the corresponding pixel for the neutrosophic background subtraction image.

2.3. Extracting Foreground Objects Based on NSS

Once the neutrosophic background subtraction image is constructed, the only task left is to find an optimal threshold based on such a neutrosophic image. In this work, the Otsu's method [18] is applied for tackling this problem. For the completeness of this work, details of the procedure are presented below:

- (1) Selecting one value as the initial threshold T_{NSS} .
- (2) Using this threshold to separate the pixels of the neutrosophic background subtraction image S_{NSS} into two classes C_1 and C_2 . For example, C_1 contains all the pixels who are smaller than T_{NSS} . Then we can calculate the variance of S_{NSS} by

$$\sigma = \sigma_1 p_1 + \sigma_2 p_2$$

where σ_1 and σ_2 are the variance of the similarity degrees of the pixels belonging to C_1 and C_2 , respectively. p_1 and p_2 are probability of the corresponding pixels.

- (3) Try another value as the threshold until find the minimal σ , and the corresponding value is considered as the optimal threshold.

After finding the optimal threshold, the foreground can be extracted by using Equation (2).

3. Experiments

We tested our algorithm on several challenging real-world video sequences. For comparison, we implemented two algorithms. Both algorithms employ the basic model introduced in section 2.1. The only difference is on the selection of the threshold T . The first one uses a fixed threshold, called FTFD for short. The other one employs the Otsu's method to find the optimal threshold based on the original background subtraction image at each frame, we call this algorithm OTFD for short. To gauge absolute performance, we also compare our results to the Euclidean distance [19] algorithm, we call this algorithm EDFD here.

3.1 Setting Parameters

The update rate α for our method, FTFD and OTFD are all set to 0.01. For Equation (12) and (15), the width of the filter window is set to 5. For the weight of each criteria,

w_g, w_m, w_{ma} are set to 0.5, 0.3, 0.2 in Equation (18), respectively. Finally, all parameters were kept constant for all experiments.

3.2. Foreground Extraction Results

Figures 2–4 show screen captures for some of the testing sequences. The second column of each figure shows the background subtraction image (BSIMG) at each time. Below is a more detailed discussion of the foreground extraction results.

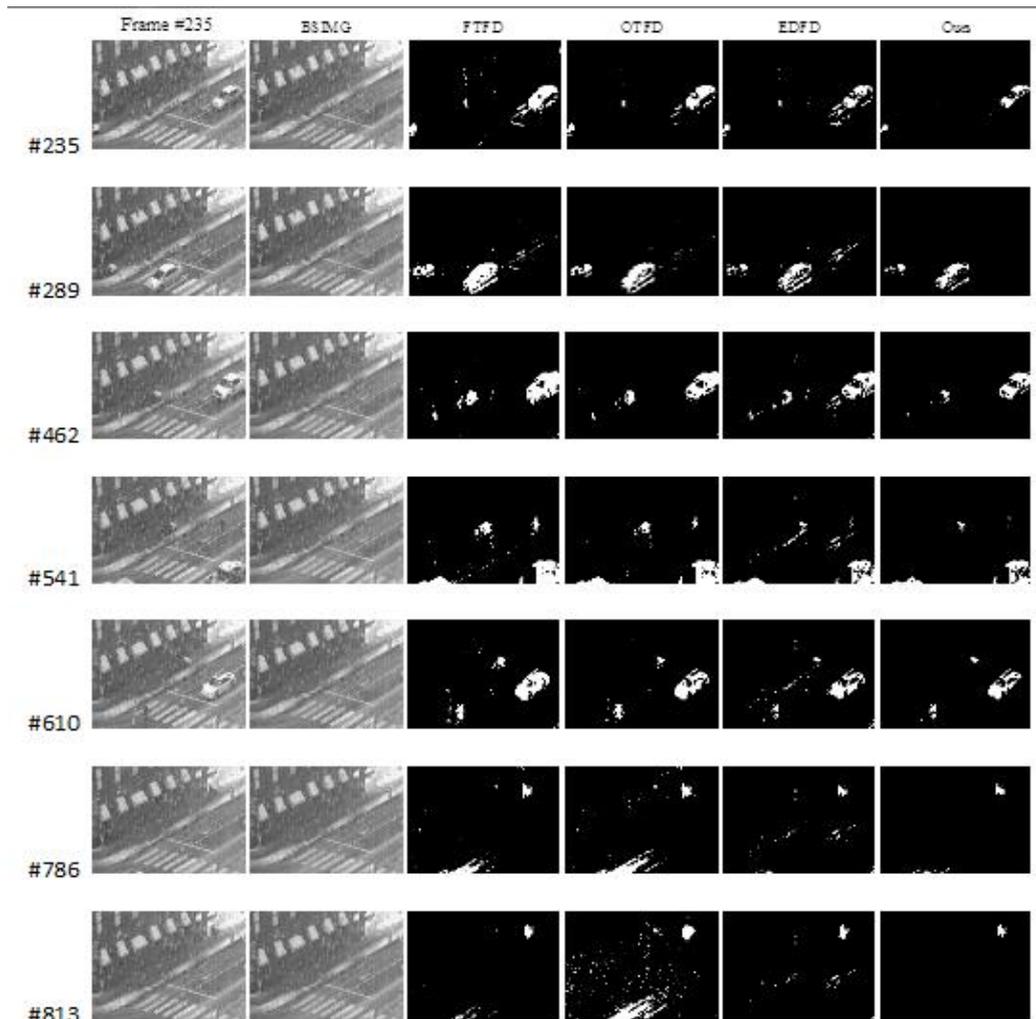


Figure 2. Screenshots of Foreground Separation Results of the Wetsnow Sequence with the Challenge of Bad Weather

wetsnow sequence: this sequence highlights the challenges of bad weather, slow movement and foreground disturbance. As shown in Figure 2, when the water droplets flow along the mirror of the camera, our method can successfully handle this, while the other three methods fail to filter out such a kind of foreground disturbance, see frame #235, #462. From frame #235 to frame #610, compared to our results, the separation results of other algorithms are noisy. As can be seen in

frame #786, there is only one person appears in the camera view and moves very slowly. The person only occupies a small area of the camera view, which leads to the failure of the OTFD. In frame #1015, before the person disappears from the view, a totally wrong threshold is employed by the OTFD to separate the foreground from the background. The EDFD also produces some noise due to the slow motion.

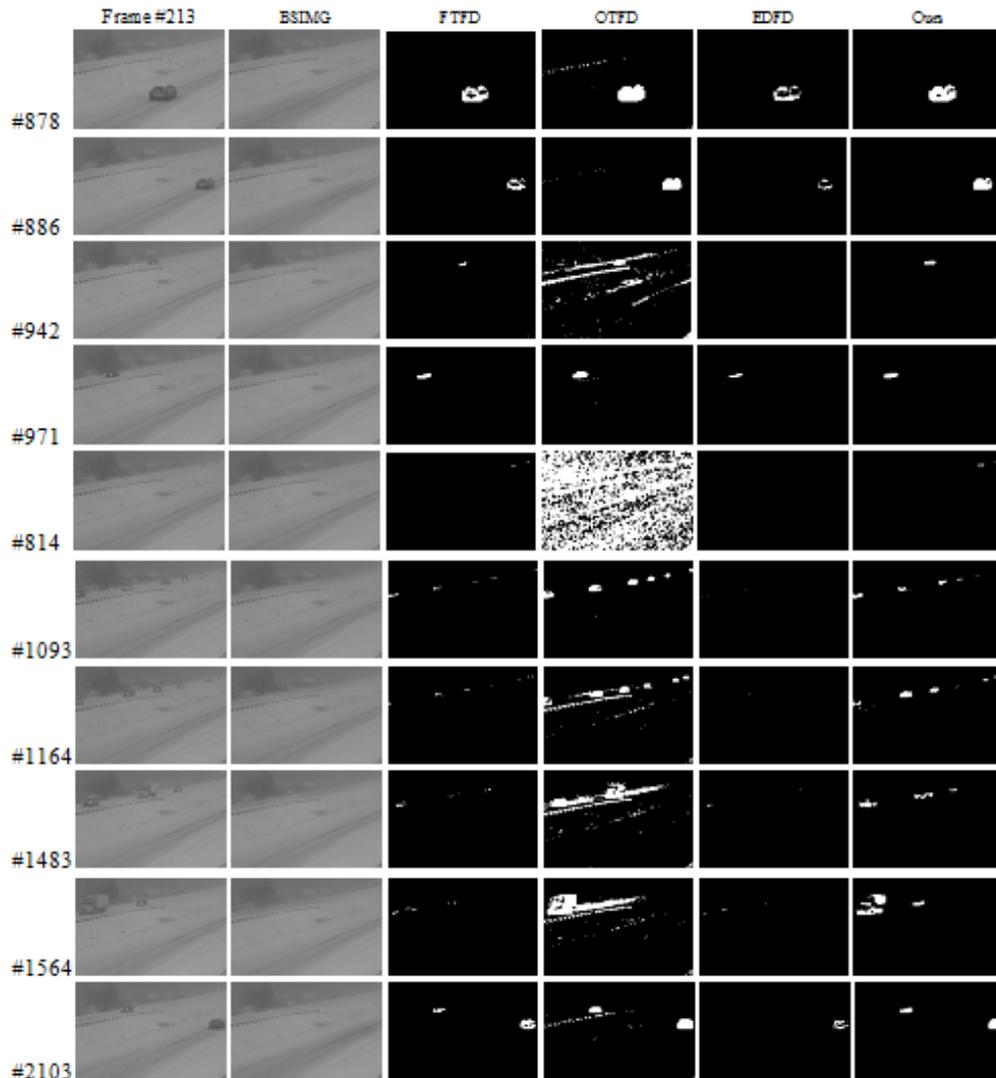


Figure 3. Screenshots of Foreground Separation Results of the *Blizzard* Sequence with the Challenge of Bad Weather

blizzard sequence: this sequence presents the challenges of bad weather, similar background, tinny camera jitter and small foreground area. Figure 3 shows some screenshots of foreground separation results of this sequence. As seen in frame #878 and #886, some noisy foreground are included in the results for OTFD, due to the small camera jitter. The EDFD gives an imperfect foreground, while our method outperforms others. In frame #814, due to the similar background, only the lamps of those two cars are detected as foreground. Noise is included in the results of OTFD on account of the small foreground area. As seen in Figure 2, our method performs the best. The FTFD sometimes perform worse due to the fixed threshold.

badminton sequence: challenge of camera jitter is presented in this sequence. As shown in Figure 4, our method performs the best, and the EDFD performs the

second best. Due to the serious camera jitter, edges jitter in each frame lead to the wrong separation of the foreground. When the camera stabilize for a while, the disturbance of the edges in each frame are successfully filtered out, see frame #1036.

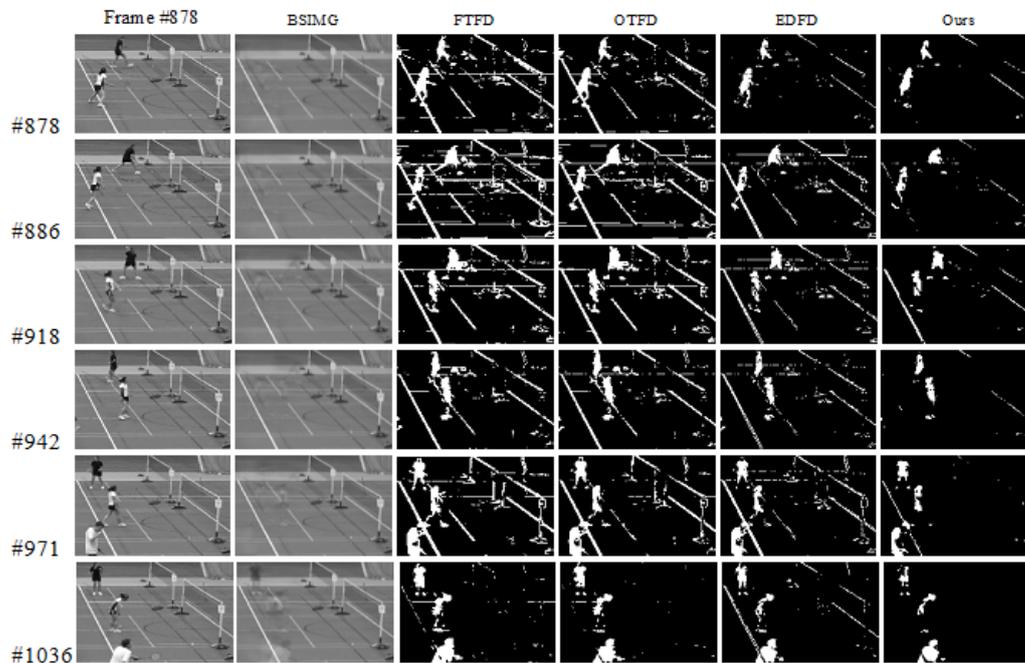


Figure 4. Screenshots of Foreground Separation Results of the Badminton Sequence with the Challenge of Camera Jitter

As seen in Figure 2-4, our algorithm performs well due to the usage of the three kinds of criteria in NS domain. The standard deviation employed by the calculation of the indeterminacy degree I contributes a lot for the performance of our method when handling the problem of camera jitter or noise pixels.

4. Conclusion

In this paper, we propose a new algorithm for foreground separation. The background subtraction image is achieved by employing a basic background model. Then three criteria are proposed, and a neutrosophic background subtraction image is calculated by using the metric of neutrosophic similarity score. Finally, an optimal threshold can be found based on such a neutrosophic image. Experimental results have shown that our method performs well when handling challenges like bad weather and camera jitter. In future, it will be our primary mission to try to further improve the robustness of foreground detection by introducing the neutrosophic scheme into a more robust background model.

Acknowledgements

This work is supported by the public welfare technology application research project of Zhejiang province under Grant No. 2016C31082.

References

- [1] Bouwmans, T., 'Traditional and recent approaches in background modeling for foreground detection: An overview', Computer Science Review, vol. 11–12, no. 0, pp. 31-66, (2014).

- [2] Thierry, B., B. Fida El, and V. Bertrand, 'Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey', *Recent Patents on Computer Science*, vol. 1, no. 3, pp. 219-237, (2008).
- [3] Stauffer, C. and W.E.L. Grimson, 'Five background mixture models for real-time tracking'. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Collins, Colorado, pp. 22-29, (1999).
- [4] Elgammal, A., R. Duraiswami, et al., 'Background and foreground modeling using nonparametric kernel density estimation for visual surveillance', *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1151-1163, (2002).
- [5] Gutchess, D., M. Trajkovic, et al. 'A background model initialization algorithm for video surveillance'. *IEEE International Conference on Computer Vision*. Vancouver, pp. 733-740, (2001).
- [6] Bouwmans, T., 'Background subtraction for visual surveillance: A fuzzy approach', *Handbook on Soft Computing for Video Surveillance*, pp. 103-134, (2012).
- [7] Maddalena, L. and A. Petrosino, 'A fuzzy spatial coherence-based approach to background/foreground separation for moving object detection', *Neural Computing and Applications*, vol. 19, no. 2, pp. 179-186, (2010).
- [8] Zhao, Z., T. Bouwmans, et al., 'A fuzzy background modeling approach for motion detection in dynamic backgrounds', *Multimedia and Signal Processing*, pp. 177-185, (2012).
- [9] Smarandache, F., *A unifying field in logics. Neutrosophy: neutrosophic probability, set and logic*, American Research Press, Rehoboth, (1999).
- [10] Guo, Y. and A. Şengür, 'A novel image segmentation algorithm based on neutrosophic similarity clustering', *Applied Soft Computing Journal*, vol. 25, pp. 391-398, (2014).
- [11] Anter, A.M., A.E. Hassanien, et al., 'Neutrosophic Sets and Fuzzy C-Means Clustering for Improving CT Liver Image Segmentation', *Advances in Intelligent Systems and Computing*, vol. 303, pp. 193-203, (2014).
- [12] Karabatak, E., Y. Guo, and A. Sengur, 'Modified neutrosophic approach to color image segmentation', *Journal of Electronic Imaging*, vol. 22, no. 1, (2013).
- [13] Zhang, M., L. Zhang, and H.D. Cheng, 'A neutrosophic approach to image segmentation based on watershed method', *Signal Processing*, vol. 90, no. 5, pp. 1510-1517, (2010).
- [14] Guo, Y. and A. Sengur, 'A novel 3D skeleton algorithm based on neutrosophic cost function', *Applied Soft Computing Journal*, vol. 36, pp. 210-217, (2015).
- [15] Guo, Y. and A. Sengur, 'NCM: Neutrosophic c-means clustering algorithm', *Pattern Recognition*, vol. 48, no. 8, pp. 2710-2724, (2015).
- [16] Guo, Y., A. Şengür, and J. Ye, 'A novel image thresholding algorithm based on neutrosophic similarity score', *Measurement: Journal of the International Measurement Confederation*, vol. 58, pp. 175-186, (2014).
- [17] Ye, J., 'Multicriteria decision-making method using the correlation coefficient under single-valued neutrosophic environment', *International Journal of General Systems*, vol. 42, no. 4, pp. 386-394, (2013).
- [18] Otsu, N., 'A threshold selection method from gray-level histograms', *Automatica*, vol. 9, no. 1, pp. 62-66, (1975).
- [19] Benezeth, Y., P. M. Jodoin, et al., 'Comparative study of background subtraction algorithms', *Journal of Electronic Imaging*, vol. 19, no. 3, pp. 033003-033003-12, (2010).

