



Smoke root detection from video sequences based on multi-feature fusion

Liming Lou¹ · Feng Chen² · Penge Cheng¹ · Ying Huang³

Received: 14 October 2021 / Accepted: 21 November 2021 / Published online: 7 March 2022
© Northeast Forestry University 2022

Abstract Smoke detection is the most commonly used method in early warning of fire and is widely used in forest detection. Most existing smoke detection methods contain empty spaces and obstacles which interfere with detection and extract false smoke roots. This study developed a new smoke roots search algorithm based on a multi-feature fusion dynamic extraction strategy. This determines smoke origin candidate points and region based on a multi-frame discrete confidence level. The results show that the new method provides a more complete smoke contour with no background interference, compared to the results using existing methods. Unlike video-based methods that rely on continuous frames, an adaptive threshold method was developed to build the judgment image set composed of non-consecutive frames. The smoke roots origin search algorithm increased the detection rate and significantly reduced false detection rate compared to existing methods.

Keywords Smoke detection · Multi-feature fusion · Search strategy · ViBe · Choquet

Introduction

Smoke monitoring from forest fires depends on visual monitoring (Alamgir et al. 2018) and developing early smoke detection technology based on video for large areas is of practical importance. Video-based smoke detection can be divided into three categories: traditional methods, deep learning-based methods, and methods combining both (Wang et al. 2020). Traditional video-based methods detect the smoke using the set features of the smoke extraction results (Nguyen et al. 2020), which can be either static or dynamic. When the recognition object is only a single smoke image frame, researchers often choose the static characteristics of smoke such as color, Local Binary Pattern (LBP), shape, wavelet, hog, irregularity, and density (Gaur et al. 2020). Among these features, color-based image segmentation has a wide range of applications in obtaining smoke candidate regions and analyzing the results (Appana et al. 2017). In addition to color space using Red–Green–Blue (RGB) (Yuan et al. 2017a, 2018; Sousa and Gamboa 2020), HSV, YUV (Prema et al. 2016), YCbCr (Wang et al. 2017; Ye et al. 2017; Yuan et al. 2017b; Sousa and Gamboa 2020), Hue-Intensity-Saturation(HIS) (Yuan et al. 2017a, 2018) and grayscale space were all investigated as smoke characteristics.

However, the complexity of the forest environment and the problems of sharpness and distance of the material obtained by videoing make it difficult to detect smoke using a single static feature (Wu et al. 2021). To address this, the dynamic characteristics from the relationship between frames, such as motion direction, texture (Zhao et al. 2021),

Project funding: The work was supported by the National Natural Science Foundation of China (grants no. 32171797 and 31800549).

The online version is available at <http://www.springerlink.com>.

Corresponding editor: Yu Lei.

✉ Penge Cheng
chengpengle@bjfu.edu.cn

¹ School of Technology, Beijing Forestry University, 100083 Beijing, People's Republic of China

² School of Nature Conservation, Beijing Forestry University, 100083 Beijing, People's Republic of China

³ Department of Civil, Construction, and Environmental Engineering, North Dakota State University, Fargo, ND 58102, USA

fixed source (Gao and Cheng 2019, 2021), and fluid characteristics of smoke, can be combined with static characteristics such as color to detect smoke, one of the major advantages of video-based smoke detection compared to photo-based detection. For such dynamic characteristics, advanced classifiers are needed to better categorize and detect smoke in video images. Currently, there are existing classifiers used for dynamic feature detection (Xia et al. 2019), including support vector machines (SVM) (Barmpoutis et al. 2013; Ye et al. 2015), AdaBoost (Yuan et al. 2015; Zhao et al. 2015), hidden Markov models (HMMs) (Savci et al. 2019), K nearest neighbors (KNN) (Zhao et al. 2021), conditional random fields (Cheng et al. 2019), Gaussian process regression (Yuan et al. 2017a), and Bayes classifier (Piccinini et al. 2008). In addition, combinations of these classifiers were also investigated.

The dynamic characteristics of smoke are different from other moving objects and using the above-mentioned traditional machine learning methods to process the motion information in the original format is difficult (Wang et al. 2020), resulting in the development of deep learning methods. The Convolutional Neural Network (CNN) can automatically extract features from the original data, and classify and optimize them to obtain better results (Li et al. 2020). He et al. (2021) developed an attention module combined with spatial and channel attention that was deeply integrated with CNN to improve detection in conditions of fog and light smoke. Li et al. (2020) reconstructed a neural network model into a new SC-CNN model, and by using a new regularized loss function-score clustering, the new model reduced over-fitting issues and improved model accuracy. Xu et al. (2019) developed a deep saliency CNN network using a combination of a deep feature map and a saliency map at both pixel-level and target-level to detect smoke. In addition, better model performance usually requires larger models. To address the compatibility of performance and model size, Jadon et al. (2019) proposed a new shallow neural network, FireNet, which achieved the optimal frame rate for fire detection to date.

To further reduce false detection and missed detection rates under changing smoke detection from forest fires, Gao and Cheng (2019) proposed the “smoke root”, which is defined as smoke source that do not change with time, believed to be the most fundamental feature of forest fire smoke that distinguishes it from other disturbances. Diffusion of smoke roots would not change over time, even in changing scenes. The stability of smoke roots is a feature that other natural phenomena such as clouds and fog do not have. To detect the smoke roots, the dynamic feature area of the fire can be obtained through the ViBe algorithm, the connected domain by opening and closing operations, and the detection of skeleton and skeleton endpoints by the use of the Zhang-Suen skeleton extraction algorithm (Zhang and

Suen 1984). The stable skeleton endpoints between consecutive frames are regarded as smoke root candidate points and were put into a two-dimensional smoke simulation engine to generate smoke used to overlap with the monitored smoke for detection. If more than 70% of the pixel values matched the simulated smoke and the monitored smoke, this region was a smoke root. Since the ViBe algorithm could not detect light smoke in a long distance, the algorithm was combined with the MSER algorithm through Bayesian fusion to form a better shape of smoke candidate regions for a complete, full-size smoke detection using video (Gao and Cheng 2021).

By considering the small diffusion range and relatively stable position of the combustion source in the early stages of a fire, it may be assumed that the source position of the smoke is relatively fixed in the continuous frames. This relatively stationary position of combustion is defined as a “smoke root”. According to the smoke root detection method developed by Gao and Cheng (2019), the premise of smoke root detection is to extract the smoke candidate regions successfully and find the stable skeleton endpoint that does not change with time, regarded as a smoke root candidate. Complete and accurate extraction of smoke candidate regions is critical for smoke root detection. Under different environmental conditions, detection results would vary with different detection algorithms. Figure 1 shows the results of the smoke candidate regions obtained by the ViBe algorithm, the smoke skeletons and the skeleton endpoints obtained using the method of Gao and Cheng (2019) in two different environmental conditions. The ViBe algorithm can extract some smoke pixels, but voids and missed detection will also occur under some environmental conditions. Although the algorithm produces a good smoke contour, it can also result in an “empty region” at the center and edges of the smoke, making it impossible to obtain the connected smoke candidate regions and limiting the search process of smoke root candidate points. To address these issues, this study integrates the dynamic and static characteristics of smoke into the ViBe algorithm to obtain more complete smoke candidate regions and develops a new smoke roots search strategy to improve detection rate and calculation efficiency.

Smoke roots are a source of smoke that are stable in consecutive frames so the continuity between frames is key for determining smoke roots. According to Gao and Cheng (2019), smoke root candidate points were identified as stable skeleton endpoints in five consecutive skeleton images. However, for a video with a rate of twenty-five frames per second or higher, the time required to obtain five image frames was only 0.2 s or lower. Such a brief time was insufficient to result in a significant change in the shape of the smoke. Therefore, many fake or false smoke roots would be determined as endpoints that did not change with time and entered into the next two-dimensional smoke simulation engine to generate candidate points, which would

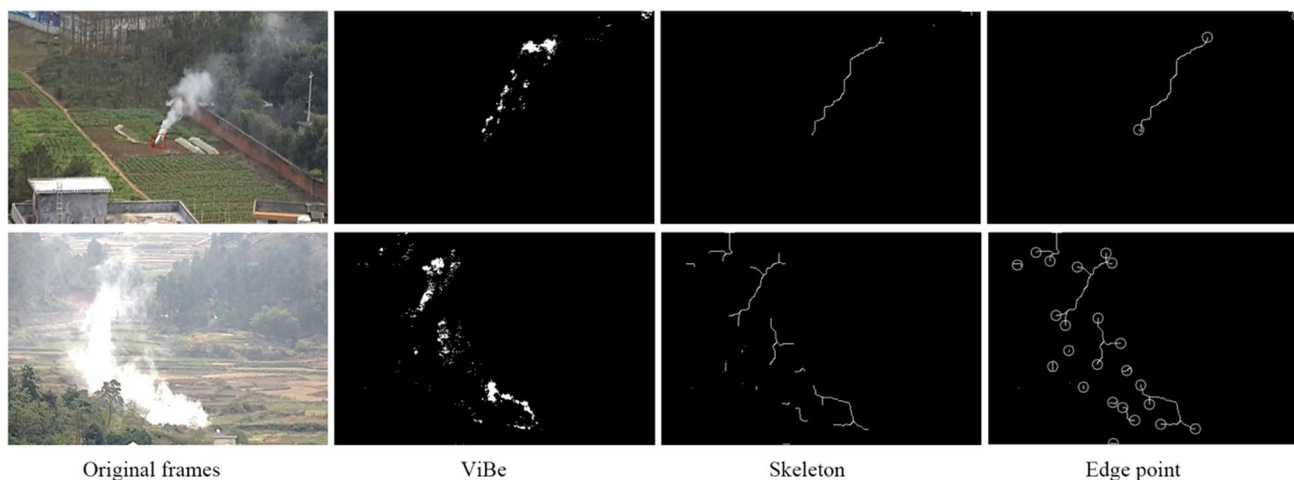


Fig. 1 Different results in different environmental conditions using Gao and Cheng (2019)’s method

significantly increase the amount of calculation and reduce the calculation efficiency.

To address this challenge, in this study, a new multi-frame discrete confidence strategy was developed to process the multi-frame connected domain images and determine the smoke root. This new strategy introduces the concept of confidence into the process of determining smoke roots, and represents the relative magnitude of the probability that the correct smoke roots appear on the image frame.

To detect the smoke roots, root candidate points are identified by analyzing the continuity of smoke roots between frames and overlapping the simulated smoke with the actual monitored smoke. Since deep learning methods only extract features from original data, they do not conform to the principle of the smoke root detection in method and original intention. Thus, although deep learning methods have been applied to recognizing dynamic features of smoke and flames, there are limited studies on applying deep learning networks to identify and extract the basic features of smoke from smoke images. In this study, a new smoke roots search algorithm, based on a multi-feature fusion dynamic extraction strategy and multi-frame discrete confidence strategy, was developed to extract smoke root features for higher detection efficiency and more accurate detection. This paper

has two major contributions: (1) A proposed new forest fire smoke roots search algorithm fusing the dynamic and static characteristics of smoke to obtain complete smoke contours through motion and grayscale detection. This new algorithm will improve detection capacity and accuracy of the edges and roots of smoke which are difficult to identify by the dynamic extraction algorithm; (2) A proposed new multi-frame discrete confidence strategy to reduce the number of smoke root candidate points or skeleton endpoints for smoke root detection. By applying the double-layer confidence selection process, this new strategy can obtain more smoke root skeleton endpoints with a high degree of confidence, thus reducing the need for large number of skeleton endpoints as required by the traditional smoke root detection method (Gao and Cheng 2019).

Materials and methods

The flowchart of the developed algorithm (Fig. 2) consists of three steps: the first is to extract the smoke regions from the continuous video frames through the multi-feature fusion method with the combination of dynamic and static stacking strategies; the second is to extract ten eligible connected

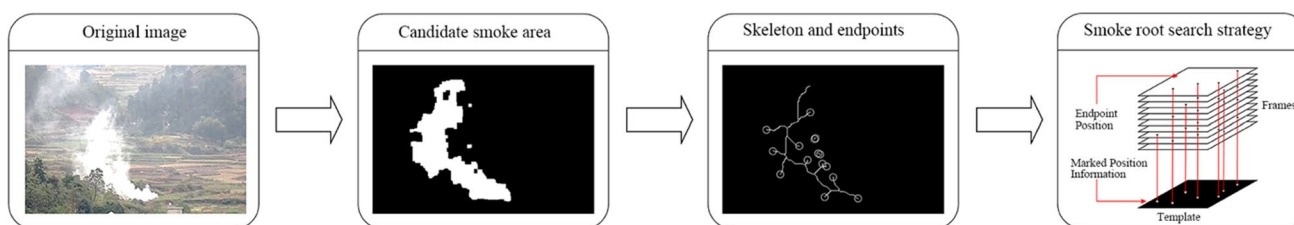


Fig. 2 Flow steps of smoke root detection during the image processing stage

domain frames as the base frames for determining the smoke root candidate points; and, the third is to extract the skeleton and skeleton endpoints of the selected connected domain frames, and screen out smoke root candidate points that have higher confidence through the multi-frame discrete confidence determination strategy. These candidate points are then put into the two-dimensional smoke simulation engine to generate the simulated smoke.

Smoke candidate regions extraction

Extracting the smoke regions from the continuous frames of the video is the first step in the smoke detection and smoke root acquisition. The ViBe algorithm was used as a foreground detection algorithm based on background updating. Gao and Cheng (2019) used this algorithm to extract the dynamic pixels from the video frame and applied the closing operation to select a large range of dynamic points from a connected domain. The closing operation was to expand the images to make the dynamic area a connected domain, identified as the smoke regions. Such a connected domain was then skeletalized and the points that were stable in the consecutive frames were considered smoke root candidate points.

Since the ViBe algorithm requires to establish and update the background pixels in the initial image frame, the image with the first frame number does not result in a good extraction. On the other hand, the dynamic extraction algorithm alone cannot identify the smoke regions under conditions of excessive smoke density. Thus, the extracted smoke regions using the ViBe algorithm may be incomplete or unable to form connected domains, resulting in ineffective detection of some smoke that is more scattered. To address this issue, this study developed a new region stacking strategy Fig. 3 using multi-feature fusion that stacks the dynamic and static candidate regions (Fig. 3).

The first step of the new stacking strategy is to extract dynamic features from the first 50 input frames to create a model image made up of the extracted ViBe dynamic features from the input frames (Fig. 3). Such extracted dynamic area contains both the correct smoke regions and part of

the background noise. An erosion operation with a core of 1×1 is performed on the extracted dynamic area to eliminate background noise and obtain a set of dynamic pixels that only belong to the smoke regions, resulting in binary images. The resulting dynamic pixel set, and the eroded input frame are then added on to the model image to obtain a binarized image, $M(i, j)$ with all dynamic pixel point sets which can be presented as:

$$M(i, j) = \begin{cases} 1, & \sum P_i(x, y) \neq 0 \\ 0, & \sum P_i(x, y) = 0 \end{cases} \quad (1)$$

where, $P_i(x, y)$ is....., x is the x coordinate and y is the y coordinate of dynamic pixel point.

After obtaining all the dynamic feature pixel points using Eq. (1), a small rectangular frame is used to frame all the smoke dynamic feature pixels on the binarized model image. The coordinates of the four end points of the rectangular box are obtained by adding 10 pixels to the pixel points of the higher smoke temperature in the vertical direction as:

$$(x_{\min}, y_{\min}), (x_{\max}, y_{\min}), (x_{\min}, y_{\max} + 10), (x_{\max}, y_{\max} + 10) \quad (2)$$

The binarized model image with the dynamic feature area of smoke can be built (Fig. 4).

After a binarized model image of the smoke is obtained, the second step is to perform fusion feature extraction

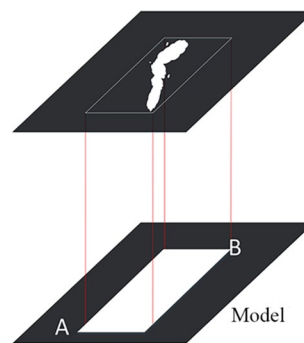


Fig. 4 Built model image

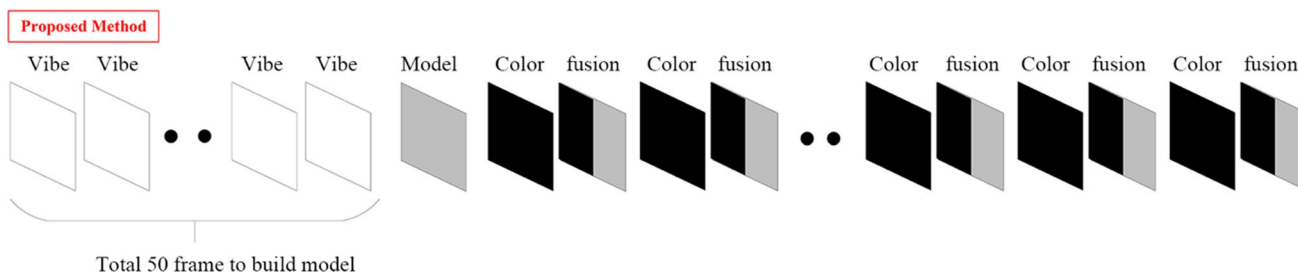


Fig. 3 Smoke candidate regions detection frame stream

on the input image and model image of each following frame (Fig. 3). The image fusion is applied to distinguish the foreground and background areas. The area enclosed by the rectangle in the input frame is considered as the foreground area, $F(x, y)$, and the area not enclosed is the background area as:

$$F(x, y) = \begin{cases} 0, & (x, y) \notin \text{the rectangular area} \\ \text{Static feature extraction,} & (x, y) \in \text{the rectangular area} \end{cases} \quad (3)$$

Within the foreground area, the foreground pixels with corresponding static characteristics of the smoke are extracted, and pixels without the static characteristics of the smoke are considered as the background pixels and are not extracted. The single-channel grayscale processing is applied on the input frame to process the images, resulting in a feature image of the binarized smoke regions represented as:

$$F(x, y) = \begin{cases} 0, & f(x, y) < \text{average} \times 1.1 \\ 255, & f(x, y) \geq \text{average} \times 1.1 \end{cases} \quad (4)$$

$$\text{average} = \frac{\sum_{i=1}^n f(x, y)}{n} \quad (5)$$

where, (x, y) belonging to the rectangular area, and $f(x, y)$ is the pixel value of the single-channel grayscale image at the position (x, y) .

Finally, the obtained feature image with binary smoke regions from Eq. (4) is closed to obtain the connected domain. Figure 5 shows three smoke image examples in different conditions obtained using the ViBe algorithm (second column) and using the new algorithm developed above (last column). The completeness and roundness of the obtained connected domain from the developed algorithm exceeds the connected domain obtained by the ViBe algorithm. In addition, the new algorithm also simplifies the smoke region generation and significantly reduces the amount of calculation needed for further obtaining smoke root candidate points.

Smoke roots determination

As seen in the first and second rows of Fig. 6, to start the new strategy for smoke root determination, ten connected domain images are selected by the adaptive threshold (AT) method. The first connected domain image is considered as the first frame of the ten connected domain images. When the connected domain change rate of the next input frame exceeds the pre-set threshold, this frame is selected as the next image until the rest of the nine frames are extracted. These ten images are then set as the base judgment image set from which smoke root candidate points are extracted. The purpose of such a process is to extract a sufficiently large change in smoke to filter out the “false smoke roots”, and to adapt to smoke with different rates of spread in different

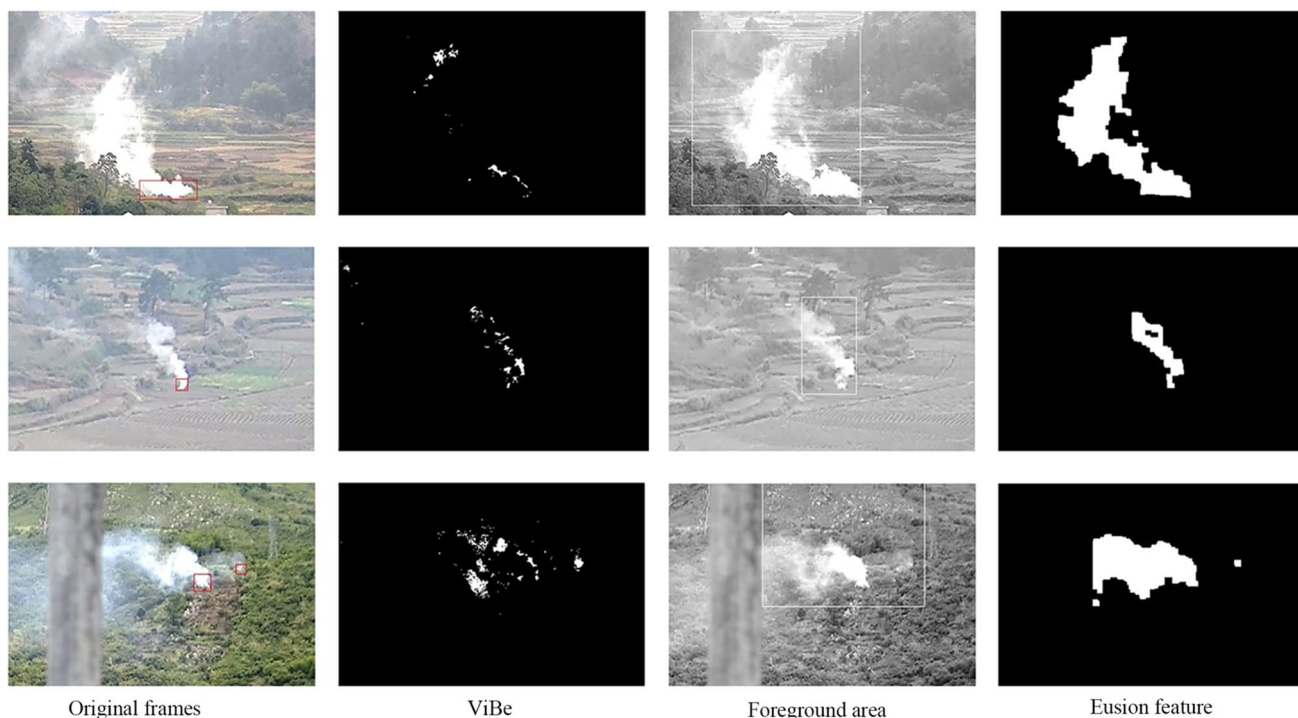
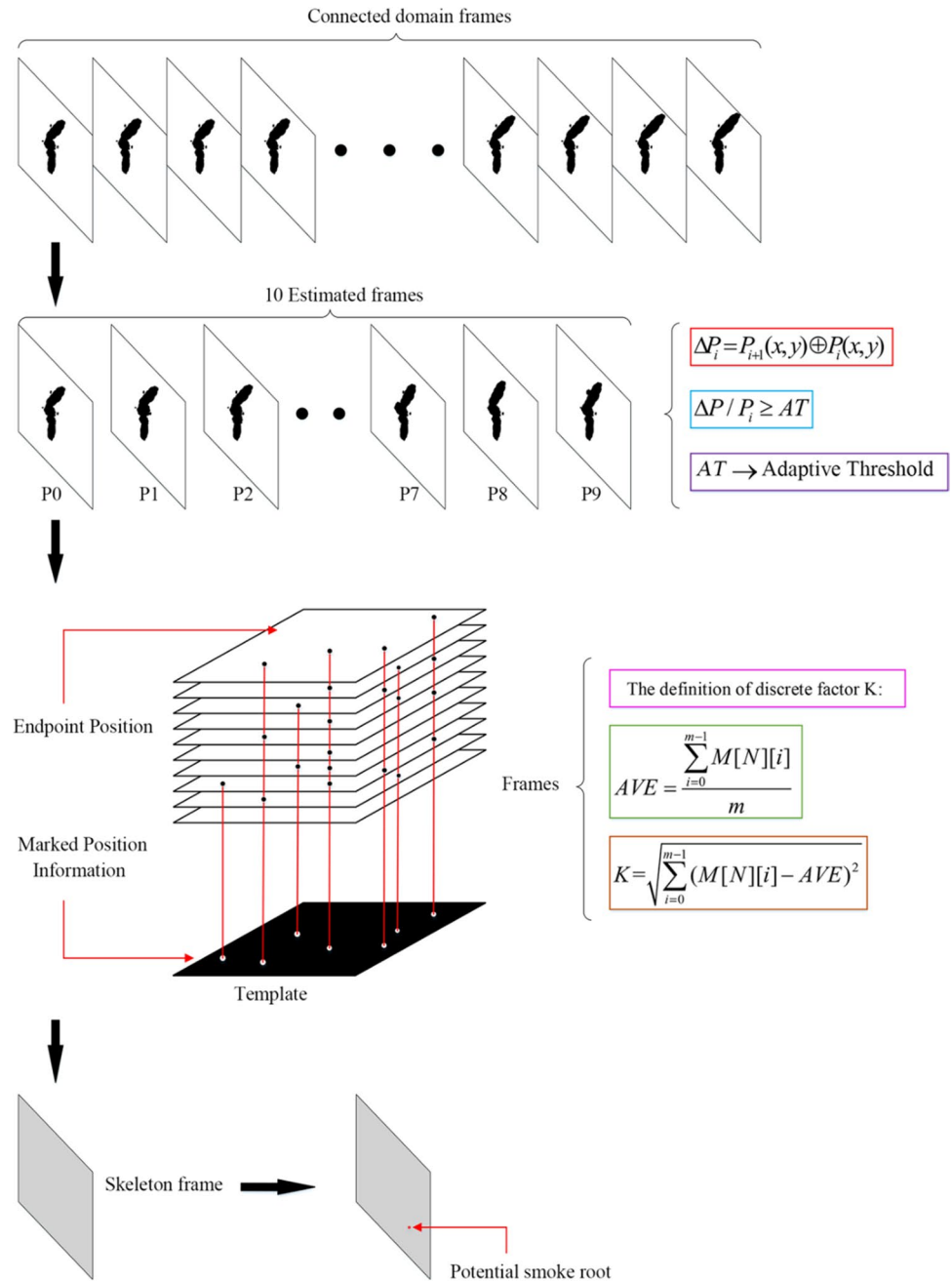


Fig. 5 Comparing the results of connected domain using ViBe algorithm and the newly developed algorithm

Fig. 6 Flowchart of the new multi-frame discrete confidence strategy to determine smoke root from connected domain frames



environments. This new strategy can detect the candidate points of the smoke roots rapidly when the smoke spreads quickly and detect the smoke roots correctly and smoothly when the smoke spreads too slowly.

After the base judgment image set are extracted, shown in the third row of Fig. 6, the skeletonization processing is then performed on the ten base judgment images to determine the smoke root candidate points, obtain the skeletons, and extract the skeleton endpoints. For the skeleton endpoint in each set of endpoints, the skeleton endpoint stacking search strategy is applied to obtain the integer set of search results corresponding to the skeleton endpoints on

each image frame. Such a stacking strategy superpositions the endpoints of the ten frames of skeleton endpoint images on a single channel model image through coordinate projection to obtain a model image with all skeleton endpoints. This stacking strategy is applied to each endpoint on each frame on the projected model image to obtain the statistical number of sets. A search strategy is then used to search the model image with $R = 5$ as the radius for the search point coordinates (i, j) . Whenever a pixel is found, the search count is increased by 1, and a two-dimensional integer group $M[N][m]$, the stacking number of pixels of the m th skeleton endpoint of the N th frame, is obtained, where N represents

the Nth frame and is between one and ten and m is the end-point meeting the searching criteria with an upper limit of the number of skeleton endpoints in the Nth frame.

With the search strategy applied, the dispersion degree of the ten integer sets is then calculated corresponding to the ten frames of images, and the skeleton frame with the largest dispersion factor is selected as the base frame. Considering that a higher probability of the correct smoke roots occurs on the frame of the image with the greatest dispersion factor, the stacking and searching process for the skeleton endpoints is repeated on the base frame to obtain the integer set of searched skeleton endpoints of the base frame. The pixels whose search results are greater than the threshold (equals 5 in this study) are identified as a smoke root with higher confidence, which indicates that the points searched in at least the threshold times of frames of skeleton endpoint images. These detected smoke roots, obtained by further screening, represent the source points of the smoke. These identified smoke root pixels are labeled for further smoke detection analysis if needed.

Datasets

To validate the new smoke detection algorithm, publicly available video datasets from forest fires were used to perform the smoke roots detection, including eighteen smoke videos with a resolution of 480×320 , and a frame rate of 25 frames per second. The input video materials used can be downloaded via [**CHENG DATASET \(2021\)](#). The new algorithm was implemented using C++ in Visual Studio 2015 and OpenCV 4.3.0 to detect the fire and to visualize the predicted image based on the smoke eigenvalue generated in the process. The results were compared with the detection obtained from the traditional ViBe algorithm from Gao and Cheng (2019) to evaluate the effectiveness of the new algorithm. The CPU used was the Intel Core i7-10710U and the GPU was the MX250.

Results

Parameter settings

To compare with the smoke root detection results from Gao and Cheng (2019), the same parameters were set for the ViBe extraction of smoke candidate points. For the new method developed in this study, the adaptive threshold (AT) is an empirical value (Table 1). Table 2 also shows the smoke detection results based on the Choquet fuzzy integral algorithm developed by Wang et al. (2017), which detected smoke rather than smoke roots. Therefore, to compare the results obtained by Wang et al. (2017), the Choquet fuzzy integral algorithm was used to extract the smoke candidate

Table 1 Adaptive threshold used in the new developed algorithm

Video	AT	Video	AT	Video	AT
T1	0.20	T7	0.20	T13	0.15
T2	0.20	T8	0.20	T14	0.20
T3	0.20	T9	0.15	T15	0.20
T4	0.20	T10	0.05	T16	0.15
T5	0.15	T11	0.15	T17	0.20
T6	0.20	T12	0.20	T18	0.18

Table 2 Forest fire smoke detection results by Wang et al. (2017) (“N” = not detected; “Y” = successfully detected)

Video	Detect successfully	Video	Detect successfully	Video	Detect successfully
T1	Y	T7	N	T13	N
T2	N	T8	Y	T14	N
T3	N	T9	Y	T15	Y
T4	Y	T10	N	T16	N
T5	N	T11	Y	T17	N
T6	Y	T12	N	T18	N

regions followed by the smoke roots extraction method in this study. Further details of comparisons between the smoke roots detected based on the Choquet fuzzy integral algorithm and the new algorithm are presented later.

Since image detection and fusion is a pixel-level process, a large number of isolated points need to be included in the connected domain. To reduce computation needs, this study applied erosion and expansion operations on the smoke foreground pixels from Gao and Cheng (2019). Table 3 lists the expansion and erosion cores used in this paper. In the smoke root search and detection process, the search radius R was set at 5.

Table 4 shows the adjusted adaptive thresholds set by the new method to obtain the base judgement image database and the number of frames required to successfully detect smoke root candidate points, which are the time and the number of detected candidate smoke roots. When a proportion of the acquired connected domain is greater than the adjusted AT compared to the previous frame of the judgement image, the next connected domain frame is extracted as the frame of judgment image of the smoke roots search strategy. The selection of different thresholds for different samples depends on the inconsistent diffusion speeds of the smoke for different samples. The connected domains of samples with slower smoke diffusion change slowly, while for the faster smoke diffusion change quickly. For slow smoke, a relatively small frame extraction threshold is needed to have enough search frames during the ten seconds of the sample video. For samples with fast smoke diffusion, their

Table 3 Smoke foreground pixels via Gao and Cheng (2019) (“N”=not detected; “Y”=successfully detected)

Test video	First available frame	Close/Erode/Dilate Kernel	Roots number	Detected successfully
T1	81	21*21 3*3 21*21	0	N
T2	65	21*21 1*1 25*25	0	N
T3	19	21*21 3*3 21*21	2	Y
T4	80	21*21 1*1 29*29	2	N
T5	46	21*21 3*3 31*31	2	N
T6	49	21*21 1*1 21*21	1	Y
T7	39	21*21 5*5 25*25	0	N
T8	17	21*21 3*3 21*21	0	N
T9	52	21*21 3*3 29*29	2	Y
T10	68	21*21 3*3 29*29	2	Y
T11	76	21*21 1*1 21*21	0	N
T12	52	21*21 3*3 25*25	2	N
T13	15	21*21 3*3 21*21	0	N
T14	15	21*21 7*7 25*25	0	N
T15	213	21*21 9*9 35*35	2	N
T16	80	21*21 5*5 25*25	2	N
T17	20	21*21 3*3 21*21	1	Y
T18	67	21*21 7*7 31*31	2	Y

Table 4 Adjusted adaptive thresholds (“N”=not detected; “Y”=successfully detected)

Test video	Roots number	Detected frame	Adjusted AT	Detected successfully
T1	4	184	0.20	Y
T2	1	203	0.20	Y
T3	1	151	0.20	Y
T4	2	127	0.20	Y
T5	1	245	0.15	Y
T6	2	60	0.20	Y
T7	1	57	0.20	Y
T8	1	47	0.20	Y
T9	1	206	0.15	Y
T10	1	125	0.05	Y
T11	1	194	0.15	N
T12	3	175	0.20	Y
T13	3	195	0.15	Y
T14	1	164	0.20	Y
T15	1	157	0.20	Y
T16	2	213	0.15	Y
T17	1	129	0.20	Y
T18	2	217	0.18	Y

connected domains change at a faster rate, and a smaller frame extraction threshold is required to make the time span of extracting search frames long enough. For smoke with a

moderate diffusion rate, the search frame is extracted with a 20% frame extraction threshold to ensure that the sample length can extract sufficiently long time span of frames and the sufficiently large inter-frame differences.

Adjusting the adaptive thresholds does not affect the acquisition of correct smoke roots but instead it will reduce the detection of fake smoke roots. The T2 sample was analyzed as an example using different ATs to obtain the smoke roots (Table 5). It shows that a larger AT requires more frames obtain ten frames of the judgment image and fewer smoke root candidate points are obtained. A smaller AT requires fewer frames to obtain the ten frames of judgment image and a greater number of smoke root candidate points are obtained. Since larger AT need longer stacked search to reach the degree of change in the connected domain, fewer “fake smoke roots” are considered as candidate points. Although the AT does not affect whether the smoke roots are detected or not, too many detection frames will increase the amount of calculation to obtain the frames of judgment image in the early stage, and more will also increase the amount of calculation in the later two-dimensional smoke simulation process. Thus, to be computationally efficient with fewer fake smoke root detection, the AT is adjusted to an appropriate value (Table 4).

Performance evaluation of the new algorithm

Two indicators used to evaluate the performance of the algorithm, include detection accuracy and the number of effective connected domains. To quantify the detection accuracy, an artificial ROI (region of interest) area was set in the smoke source area of the image as the reference smoke roots, and used to compare whether the detected smoke roots by different detection methods were correct. The ROI areas for each forest fire event were marked with a red frame in the image frame, and the coordinates of these areas were stored for later comparison. If the detected smoke roots candidate points fell in the ROI area, the detection was considered successful, which was also the detection accuracy index proposed by Gao and Cheng (2019). The shape of the ROI area is rectangular but the size for different fires are different

Table 5 Smoke root detection of different ATs

Test video	AT	Detected Frame	Roots number
T2	0.25	255	1
T2	0.20	203	1
T2	0.15	152	3
T2	0.10	96	3

(Table 6). Among the available 27 videos, 18 showed a forest fire.

Table 7 compares the detection accuracy of three different smoke detection algorithms, including the ones proposed by

Table 6 Types of smoke, number of frames, and the ROI sizes in the 18 tests videos with fire

Test video	Type	Frame	ROI size
T1	Mountain smoke	269	20*20
T2	Mountain smoke	269	20*20
T3	Early smoke on flat ground	269	20*20
T4	Early smoke on flat ground	263	20*20
T5	Smoke from residential areas	258	33*33
T6	Low-definition forest smoke	266	20*20
T7	Smoke surrounded by intrusive colors	271	20*60
T8	Smoke surrounded by intrusive colors	271	20*60
T9	Smoke with low definition	269	20*20
T10	Smoke with low definition	264	20*20
T11	Smoke partially obscured by trees	250	60*60
T12	Vegetable field smoke	250	20*20
T13	Vegetable field smoke	253	20*20
T14	Long-distance residential area smoke	250	20*20
T15	Obstructed smoke	251	22*22
T16	Double ignition point smoke	256	22*22
T17	Terrace smoke	252	17*17
T18	Smog with a tendency to spread	250	90*30

Table 7 Comparison of detection results between different algorithms (“N” = not detected; “Y” = successfully detected)

Videos	Algorithm used by Gao and Cheng (2019)		Algorithm used by Wang et al. (2017)		The new algorithm	
	Roots	Successfully	Roots	Successfully	Roots	Successfully
T1	0	N	5	N	4	Y
T2	0	N	4	Y	1	Y
T3	2	Y	2	Y	1	Y
T4	2	N	9	Y	2	Y
T5	2	N	6	N	1	Y
T6	1	Y	2	Y	2	Y
T7	0	N	7	Y	1	Y
T8	0	N	12	Y	1	Y
T9	2	Y	9	Y	1	Y
T10	2	Y	5	N	1	Y
T11	0	N	2	N	1	N
T12	2	N	0	N	3	Y
T13	0	N	0	N	3	Y
T14	0	N	0	N	1	Y
T15	2	N	1	Y	1	Y
T16	2	N	8	Y	2	Y
T17	1	Y	2	Y	1	Y
T18	2	Y	17	Y	2	Y
Total		31.6%		52.6%		94.9%

Gao and Cheng (2019) and Wang et al. (2017), and the one developed in this study. The results were obtained using the foreground extraction algorithms of each method to obtain smoke candidate regions, followed by extracting the smoke root candidate points using the same method and parameters. From Table 7, it is shown that the new algorithm can detect smoke roots with an accuracy of 94.9%, a significant improvement compared to those of Gao and Cheng (2019) and Wang et al. (2017) algorithms of 31.6 and 52.6%, respectively.

In addition to the detection accuracy, the fusion image of the binarized smoke candidate regions can also be used to qualitatively validate the effectiveness of the new algorithm. Figure 7, compare the original image (first row), the dynamic and static feature fusion image obtained using the new algorithm (second row), the results from the ViBe algorithm (third row), and the results from the Choquet fuzzy integration algorithm (fourth row). The new algorithm can identify smoke root areas more accurately and more completely with less false smoke roots compared to the other traditional algorithms.

To quantify the effectiveness of different algorithms (Fig. 8), the number and proportion of effective connected domain frames can be used as the second validation index. These are defined as the connected domain that can obtain the skeleton endpoints of the correct smoke root candidate points after skeletonization. An example of ineffective and effective connected domain frames, obtained using different



Fig. 7 Comparison between the original images (first row), the fusion images from the developed method in this study (second row), from the ViBe algorithm (third row), and from the Choquet algorithm (fourth row) for T1, T5, T6, T11

methods in the same frame is shown in Fig. 9. The frame on the left is considered ineffective since it cannot be directly refined through the skeleton refinement, and the correct smoke roots cannot be effectively extracted through the smoke roots search strategy. Thus, the ineffective connected domain frames, defined as the foreground frames, require preprocessing to obtain a complete connected domain through opening and closing operations or as the frames that cannot obtain a complete connected domain through basic opening and closing operations. The connected domain on the right of Fig. 9 is effective since it can be directly processed through skeletonization to obtain continuous smoke root candidate points that can effectively extract the correct smoke roots.

With the effective connected domain frames defined, the positive detection (PD) and the detection rate (DR) are used to compare the detection effectiveness between different algorithms.

Table 8 compares the positive rate and the detection rate for the three different algorithms including those of Gao and Cheng (2019), Wang et al. (2017), and the developed algorithm in this study. It illustrates that the new algorithm

improves significantly on detection effectiveness of the smoke roots, with detections rate of 100% for most of the test video and minimum detection rate of 73.3%, compared to a minimum detection rate of 11.7 and 0% for those of Gao and Cheng (2019) and Wang et al. (2017), respectively.

Discussion

To further investigate the possible influencing factors for the detection accuracy of different smoke root detection algorithms, Figs. 10, 11, 12 and 13 show the number of detected true roots and false detection for the three algorithms. Comparing these figures, shows that the proposed new method not only improves the detection rate of true smoke roots, but can detect smoke roots in some challenging environments which the other two algorithms cannot. The test videos T1 and T13 (Figs. 7, and 8) were selected as examples to further explain such conditions. In T1 and T13, there was regular smoke in both scenes and a few light-colored buildings in the background. Under such conditions, the ViBe algorithm failed to extract any smoke roots because

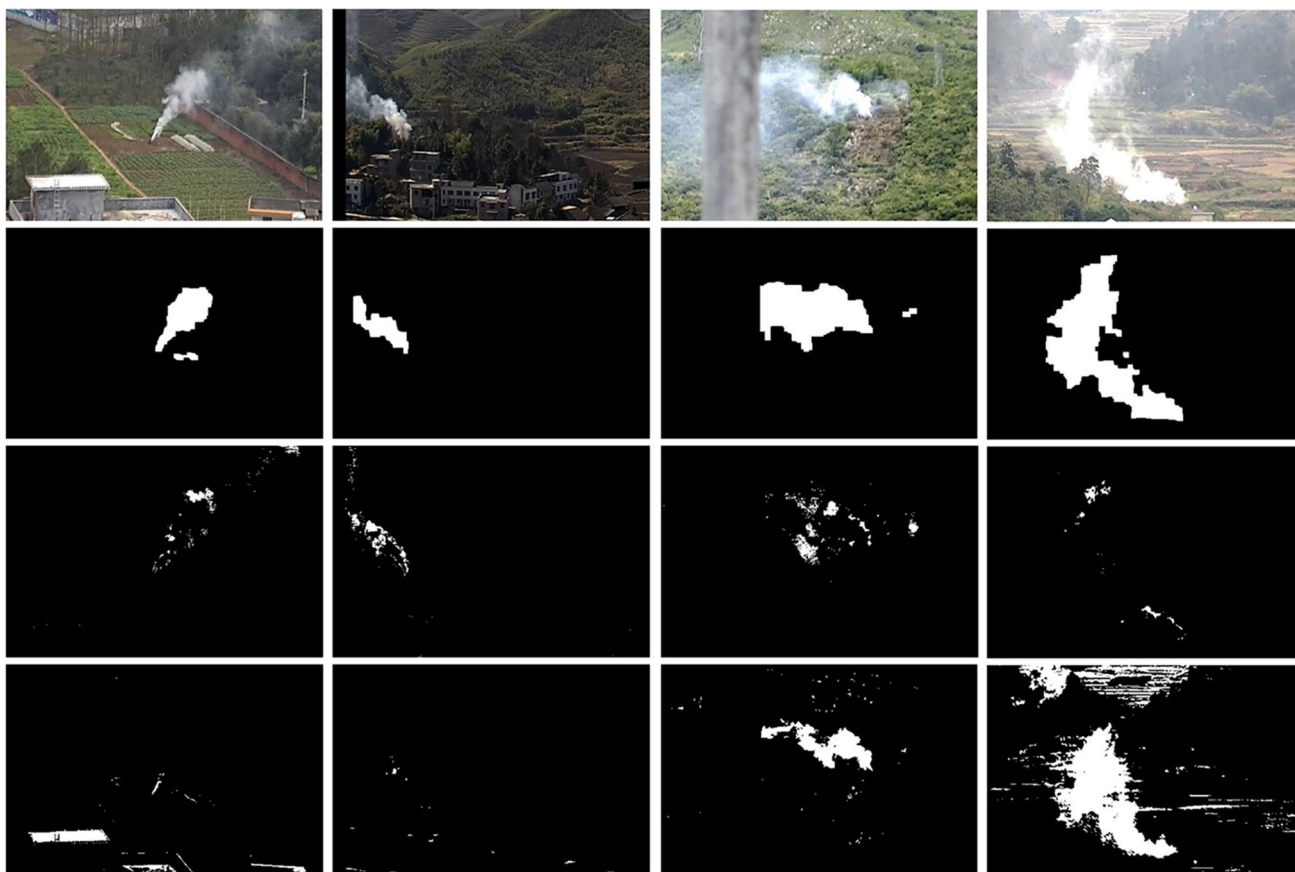


Fig. 8 Comparison between the original images (first row), the fusion images from the developed method in this study (second row), from the ViBe algorithm (third row), and from the Choquet algorithm (fourth row) for T13, T14, T16, T18

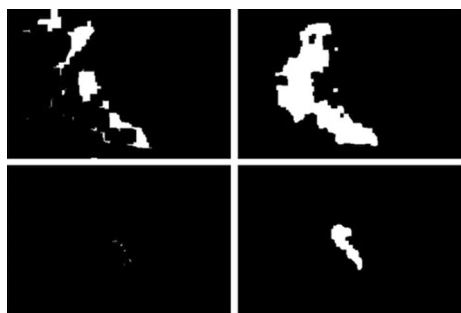


Fig. 9 Ineffective (left) and effective (right) connected domain frames

it extracted foreground information based on differences between front and rear frames. The high grayscale pixels in the center of the smoke mixed with the background area and were missed, and the edge of the smoke would not be detected due to the slight change in color. Therefore, the thin smoke and the background in the central area of the smoke that maintained a stable attitude, induced interference to the ViBe algorithm. The diffuse and thin smoke displayed by

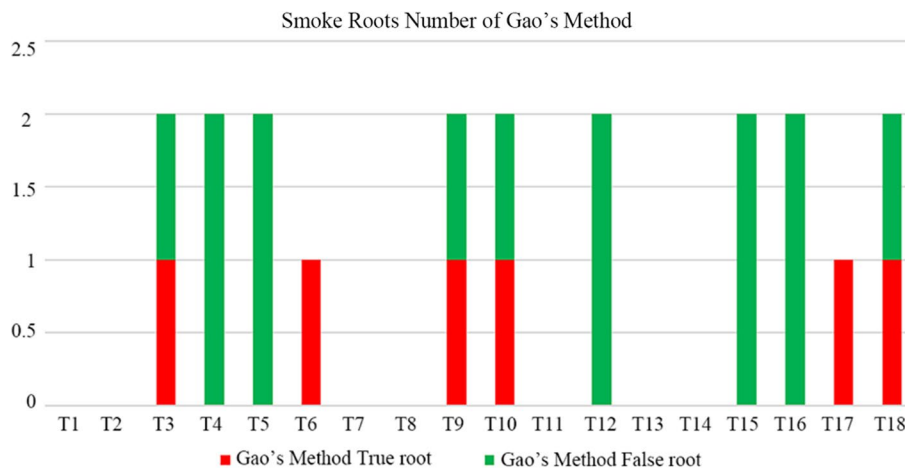
T5 (Fig. 7) also showed few smoke regions using the ViBe algorithm, verifying this finding. The algorithm developed in this study further integrates static features based on the ViBe algorithm, so that the pixels in the smoke regions that do not conform to the characteristics of ViBe can also enter the fusion image by conforming to the static characteristics of the smoke. Therefore, under the same circumstances, the proposed method in this paper can extract the smoke regions more completely.

Wang et al. (2017) Choquet fuzzy integral algorithm could not successfully extract the smoke roots in some of the scenarios because it was significantly influenced by background interference. The light-colored background object in T13 interfered considerably with the Choquet fuzzy integral algorithm, making it unable to distinguish the light-colored background from the smoke regions. The resulting foreground smoke regions were too far away from the correct smoke regions as shown in T13 (Fig. 8). The algorithm developed in this paper, on the other hand, used a dynamic area frame containing all pixels to set the extraction range of the smoke regions after obtaining the dynamic pixel set. Such a procedure significantly reduces interference outside

Table 8 Comparison of positive rate and detection rate for effective connected domain frames between the three different algorithms

Video	Smoke/ROI (frame)	Available frame of GAO and Cheng (2019) method		Available Frame of Wang et al. (2017) method		Available frame of proposed method	
		PD	DR	PD	DR	PD	DR
T1	269	32	11.7%	28	10.4%	218	80.9%
T2	269	203	75.5%	105	39.03	269	100.00%
T3	269	103	38.29	251	93.31%	269	100.00%
T4	263	185	70.3%	263	100%	257	97.77%
T5	258	86	33.3%	27	10.47%	189	73.26%
T6	266	217	81.6%	127	47.74%	266	100.00%
T7	271	232	85.6%	271	100%	271	100.00%
T8	271	160	59.0%	271	100%	271	100.00%
T9	269	217	80.7%	269	100%	269	100.00%
T10	264	133	50.4%	1	0.38%	264	100.00%
T11	250	173	69.2%	174	69.60%	250	100.00%
T12	250	173	58.8%	226	90.40%	250	100.00%
T13	253	238	94.1%	0	0%	253	100.00%
T14	250	235	94.0%	0	0%	253	100.00%
T15	251	39	15.5%	119	47.41%	251	100.00%
T16	256	111	43.4%	256	100%	256	100.00%
T17	252	233	92.5%	252	100%	252	100.00%
T18	250	183	73.2%	250	100%	250	100.00%

Fig. 10 Smoke root numbers by Gao and Cheng (2019) method



the dynamic area. However, if the interference is too close to the smoke source contained in the dynamic area frame, it would still affect the detection accuracy of the algorithm. Since such interference does not change the basic outline of the smoke regions, it will be reflected as additional false smoke roots and does not affect the detection results of the correct smoke roots.

In addition, when comparing the detection results of T6, T14, and T18 (Figs. 7, 8, 10, 11, 12) it can be seen that the overall brightness of the environments where the smoke is located had a greater impact on the extraction of the foreground area using the Choquet fuzzy integral algorithm. In

T18 (Fig. 8), both the outline details of the smoke and the brighter background were extracted at the same time. In T14 (Fig. 8), although the dynamic and static characteristics of the smoke were obvious, only a few pixels were successfully extracted, resulting in an unstable smoke region. The contrasting T6 (Fig. 8) had a low pixel rate, but the overall area was brighter. Under the premise that the presence of fog made the smoke more difficult to distinguish, using the Choquet fuzzy integration algorithm, the smoke near the smoke source contour can still be extracted.

In Figs. 10, 11 and 12 for T11, none of the algorithms can detect the correct smoke roots. Looking at Fig. 8, T11

Fig.11 Smoke roots number of Wang et al. (2017) method

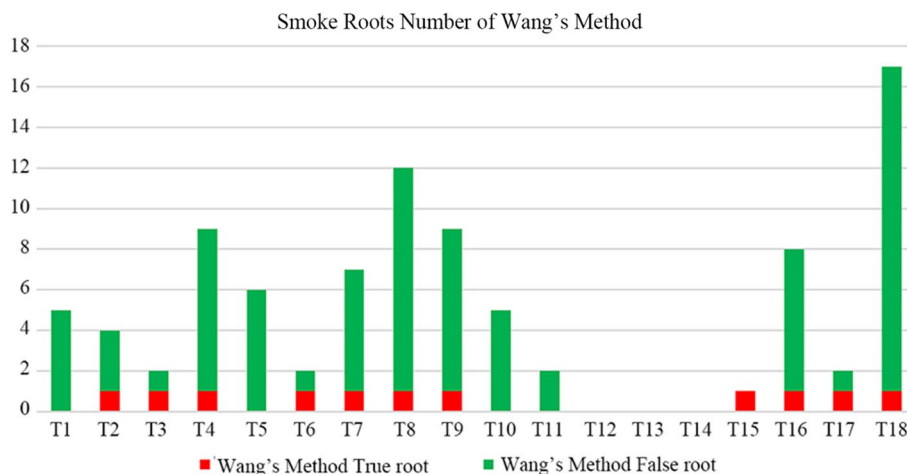


Fig. 12 Smoke roots number of proposed method

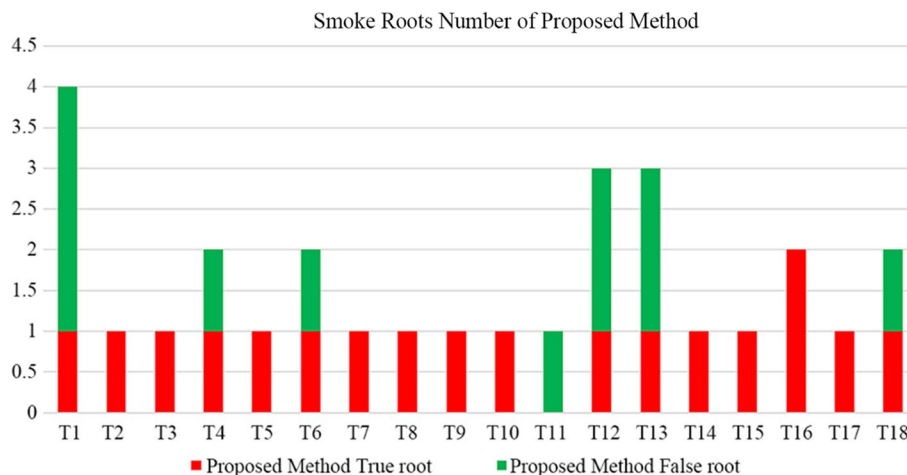
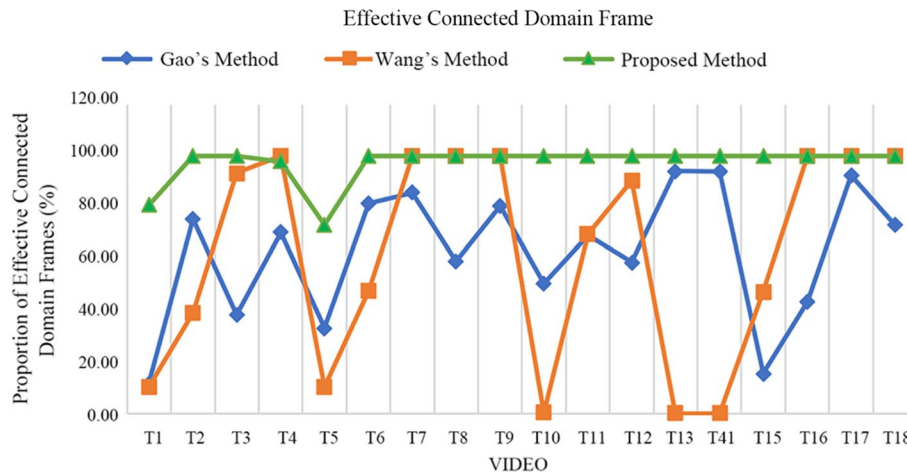


Fig. 13 The proportion of effective connected domain frame in the total frame proportion



might have a problem with the scene itself, because the smoke roots were not obvious and the portion of the smoke covered by trees was almost equal to the portion that can be observed, which made the delineation of the ROI area doubtful.

From Figs. 10, 11 and 12, it also can be seen that compared with the Wang's method, the algorithm developed in this study greatly reduced the number of fake smoke roots, which can significantly reduce the workload in the stage of verifying smoke roots candidate points by generating

simulated smoke using the two-dimensional smoke simulation engine. For example, T16 (Fig. 8) was a video of forest fire smoke with two smoke sources. Figure 11 shows that the Gao and Cheng (2019) method detected two false smoke root candidate points, Fig. 11 shows that the Wang et al. (2017) method falsely detected more than six smoke root candidate points, and Fig. 12 indicates that the developed method detected all smoke root candidate points correctly. The smoke regions obtained by Wang et al. (2017) was a smoke profile with a certain concentration, which was the core diffusion trajectory of the smoke, and its changing shape and speed were relatively slow. The second smoke source on the right of T16 just began to smoke and only a small smoke moved. Since the Wang et al. (2017) method requires a certain minimum smoke concentration to be detected, the second source of smoke on the right of T16 could not be detected using their method. By using the method developed in this paper, the smoke regions showed the contour of diffuse smoke, and its diffusion shape and speed were faster. Therefore, under similar conditions as in T16, the method developed in this study can obtain fewer candidate smoke roots for a higher quality of smoke root detection. It also can extract the smaller smoke source on the right of T16 by extracting its static features, and both smoke sources in T16 can be detected correctly. Also, when comparing smoke videos of T18 and T16 (Fig. 8), in addition to the difference in smoke contours, there was also significant background interference in T18. For T18, Figs. 10, 11 and 12 indicate that the methods of Gao and Cheng (2019), Wang et al. (2017) and the method developed in this study detected one, sixteen, and one false smoke root candidate points, respectively. Background interference that does not change with the frame will also increase the number of fake smoke root candidate points as in the Wang et al. (2017) method, which was addressed by the developed method. Although the number of smoke root candidate points obtained by Gao and Cheng (2019) is smaller, it detected far fewer correct smoke roots as indicated clearly by comparing Figs. 10, 11 and 12.

In addition to the detection accuracy, the effective connected domain is another evaluation indicator for different algorithms and refers to the complete and connected domain that can be obtained through the skeletonization algorithm to reflect the correct smoke roots. Figure 13 compares the efficiency of the connected domain frames for the three investigated methods. The efficiency of the connected domain frames obtained by the method developed method in this study is an improvement compared with the other two methods.

The proportion of the effective connected domain for all the 18 test videos varied significantly between different videos ranging from 20 to 80% (Fig. 13). The method of Gao and Cheng (2019) obtained the smoke foreground images directly through the ViBe algorithm but could not effectively

obtain the connected domain to be skeletonized when processing part of the samples. For their method, for different test videos, the process of preprocessing the images into connected domains requires separate adjustment of parameters for each sample to obtain a good, connected domain (Table 3). Even after tuning each sample individually to have better performance, the proportion of effective connected domain frames was still low. In addition, the background database update and establishment process of the ViBe algorithm also makes it difficult to obtain the correct smoke regions image for the first frame. Since the follow-up smoke roots search strategy of their method needs to operate on the connected domain of five consecutive frames, there are also relatively high requirements for the continuity of the effective connected domain. This makes the proportion of consecutive connected domain frames that can obtain the correct smoke root candidate points smaller, which directly lowers the success rate of the method.

For the method of Wang et al. (2017), a proportion of the effective connected domain for all the 18 test videos varied even larger to 100% (Fig. 13). Their method obtained images of the smoke regions through the Choquet fuzzy integration algorithm. The effectiveness of this algorithm is suitable for different, but not all, smoke environments. For inapplicable conditions, the effective connected domain frame proportion is low to 0%, and for applicable conditions, the proportion is up to 100%. The Choquet algorithm is stable processing inter-frame effects, unlike the ViBe algorithm, and integrates and improves as processing continues.

The method developed in this study can provide the proportion of the effective connected domain for all 18 test videos between 80 and 100% with small variations. The method established a model to frame the dynamic area of the smoke after obtaining the ViBe smoke foreground image and extracted the static feature area from such a dynamic area. This process enabled the method to obtain good, connected domains that can be skeletonized into effective skeletons for different sample videos. More importantly, the subsequent adaptive threshold extraction of the skeleton frame strategy, and the multi-frame discrete confidence determination strategy, did not require the continuity of the connected domain. Using this method, high-confidence smoke root candidate points can still be obtained even when with a few ineffective frames, effectively reducing the probability of a missed detection.

Conclusion

In this study, a new smoke root detection method was developed to extract smoke regions and smoke roots effectively. The new method retains the dynamic feature smoke regions obtained through the ViBe algorithm, uses the ViBe feature

image to establish a model image to determine the smoke dynamic area frame, and obtains the smoke regions through the static feature extraction method. Such a process solves the problem of missing the detection of the smoke regions induced by voids or empty spaces. The smoke dynamic area frame excludes areas outside the smoke regions for a better outline and shape, and provides a solution to detect smoke more accurately under complex conditions. In addition, the smoke region images obtained are extracted by the adaptive threshold in real time. Ten frames of smoke connected region images that meet the extraction conditions are used for skeletonization, followed by endpoint extraction and a multi-frame discrete confidence determination strategy to extract smoke root candidate points. Such a process will significantly improve the detection accuracy and rate of smoke roots. Based on the analysis of 18 test videos and a comparison between three different smoke detection methods, including those of Gao and Cheng (2019), Wang et al. (2017), and the one developed in this study, the following conclusions can be made:

- (1) The results show that the minimum detection rate based on the new model significantly increased to 94.9%.
- (2) The proportion of effective connected domain in smoke frames can be increased to 100% and the new model under similar environmental conditions achieved similar improvement. Even if the Choquet algorithm is used to obtain the smoke regions, the detection rate increased from 38.9 to 52.6% using the judgment strategy developed in this study.
- (3) The new method also requires fewer smoke root candidate points for more accurate detection, resulting in higher computational efficiency in generating simulated smoke for the two-dimensional smoke simulation engine to verify the smoke root candidate points.

The new method can solve the challenges of reducing the impacts of smoke voids and background interference in the extraction of the smoke regions, and can be widely applied for forest fire detection.

References

- Alamgir N, Nguyen K, Chandran V, Boles W (2018) Combining multichannel color space with local binary co-occurrence feature descriptors for accurate smoke detection from surveillance videos. *Fire Saf J* 102:1–10. <https://doi.org/10.1016/j.firesaf.2018.09.003>
- Appana DK, Islam R, Khan SA, Kim JM (2017) A video-based smoke detection using smoke flow pattern and spatial-temporal energy analyses for alarm systems. *Inf Sci* 418–419:91–101. <https://doi.org/10.1016/j.ins.2017.08.001>
- Barnpoutis P, Dimitropoulos K, Grammalidis N (2013) Real time video fire detection using spatio-temporal consistency energy. *IEEE AVSS 2013*. <https://doi.org/10.1109/AVSS.2013.6636667>
- Cheng SH, Ma JY, Zhang SJ (2019) Smoke detection and trend prediction method based on deeplabv3+ and generative adversarial network. *J Electron Imaging* 28:033006–033006. <https://doi.org/10.1117/1.JEI.28.3.033006>
- Gao Y, Cheng PL (2021) Full-scale video-based detection of smoke from forest fires combining ViBe and MSER algorithms. *Fire Technol*. <https://doi.org/10.1007/s10694-020-01052-3>
- Gao Y, Cheng PL (2019) Forest fire smoke detection based on visual smoke root and diffusion model. *Fire Technol* 55(5):1801–1826. <https://doi.org/10.1007/s10694-019-00831-x>
- Gaur A, Singh A, Kumar A, Kapoor K (2020) Video flame and smoke-based fire detection algorithms: a literature review. *Fire Technol* 56(5):1943–1980. <https://doi.org/10.1007/s10694-020-00986-y>
- He LJ, Gong XL, Zhang SR, Wang LJ, Li F (2021) Efficient attention based deep fusion CNN for smoke detection in fog environment. *Neurocomputing* 434:224–238. <https://doi.org/10.1016/j.neucom.2021.01.024>
- Jadon A, Omama M, Varshney A, Ansari MS, Sharma R (2019) FireNet: a specialized lightweight fire & smoke detection model for real-time iot applications, CVPR (cs.CV) (May 2019). <https://arxiv.org/abs/1905.11922>
- Li CH, Yang B, Ding H, Shi HL, Jiang XP, Sun J (2020) Real-time video-based smoke detection with high accuracy and efficiency. *Fire Saf J* 117:103184. <https://doi.org/10.1016/j.firesaf.2020.103184>
- Nguyen VT, Quach CH, Pham MT (2020) Video smoke detection for surveillance cameras based on deep learning in indoor environment. In: 2020 4th International Conference on Recent Advances in Signal Processing, Telecommunications Computing (SigTelCom), pp. 82–86. <https://doi.org/10.1109/SigTelCom49868.2020.9199056>
- Piccinini P, Calderara S, Cucchiara R (2008) Reliable smoke detection in the domains of image energy and color. In: 2008 15th IEEE International Conference on Image Processing, pp. 1376–1379. <https://doi.org/10.1109/AVSS.2013.6636667>
- Prema CE, Vinsley SS, Suresh S (2016) Multi feature analysis of smoke in YUV color space for early forest fire detection. *Fire Technol* 52(5):1319–1342. <https://doi.org/10.1007/s10694-016-0580-8>
- Savci MM, Yildirim Y, Saygili G, Toreyin BU (2019) Fire detection in h.264 compressed video. *ICASSP 2019*:8310–8314. <https://doi.org/10.1109/ICASSP.2019.8683666>
- Sousa JVR, Gamboa PV (2020) Aerial forest fire detection and monitoring using a small UAV. *KEG* 5:242–256. <https://doi.org/10.18502/keg.v5i6.7038>
- Wang SD, He YP, Yang HY, Wang KX, Wang J (2017) Video smoke detection using shape, color, and dynamic features. *J Intell Fuzzy Syst* 33:305–313. <https://doi.org/10.3233/JIFS-161605>
- Wang SW, Yu F, Zhou CL, Jiang MH (2020) Straw burning detection method based on improved frame difference method and deep learning. *IEEE ICIVC* pp. 29–33. <https://doi.org/10.1109/ICIVC50857.2020.9177456>
- Wu XH, Cao YC, Lu XB, Leung H (2021) Patchwise dictionary learning for video forest fire smoke detection in wavelet domain. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-020-05541-y>
- Xia X, Yuan FN, Zhang L, Yang LZ, Shi JT (2019) From traditional methods to deep ones: review of visual smoke recognition, detection, and segmentation. *Int J Image Graph* 24(10):1627–1647. <https://doi.org/10.11834/jig.190230>
- Xu G, Zhang YM, Zhang QX, Lin GH, Wang Z, Jia Y, Wang JJ (2019) Video smoke detection based on deep saliency network. *Fire Saf J* 105:277–285. <https://doi.org/10.1016/j.firesaf.2019.03.004>
- Ye W, Zhao JH, Wang S, Wang Y, Zhang DY, Yuan ZY (2015) Dynamic texture-based smoke detection using surfacelet transform and HMT model. *Fire Saf J* 73:91–101. <https://doi.org/10.1016/j.firesaf.2015.03.001>

- Ye SP, Bai ZC, Chen HF, Bohush R, Ablameyko S (2017) An effective algorithm to detect both smoke and flame using color and wavelet analysis. *Pattern Recognit Image Anal* 27(1):131–138. <https://doi.org/10.1134/S1054661817010138>
- Yuan FN, Fang ZJ, Wu SQ, Yang Y, Fang YM (2015) Real-time image smoke detection using staircase searching-based dual threshold AdaBoost and dynamic analysis. *IET Image Process* 9(10):849–856. <https://doi.org/10.1049/iet-ipr.2014.1032>
- Yuan C, Liu ZX, Zhang YM (2017a) Aerial images-based forest fire detection for firefighting using optical remote sensing techniques and unmanned aerial vehicles. *J Intell Robot Syst* 88(2–4):635–654. <https://doi.org/10.1007/s10846-016-0464-7>
- Yuan FN, Xia X, Shi JT, Li HD, Li G (2017b) Non-linear dimensionality reduction and Gaussian process-based classification method for smoke detection. *IEEE Access* 5:6833–6841. <https://doi.org/10.1109/ACCESS.2017.2697408>
- Yuan C, Liu ZX, Zhang YM (2018) Learning-based smoke detection for unmanned aerial vehicles applied to forest fire surveillance. *J Intell Robot Syst* 93(1–2):337–349. <https://doi.org/10.1007/s10846-018-0803-y>
- Zhang TY, Suen CY (1984) A fast parallel algorithm for thinning digital patterns. *Commun ACM* 27(3):236–239. <https://doi.org/10.1145/357994.358023>
- Zhao YQ, Zhou Z, Xu MM (2015) Forest fire smoke video detection using spatiotemporal and dynamic texture features. *JECE* 2015:1–7. <https://doi.org/10.1155/2015/706187>
- Zhao N, Wang XW, Yin SL (2021) Research of fire smoke detection algorithm based on video. *IJEIJ* 13:1–9. [https://doi.org/10.6636/IJEIE.202103_13\(1\).01](https://doi.org/10.6636/IJEIE.202103_13(1).01)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.