

Article

Smoke Object Segmentation and the Dynamic Growth Feature Model for Video-Based Smoke Detection Systems

Md Rashedul Islam ^{1,*}, Md Amiruzzaman ², Shahriar Nasim ³ and Jungpil Shin ^{1,*}

¹ School of Computer Science and Engineering, University of Aizu, Fukushima 965-0006, Japan

² College of Aeronautics and Engineering, Kent State University, Kent, OH 44242, USA; mamiruzz@kent.edu

³ Department of Computer Science and Engineering, University of Asia Pacific, Dhaka 874-8577, Bangladesh; 14101096@uap-bd.edu

* Correspondence: rashed.cse@gmail.com (M.R.I.); jpshin@u-aizu.ac.jp (J.S.)

Received: 7 June 2020; Accepted: 29 June 2020; Published: 30 June 2020



Abstract: This article concerns smoke detection in the early stages of a fire. Using the computer-aided system, the efficient and early detection of smoke may stop a massive fire incident. Without considering the multiple moving objects on background and smoke particles analysis (i.e., pattern recognition), smoke detection models show suboptimal performance. To address this, this paper proposes a hybrid smoke segmentation and an efficient symmetrical simulation model of dynamic smoke to extract a smoke growth feature based on temporal frames from a video. In this model, smoke is segmented from the multi-moving object on the complex background using the Gaussian's Mixture Model (GMM) and HSV (hue-saturation-value) color segmentation to encounter the candidate smoke and non-smoke regions in the preprocessing stage. The preprocessed temporal frames with moving smoke are analyzed by the dynamic smoke growth analysis and spatial-temporal frame energy feature extraction model. In dynamic smoke growth analysis, the temporal frames are segmented in blocks and the smoke growth representations are formulated from corresponding blocks. Finally, the classifier was trained using the extracted features to classify and detect smoke using a Radial Basis Function (RBF) non-linear Gaussian kernel-based binary Support Vector Machine (SVM). For validating the proposed smoke detection model, multi-conditional video clips are used. The experimental results suggest that the proposed model outperforms state-of-the-art algorithms.

Keywords: smoke detection; pattern recognition; background subtraction; Gaussian's mixture model (GMM); HSV color segmentation; smoke growth features; support vector machine (SVM)

1. Introduction

Fire accidents cause a great impairment to human life, the economy, the environment, and ecology. Detecting a fire in its early stage can prevent mass destruction and save thousands of lives and valuable assets. Smoke components represent the beginning of the early stage of a fire and a forewarning to possibly catastrophic incidents. Thus, detecting smoke at an early stage can potentially provide crucial information to prevent a fire event as well as minimize the damage and, consequently, save lives and properties.

The conventional fire alarm system is usually built with a point sensor, which works in the manner of heat transactions through the sensor. While a small area can be covered using this technique, a large area would still remain vulnerable to fire. This occurs because the smoke propagates in various directions and takes time to reach the sensor in a big area, which ultimately fails to give a timely warning. Besides, all sorts of interference in the sensor could delay the alarm or raise a false alarm [1].

Nowadays, the uses of different sensors and multimedia data processing have a great impact on develop smart homes, smart cities, smart industries, smart hospitals, smart agriculture, and others. Internet of multimedia Things (IoMT) devices and data processing is much more popular, especially video image processing to monitoring system [2]. As a part of the development of a smart home or industry, a video image-analysis-based smart fire alarm system is demandable. With increased use of modern technology and security concerns, high-resolution video cameras are available at the house, parking place, playground, office, industry, and even on the road. The video cameras have a high frame rate, small response time, and computational power with respect to smoke sensors, which can provide cost-effective solutions by covering wide areas. Moreover, it is easy to embed high functional and robust video image processing smoke alarm systems, which might reduce high false alarm.

Considering the above-mentioned limitation and advantage of a video camera, a video frame processing-based fire detection system has been developed in the last few years. Its effectiveness in the detection of a fire quickly made it popular. Relatively wider area coverage, better accuracy, and a lower rate of false alarms made it more robust and dynamic. Thus, the main focus of this research is to develop a robust and efficient smoke detection system using video processing, smoke pattern recognition, and machine learning techniques.

Many methods exist for the detection of smoke in the computer image processing sector. However, most of them experimented by combining several methods to improve performance and reliability. Most methods are similar to conventional smoke detection approaches. They include motion detection, a region of particle analysis, smoke volume energy [3–5]. Growth rate analysis from the dynamic volume of an increased number of smoke pixel has a great impact on detecting smoke. The differences in the efficacy of those algorithms are observed in different stages. All of the methods and techniques mainly consist of three steps: the preprocessing step; the smoke feature extraction step; and the smoke identification step.

In the preprocessing step, determining the region of interest (ROI) serves as the fundamental task for identifying and analyzing the smoke pixels and the eligible regions from the video frames taken for the experiment. Many researchers have used the color segmentation method (CS) [6–11] to segment the ROI from the input videos taken by static cameras. This is generally done by transforming the RGB color space into HSV color space [7], YUV color space [9], YCbCr color space [11], or HIS color space [4] color spaces. The intensity value of an image pixel, saturation, and hue in terms of variations are visualized and punctuated by the HSV color space, which helps to perform the segmentation more easily. Optical flow [8] and frame differencing [5] are primarily used to analyze the vulnerable area to fire. Despite this, the frame differencing does not give optimal performance for background subtraction. As a result, in this paper, a hybrid smoke segmentation combination of the GMM [12] moving foreground detection and HSV color segmentation are considered for removing the complex background and unwanted non-smoke moving objects in the preprocessing steps.

In recent decades, several researchers have conducted studies in this area. For example, Y. Cappellini [6] built an intellectual system for programmed fire detection in the forests. G. Healey proposed a system to detect a real-time fire [13]. H. Yamagishi worked on an algorithm where a color camera is used for fire flame detection [7]. However, fire detection was based on flame detection, which proved inaccurate and over-sensitive to color features. K. Dimitropoulos et al. proposed a spatio-temporal flame and texture modeling for fire detection [14]. In this instance, smoke detection is both a prior demand and more challenging than detecting fire. Many researchers are working to detect smoke by efficiently extracting features from the smoke image with dynamic movement characteristics. Y. Chunyu [15] has worked on such a video-based fire and smoke detection system using motion and color features. I. Kolesov [9] worked with ideal mass transport based on visual flow and neural systems for fire and smoke detection from video. Wavelet in smoke, image processing for automatic smoke detection, and adoptive background modeling for real-time tracking [10,12,16] have been developed for building a very effective and accurate fire alarm system. D. K. Appana et al. [17] used optical flow characteristics for fire alarm systems in which they used combined features from the Gabor filter-based

edge orientation and the smoke energy components of Spatial-temporal frequencies. The proposed model of that paper considered HSV color analysis-based smoke region segmentation and the frame difference-based smoke flow pattern using the Gabor filter. Table 1 presents the comparative study of state-of-the-art research.

Table 1. A comparative study of state-of-the-art research.

Ref.	Aim of the Research	Method	Pros	Cons
[8]	Fires detection based on video analysis by surveillance cameras	Foreground masking, Background subtraction, Optical flow analysis using color evaluation, shape variation, and movement evaluation.	A hybrid combination of color evaluation, shape variation, and movement evaluation for optical flow shows the effective results.	Background subtraction and foreground masking based on frame differencing is vulnerable to a dynamic changing environment.
[9]	Fire and smoke detection using video	Optimal Mass Transportation (OMT) for extracting optical flow descriptor of RGB video frame and Neural Networks for classifying smoke/fire	OMT is useful for detecting smoke or fire on a similar colored background.	There is no background subtraction process is introduced and smoke color moving object might misguide the detection process
[14]	Flame modeling for wildfire detection using a video signal	Background subtraction using non-parametric model, Spatio-temporal features such as color probability, flickering, spatial and Spatio-temporal energy, and dynamic texture analysis for wildfire detection	Codebook with the combination of various Spatio-temporal and dynamic texture analysis construct a strong feature vector to classify fire using SVM.	This model is for flame and fire detection which need to be enhanced for smoke detection. Also, feature extraction using several flame movement descriptors demands high computation power.
[17]	optical flow characteristics for fire alarm systems	Combined features from the Gabor filter-based edge orientation and the smoke energy components of Spatial-temporal frequencies. SVM is used for smoke classification.	HSV color segmentation is effective for detecting a smoke object. Gabor filter-based edge orientation of frame differencing and Spatial-temporal energy of frame shows good smoke classification result.	Background segmentation on a static frame is suboptimal for a dynamically changing environment. Smoke descriptor based on temporal frame differencing might show some false alarm.
[18]	Motion modeling and dynamic texture recognition for smoke detection	HSV color segmentation for candidate smoke regions detection, Spatio-temporal energy analysis, and histograms of oriented gradients and optical flows (HOGHOFs)	spatio-temporal energy analysis, (HOGHOFs) show effectiveness for moving smoke detection	HOGHOF descriptors for smoke motion modeling are sub-optimal for a smoke-like moving object

Nevertheless, those papers considered single-frame-based smoke region segmentation and smoke descriptor calculation, and, because of this, the performance is suboptimal. Very few researches considered the dynamic background and smoke-colored moving object. Smoke flow based on consecutive frame differencing is not enough for detection complex smoke explosion. There is no consideration for the optical flow of non-smoke but rather the smoke-like object. Furthermore, the smoke particles' growth differs from non-smoke.

To address the challenges, the proposed model of this paper considers the advanced candidate smoke region segmentation and dynamic smoke growth feature extraction based on temporal frames of a continuous video sequence.

In this paper, the combination of GMM-based adaptive moving object detection and HSV color segmentation is used in the preprocessing stage for segmenting the moving smoke-only object from the complex background and non-smoke moving objects. After preprocessing, the smoke growth features are extracted using proposed frame-block segmentation-based smoke growth analysis. In particular, the Spatial-temporal energy features are extracted from selected temporal frames. Finally, the smoke features are classified by an RBF non-linear kernel-based Support Vector Machine (SVM) classifier.

This paper is organized in the following pattern. Section 2 discusses the details about the proposed model, including smoke growth analysis and feature extractions. Section 3 evaluates the performance of the proposed model. Section 4 concludes this paper.

2. Proposed Model

In this proposed model, the hybrid smoke segmentation combines moving foreground detection and smoke object segmentation. Block segmentation-based dynamic smoke growth analysis for smoke feature extraction, along with the spatial frame energy, is proposed here. Finally, an SVM algorithm is used to classify moving objects, which here is smoke. The proposed approach focused on the growth area of the smoke and characteristics features of smoke growth against time for classifying smoke and non-smoke objects. A flowchart of the proposed smoke detecting method appears in Figure 1 and is further discussed in the following subsections in greater detail.

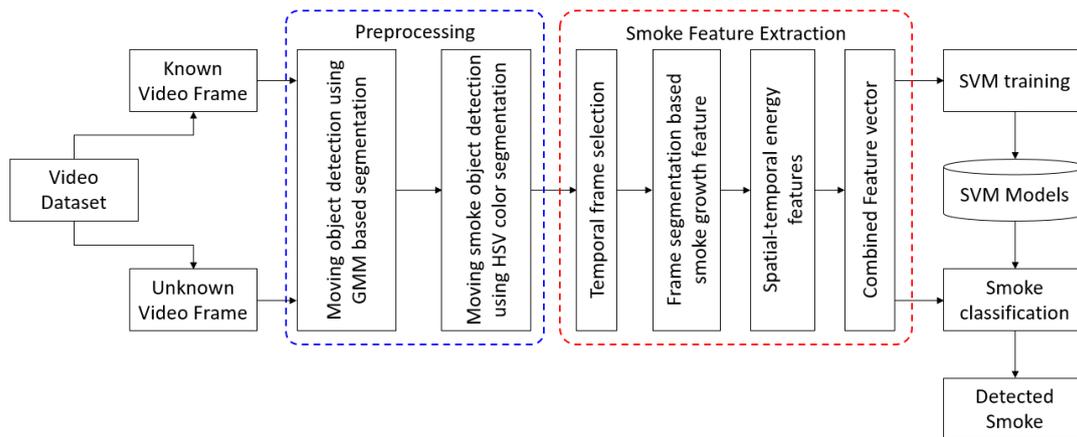


Figure 1. Proposed smoke detection model using video analysis.

2.1. Preprocessing and Hybrid Segmentation

This preprocessing consists of two steps. First, the GMM-based moving object detection is used for segmenting the moving objects from a complex background. Second, the segmented moving object is further preprocessed by HSV color segmentation for separating the smoke-like moving object. Next, the process is described in detail in the following subsections.

2.1.1. Moving Foreground Detection Using GMM Segmentation

In the case of smoke detection, the moving object on a background is the primary concern. A color-based segmentation subtracts the color-matched static object. Seeing this, moving foreground segmentation is quite essential. Background subtraction is processed to identify the foreground object for the concerned video frames. It is the most popular approach to detect moving image content from the video frame of a static camera.

A popular approach to background subtraction is frame differencing, where the frame difference is done between the current video frame and a reference frame. Still, there is a significant challenge in a real environment, where the background image is not fixed perfectly, i.e., the change of gradual and sudden illumination, change of background geometry, motion change, and others. Moreover, the frame differencing method is extremely sensitive to the threshold value and the frame rate based on the speed of foreground objects.

An efficient moving object subtraction can handle the long-term background change as well as a change of lighting intensity. To overcome the challenges, the Gaussian Mixture Model (GMM) provides the optimal solution. Seeing this, the GMM method is applied in this research for separating the moving foreground object of a video frame. The basic steps of this algorithm are preprocessing, background modeling, foreground identification, and data validation. Background modeling is the key concern of this algorithm. The GMM foreground detection methods are two types, i.e., non-adaptive and adaptive. The adaptive model is the most popular because, unlike the non-adaptive model, it maintains the background model over time.

In the initial step of the GMM foreground subtraction model, T number of frames are first considered to contract the Gaussian in the pixel-wise level and K number of Gaussian is used to represent a pixel X_t at time T . The probability of a pixel X can be formulated as Equation (1).

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \eta(X_t, \mu_{i,t}, \sigma_{i,t}^2) \quad (1)$$

where ω_i represents the weight of i -th Gaussian, $\mu_{i,t}$ represents the mean of i -th Gaussian of t -th frame, and σ_{i,t_2} covariance. The η probability density function of Gaussian is formulated in Equation (2).

$$\eta(X_t, \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{-\frac{(X_t - \mu)^2}{2\sigma^2}} \quad (2)$$

In these consecutive steps, the K number of Gaussians distributions are classified according to the value of ω/σ . Finally, the first B Gaussians are selected as the background model, where B is formulated as Equation (3).

$$B = \arg \min_b \left(\sum_{k=1}^b \omega_k > T \right) \quad (3)$$

where T is the minimum limit of fraction and the background models are selected when ω_1 to ω_b is exceeded to the value of T . In this iterative process, the weight of i -th Gaussian of K distribution of time t is updated according to the following equation:

$$\omega_{k,t} = (1 - \alpha)\omega_{k,t-1} + \alpha(M_{k,t}) \quad (4)$$

Figure 2 presents the overall process of GMM foreground moving object subtraction. The GMM method effectively separates the moving object; however, non-smoke objects might present inside the frame. With this, smoke or smoke-like objects need to be separated using color segmentation.

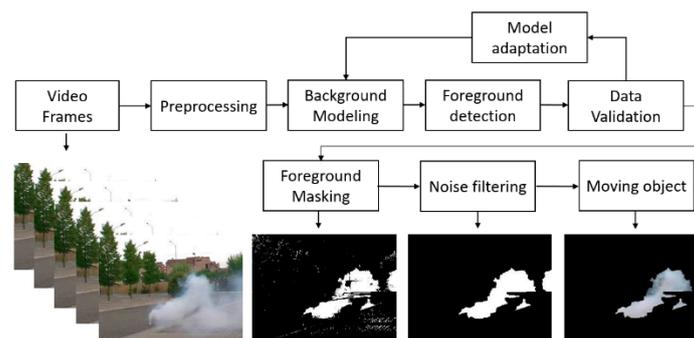


Figure 2. Gaussian Mixture Model (GMM) method for foreground moving object detection.

2.1.2. Smoke-Like Moving Object Separation Using HSV Color Segmentation

The moving object detection process efficiently segments the foreground moving object. There are other moving objects, however, which can misguide the process. Thus, removing the non-smoke subject might increase the performance of the smoke detection model. Hence, filtering only the object that is related to smoke becomes vitally important.

With a proper supply of resources and elements in a fire incident, fuels come in contact with oxygen in the air and create a combusting matter consisting of these two elements, leaving burnt residues that generate flame and smoke. The features of smoke can differ depending on the temperature at the time of burning, chemical elements, and surroundings. Oxygen and other elements are needed to ignite a fire. When the temperature is low, then the color of smoke can be bluish-white to white. The color changes to grayish-black to black when the temperature is rising and this color remains unchanged until combustion occurs. The foreground image frames, generally represented by the RGB color method, help to distinguish the different colors of the frames. However, regardless of these guidelines, images can still have problems in terms of nonlinear visual perception and illumination dependency [2]. To identify smoke, and smoke-like moving objects, the color segmentation is performed by finding the pixels that match the color from a specific frame of smoke. In this process, HSV color model analysis is used to transform the RGB color space, creating a threshold of hue (H), saturation (S),

and the components of value (V), which are formulated in Equation (5). The threshold limits of the saturation are s_{low} and s_{high} followed by 0 and 0.28. and the limits of the value (V) are v_{low} and v_{high} followed by 0.38 and 0.985. As no threshold to the hue components was applied, the value of h_{low} and h_{high} are 0 and 1, respectively.

$$F_{candidate}(i, j) = \begin{cases} 1, & f(\text{conditions} == \text{true}) \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$$\text{where } \text{conditions} = \begin{cases} h_{low} < H(i, j) < h_{high}, \\ S_{low} < S(i, j) < S_{high}, \\ v_{low} < V(i, j) < v_{high} \end{cases}$$

If the condition of Equation (5) is satisfied, the color of smoke is identified, where $F_{candidate}(i, j)$ is at the spatial location (i, j) . $H(i, j)$, $S(i, j)$, and $V(i, j)$ are followed by the hue, and saturation, as well as the value components of a pixel in the spatial space. The saturation limits and values are attained from the experimental statistical data, which are made with training videos, and could satisfy to discover smoke areas for the videos that are used for the experiment. The non-smoke regions are removed by a pre-processing method in which the extra unnecessary objects were removed, and the frames are being made smoother with the help of morphological closing method. Then, those areas were filled with the 2D four-connected neighborhood. Figure 3 presents the process of HSV color segmentation for removing the non-smoke object.

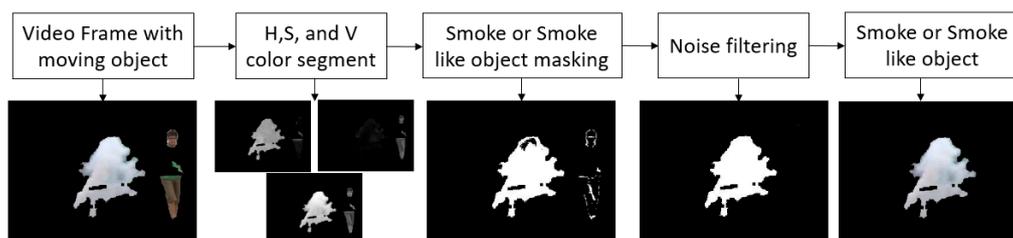


Figure 3. Removing non-smoke object using HSV color segmentation.

2.2. Smoke Feature Extraction

Smoke has a dynamic property to expand in a different direction as a diffusion process. Generally, smoke expands dynamically and changes its shape due to the effect of wind. For extracting the properties of the dynamic changes of smoke, the following subsections describe the temporal frame selection, a novel block segmentation-based smoke growth analysis, smoke growth features extraction, and spatial-temporal energy featuring extraction.

2.2.1. Temporal Frame Selection

Modern video capturing technology produces a large number of frames with a high resolution. Nonetheless, it is computationally expensive to consider all the frames in a video clip for processing. What is more, the expansion of smoke might not be significant in subsequent frames, which proves quite difficult to consider for the identification model. To overcome these issues, this research considers the selected temporal frames. To select the temporal frames, we considered the captured video is F frame per second (f/s), the frame selection interval is F/n , and considered video time T is N/n , where N is a total number of considered frames. In this study, two selected frames per second ($n = 2$) and 4 selected frames ($N = 4$) from 2 s ($T = 2$) are considered for analyzing the smoke growth. Figure 4 presents the temporal frame selection process. According to the proposed model, we are considering 4 consecutive frames within 2 s. The smoke features are extracted from considered frames and classify smoke based on the previously trained model. For the experiment, we developed the simulation program using Matlab, which is not so much faster. We hope that the whole feature extraction and

classification process will be finished within the next 2 s before coming to the next batch of frames using a faster coding technique.

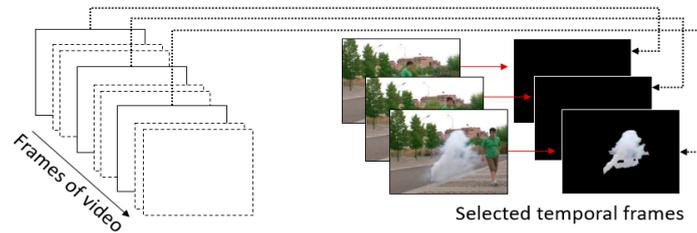


Figure 4. Extraction and selection of temporal frames.

2.2.2. Frame Blocks Segmentation

As a dynamic behavior, the smoke appears and expands dynamically both upward and in other directions. The dynamic growth of smoke can be determined by the analysis of the frame content of a video using a static camera. As indicated, several studies have been undertaken in this area. Normally, the growth of smoke is determined by calculating the frame differencing [16]. Several limitations exist, however, especially when subtracting frame content from another previous frame, which sometimes generates an unwanted smoke-like spot; such an occurrence might misguide the smoke identification process. In addition, subtracted frame output could fail to identify the behavior and direction of growth, which confuses the model for the smoke-like color, but not non-smoke objects. To quell this problem, this paper proposes a new block segmentation-based smoke growth analysis process in which the frames are segmented into blocks and a measure presents a smoke component in each block. The improvement of block content between consecutive blocks helps to determine the growth behavior and direction of smoke or smoke-like objects. This process consists of two parts. The following subsections describe the process of frame segmentation-based smoke growth analysis.

In the Frame-Block Segmentation (FBS) process, the video frames are segmented into multiple blocks and the smoke density of blocks are calculated. This process helps to analyze the presence of the smoke and smoke movement behavior analysis. Equation (6) presents the formula of FBS. In this paper, selected frames are segmented into 16 density blocks, which is later used for further analysis of smoke growth.

$$\begin{aligned}
 SB_{i,j} &= [(x_{start} : x_{end}), (y_{start} : y_{end})] \\
 x_{start} &= (\text{mod}(i, 4) - 1) \times (w/4) \\
 x_{end} &= \text{mod}(i, 4) \times ((w/4) - 1) \\
 y_{start} &= (\text{mod}(j, 4) - 1) \times (h/4) \\
 y_{end} &= \text{mod}(j, 4) \times ((h/4) - 1)
 \end{aligned} \tag{6}$$

where SB is the segmented block, w is the width, h is height, x_{start} and x_{end} are start and end position of x-axis of each block, y_{start} and y_{end} are start and end position of y-axis of each block, $i = 1..4$, and $j = 1..4$.

In the segmented blocks, the smoke or smoke-like objects are represented using an image component with the non-zero-pixel value. Consequently, the smoke density ratio (D_i) of non-zero elements is calculated of each (i -th) segmented block. The density of the non-zero element in a frame is calculated as (total number of non-zero pixels (n_i)/ total number of pixels of i -th block (N_i)) and multiplied by 100. Figure 5 presents the FBS process, SBs , and smoke density ratio (D_i) of the blocks.

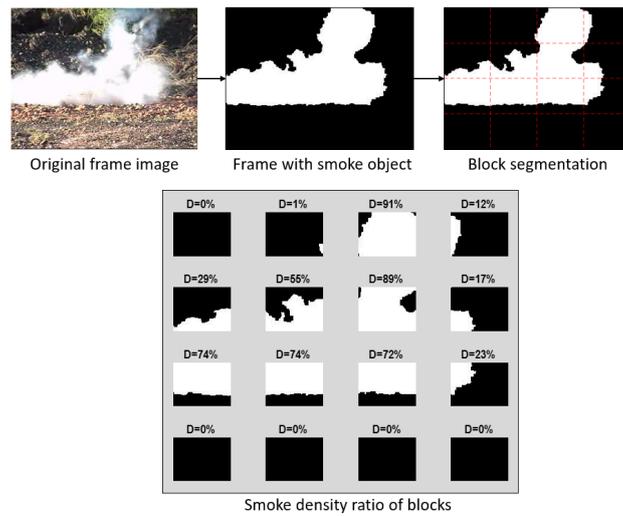


Figure 5. Segmented blocks of a frame and smoke density ratio (D_i).

2.2.3. Smoke Growth Segmented Frame Block

Commonly, smoke expands upward and in other directions with the influence of wind. Alternatively, the other smoke-like moving objects, i.e., car light movement, move in a horizontal direction, whereas a smoke-like static object, i.e., an electric bulb, has a static position.

To analyze the smoke growth, a segmented smoke block of consequence temporal selected frames is compared. In this process, the temporal selected frames (TF_j) are segmented using the *FBS* process. The smoke density ratio (D_i) is calculated for each block of each frame. Figure 6 presents the smoke density ratio in each block of each temporal frame. Finally, the smoke growth rates (*SGR*) of smoke growth frames (*SGF*) are calculated by differencing the D_i of corresponding *SB* of consequence in 2 frames. At the end of this growth analysis, *SGFs* are generated according to equation (7), where $TF_j D_i$ is smoke density ratio of i -th segmented block of j -th temporal frame. Figure 7 illustrates the process of calculation of smoke growth rate, where TF_1, TF_2, TF_3 are selected temporal frames and SGF_1 and SGF_2 are smoke growth frames.

$$\begin{aligned}
 & \text{if}(TF_{j+1}D_i - TF_jD_i > 0) \\
 & \quad SGF_jSGR_i = TF_{j+1}D_i - TF_jD_i \\
 & \text{else} \\
 & \quad SGF_jSGR_i = 0
 \end{aligned}
 \tag{7}$$

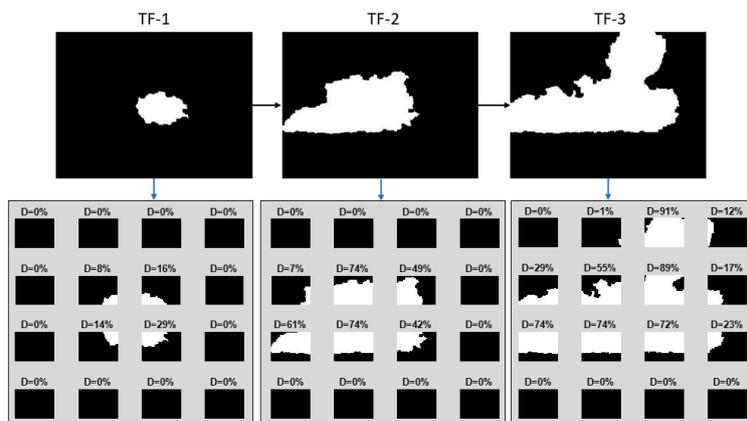


Figure 6. Smoke density ratio each block of each temporal frame.

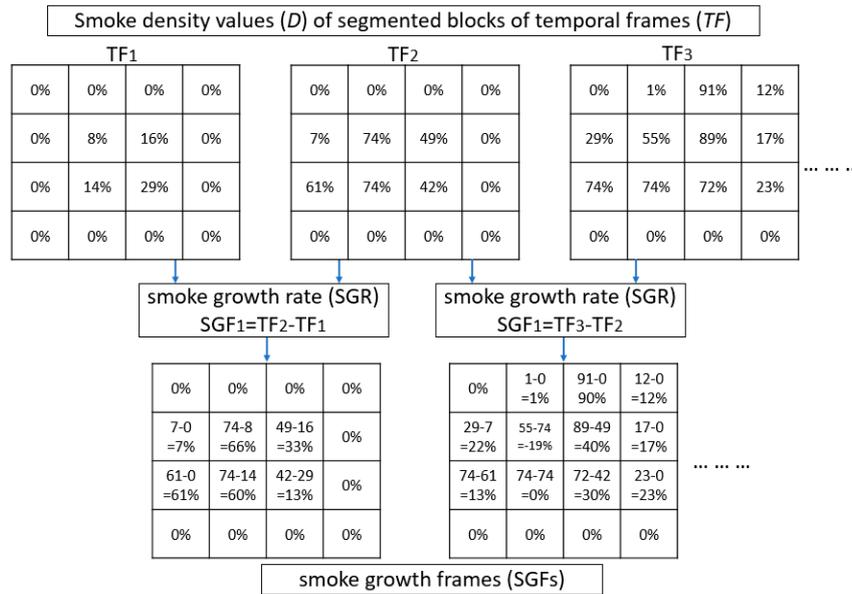


Figure 7. Process of smoke growth frame extraction, where SGF_1 and SGF_2 are smoke growth between TF_1, TF_2 and TF_2, TF_3 , respectively.

2.2.4. Smoke Growth Features

During the temporal consequence of smoke growth, the ratio of smoke growth of different segmented blocks varies. Moreover, some stationary and static movement smoke-like objects have distinguishable growth patterns compared with the smoke object. The localization of smoke growth within the segmentation block and its patterns is potential information for detecting smoke. The smoke growth rate (SGR) of the different segmented blocks is therefore considered for extracting features. In this feature extraction process, the SGR values of blocks of SGFs are averaged and construct the average smoke growth frame (ASGF). Finally, the block values of ASGF are organized into a feature vector. Figure 8 represents the process of the smoke growth feature extraction.

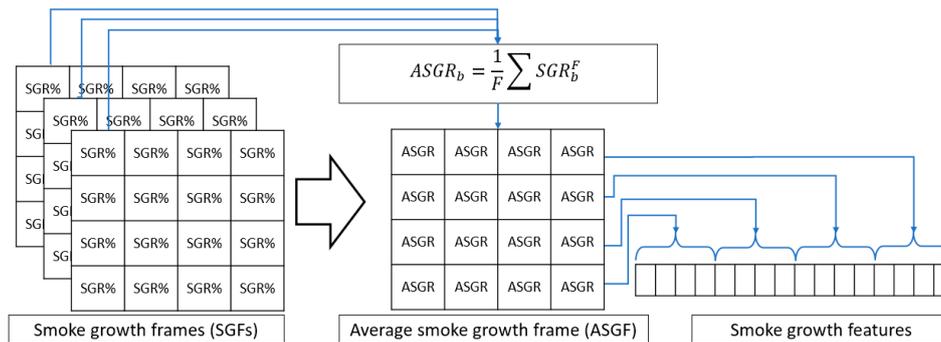


Figure 8. Smoke growth feature extraction from SGFs.

2.2.5. Spatial-Temporal Energy Features

Initially, the smoke object has mostly semi-transparent or very few image regions. In time, the transparency of smoke decreases while the smoke image region increases. Based on the level of opaqueness, the frame spatial energy is decreasing while the smoke volume is increasing. This parameter is also very effective to differentiate the smoke versus non-smoke moving objects. From this point of view, the spatial-temporal energies of video frames are calculated using 2D wavelet transform analysis.

In 2D wavelet transform analysis, the one-dimensional wavelet decomposition is initially applied to the columns. In this process, the low-pass and high-pass are useful to generate sub-images containing

the high frequency and low-frequency components, respectively. In the following step, the two sub-band images are again decomposed using the one-dimensional wavelet decomposition, which makes the four sub-band images, i.e., approximation, horizontal details, vertical details, and diagonal details. Figure 9 presents the wavelet decomposition of a 2D image and its output. For calculating the frame energy, the energy of each pixel is calculated by summing the energy of pixel of horizontal, vertical, and diagonal components. Ultimately, the frame average energy is calculated by averaging the energies of all pixels. Equation (8) presents the formula for calculating the energy of each pixel, and Equation (9) presents the formula of a selected frame. Figure 10 presents the spatial energies of frames with non-smoke and smoke moving objects.

$$E(i, j) = \omega_{\varnothing\varphi}(i, j)^2 + \omega_{\varnothing\varnothing}(i, j)^2 + \omega_{\varphi\varnothing}(i, j)^2 \tag{8}$$

$$E_n = \frac{1}{m * n} \sum_{i,j} E(i, j) \tag{9}$$

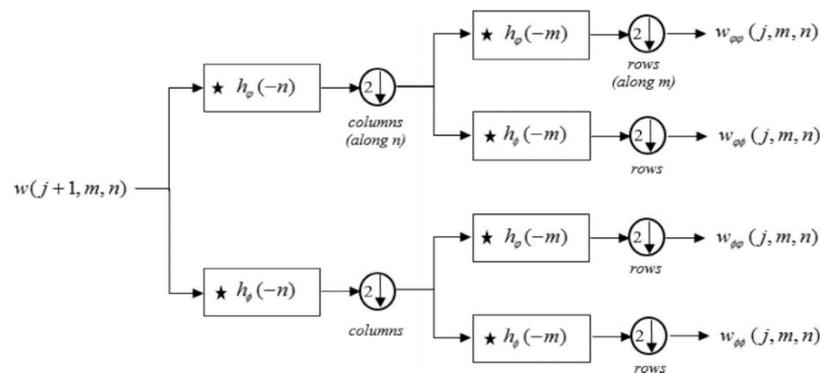


Figure 9. Wavelet decomposition of a 2D frame image.

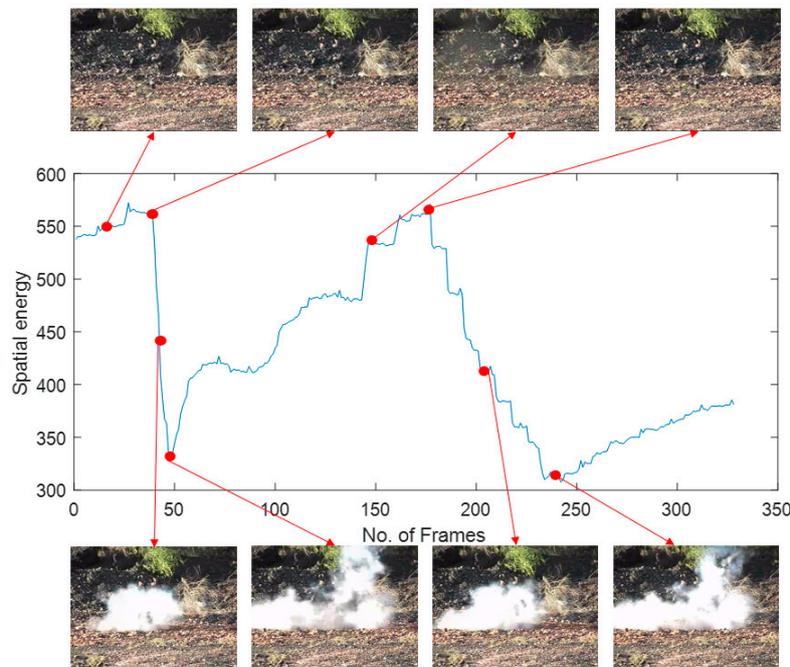


Figure 10. Spatial energies of frames with non-smoke and smoke moving object.

2.2.6. Final Feature Vector

Finally, the total feature vector is constructed by combining the smoke growth features and spatial-temporal energy features. As discussed in the previous subsection, the total number of SGFs of ASGF is 16, which represents the smoke growth features. Besides this, five spatial-temporal energy features are calculated from four temporal selected frames. Because of this, the total feature vector size is 20. Finally, the total feature vector of considered video clips is used for training for and verification of smoke detection.

2.3. Smoke Identification Using SVM

In general, smoke detection is a binary classification problem. Many state-of-the-art machine learning and classification models exist. Based on the problem set, different classification model shows satisfactory performance. The k-NN is a simple non-parametric classification algorithm. The k-NN classification algorithm does not use the stored pre-trained model at classification. During the classification process, it measures the distance value of neighbors of different classes and classifies a test sample based on the distance voting of the training sample. The k-NN is very simple; however, k-NN might not be good for complex and non-linear feature distribution. Moreover, it requires distance calculation at every time. Additionally, fixing the value of the number of nearest neighbor k is quite challenging [19]. On the other hand, CNN is an exceptionally effective machine learning and classification model, especially for image classification problems. The CNN executes a high number of convolutions in different layers and extracts high profiled features based on the context of images. Those features are very much effective for image classification. However, it demands a high computation process [20]. Alternatively, the SVM is commonly used in machine learning and classification models. In the training process, this algorithm finds the optimal hyper-line (linear or non-linear) and select the minimum number of support vectors to construct the trained model. During the classification process, the stored trained model is used for classifying the test sample. The SVM is a binary classifier and it is much suitable for the two-class classification process. If we can provide good extracted features, then the SVM shows very good results without high computation. The SVM has a linear and a different non-linear kernel, which are suitable for the classification problem accordingly. However, the RBF non-linear kernel is very much effective for most of the non-linear classification problems. Resultantly, the SVM is used in this proposed model for detection smoke.

The SVM is a non-probabilistic classifier, which can separate the provided data into two classes utilizing the optimal hyper-plane and maximizing the margin in high-dimensional feature space [17,21]. Usually, the linear SVM linearly separates the sample into classes, but several non-linear kernels of SVM, i.e., polynomial, Gaussian radial basis function, hyperbolic tangent, and others, are highly effective for classifying the complex non-linear problem. In this paper, the radial basis function (RBF) non-linear kernel is used in this proposed model for detection smoke. The RBF non-linear kernel is formulated as Equation (10).

$$k(sv_i, sv_j) = \exp\left(-\frac{\|sv_i - sv_j\|^2}{2\sigma^2}\right) \quad (10)$$

where $k(sv_i, sv_j)$ is the kernel function, and sv_i and sv_j are the input data, and parameter σ is a set by the user. The σ used here to determine the width of the kernel function k . Here, note that, if small σ values are used, then overtraining may occur. Again, if σ values are large, then the basis function puts an oval around the points without describing their shapes or patterns. Hence, seeing this, it is clear that σ values impact the classification accuracy. Therefore, in this study, optimal σ values were used to recognize smoke effectively.

3. Experimental Result and Evaluation

3.1. Experimental Setup and Video Dataset

In order to evaluate the proposed model, a standard experimental video clip dataset has been prepared. For the social and environmental restriction, we could not make any smoke or fire event inside or outside the university area. Thus, the video dataset contains eight videos, which are collected from the Bilkent and Visor online benchmarked video repositories. Table 2 includes a summary of the experimental video dataset. The frame resolution of the videos is 320×240 . Figure 11 shows the screenshots of the video collected from the Bilkent and Visor online benchmarked video repositories. Each video of the experimental dataset is split into 2-second video clips and stored into two classes, i.e., smoke clips and non-smoke clips. For example, if a video is 25 f/s then $25 \times 2 = 50$ frames consist of each clip. For this experiment, the Matlab 2018a version is used to build a simulation program, video analysis, feature extraction, and classification model.

Table 2. Summary of experimental videos.

Video #	Video Name	f/s	Time	No. of Frames
V_Bil_01	Bilkent/sBehindtheFance	10.00	1 min 3 s	630
V_Bil_02	Bilkent/sEmptyR1	16.67	28 s	466
V_Bil_03	Bilkent/sParkingLot	25.00	1 min 9 s	1725
V_Bil_04	Bilkent/sWasteBasket	10.00	1 min 30 s	900
V_Vis_01	Visor/movie13	25.00	1 min 20 s	2000
V_Vis_02	Visor/movie14	25.00	1 min 26 s	2150
V_Vis_03	Visor/burnout	25.00	1 min 28 s	2200
V_oth_01	other/IndoorVideo	14.99	1 min 20 s	1199



Figure 11. Screenshots of the videos considered in the experiment.

3.2. Experimental Process

In this proposed model, the video frames are initially preprocessed and remove the background for detecting the moving foreground object. The GMM-based moving object detector detects the moving foreground efficiently, which extracts all moving objects in the frame. Because moving smoke is the main concern of this detecting process, the filtered frames with moving objects are further processed for recognizing only smoke objects using HSV color segmentation. The preprocessing step segments the smoke or smoke-like object of the frame for further analysis. Figure 12 shows some preprocessed frames with smoke and non-smoke objects.

In the following step, temporal frame selection has been done to trade-off the image processing computational overhead and enough smoke growth for detection. To select the temporal frames, two frames per second and four selected frames are considered for analyzing the smoke growth.

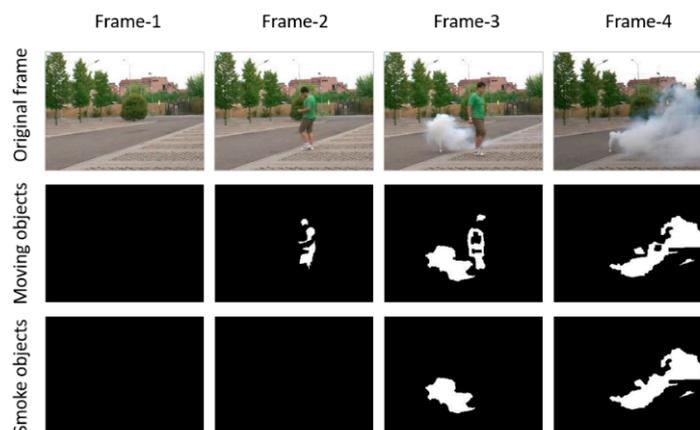


Figure 12. Foreground moving object and smoke object detection.

The next and most important step of the proposed algorithm is the frame segmentation-based smoke growth analysis. The temporal selected frames are segmented into 16 blocks and the smoke density ratios (Di) are calculated of each block of each frame. The smoke particles gradually fill the block. With a certain amount of smoke particles, almost all the blocks of the segmented are covered. Depending on the presence of smoke or smoke-like objects at the segmented block, the Di value is increased. In the experiment's smoke frames, it is noticed that the di values of upward segmented blocks are increased in consecutive temporal frames. Alternatively, moving smoke-like objects, such as car lights or a moving person, give similar changes of smoke, but the expansion of the Di value only moved in a horizontal direction. Moreover, the Di value of video frames for light bulbs remain at the same-segmented block. Figure 13 presents the smoke moving objects and non-smoke moving objects. The stationary smoke-like object was already filtered out by moving foreground object detection.

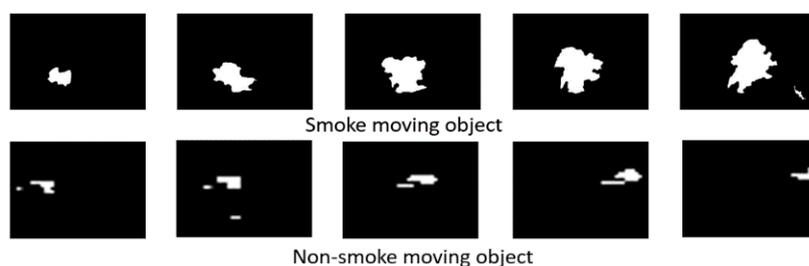


Figure 13. Spatial energies of temporal frames with non-smoke and smoke moving object.

After getting the Di , the smoke growth rate (SGR) is calculated by differencing the Di of the corresponding segmented block of consequence (two frames) and constructs the smoke growth frame (SGF). At the end of this growth analysis, three SGFs are generated, where $SGF_1 = TF_2 - TF_1$, $SGF_2 = TF_3 - TF_2$, and $SGF_3 = TF_4 - TF_3$. At the end of the smoke growth calculation, the 16 ASGFs are calculated for each block from all selected temporal frames to extract features. The ASGF values of the density block are serialized in vector to generate smoke growth features vector. Figure 14 presents non-smoke and smoke growth feature vectors.

To extract the spatial-temporal energy of a frame, the average energy of horizontal, vertical, and diagonal components of 2D wavelet transformation is calculated. As discussed in the previous section, the energy of the frame is varied based on the level of the opaqueness of the smoke volume. Due to this, the spatial energies of selected temporal frames are considered as spatial-temporal energy features. Figure 15 presents the Spatial energies of frames with non-smoke and smoke clips for a video with 25 frames/second, where red presents energy of frames with some and the color blue presents otherwise.

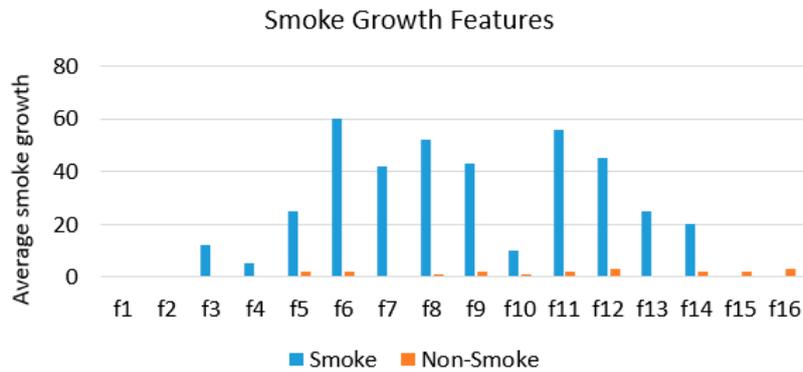


Figure 14. Spatial energies of temporal frames with non-smoke and smoke moving object.

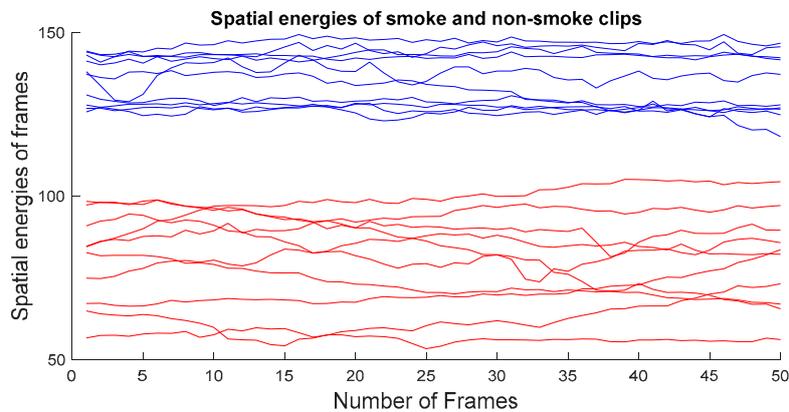


Figure 15. Spatial energies of frames with non-smoke and smoke clips, where the color red presents the energy of frames with some and the color blue presents otherwise.

Finally, to generate the feature vector for classifying smoke, two types of features are combined: a) smoke growth features and b) spatial-temporal energy features. In total, 20 features combining with 16 ASGFs and 4 (four) energies are considered as a feature vector for the classification process. For evaluating the proposed model, the extracted phenomena of smoke growth are used for training and classification using SVM. To increase the reliability of the experimental results, 2-fold cross-validation is used. The 2-fold cross-validation splits the extracted feature vector of dataset randomly into three subsets, and each subset contains $(N_{samples}/2) \times N_{classes} \times N_{features}$ feature vectors. In this experiment, $N_{classes}$ is 2 (two) and $N_{features}$ is 20. In the evaluation process using SVM, the entire training and testing process is iterated in two instances. In each iteration, one subset of the feature vector is used from training and one is used for testing. In this study, the classification accuracy (CA) is calculated from the confusion matrix for assessing the performance in each evaluation. The calculation CA is formulated in Equation (11), where the $N_{tsample}$ is a total test sample and N_{TP} is the number of true positives (number of data points of class A that are correctly classified as class A) of a class. Table 3 presents the experimental parameters of this proposed model. Finally, the average of the CA of all classes over the iterations is considered as final evaluation criteria.

$$CA = \frac{\sum_{N_{classes}} N_{TP}}{N_{tsamples}} \times 100(\%) \quad (11)$$

Table 3. Experimental parameters.

Parameter Name	Notation	Value
HSV color segmentation		
Min threshold of hue (H)	h_{low}	0
Max threshold of hue (H)	h_{high}	1
Min threshold of saturation (S)	s_{low}	0
Max threshold of saturation (S)	s_{high}	0.28
Min threshold of value (V)	v_{low}	0.38
Max threshold of value (V)	V_{high}	0.985
Temporal frame selection		
Frame per second	f/s	Based on video
Selected frame per second	n	2
Total number of considered frame	N	4
Total time duration	T	2
Frame block segmentation		
Segmented density block	SB_{ij}	$16, i = 4, j = 4$
Special temporal energy		
Level of wavelet transformation		3
Feature vector		
Total number of features	$N_{feature}$	20
Total number of classes	N_{class}	2

3.3. Experimental Result and Evaluation

To validate the proposed model, the experimental model is contrasted with state-of-the-art smoke detection models. Table 4 presents the details of compared methods for validation.

Table 4. List of compared algorithm.

Algorithms	Ref #	Title	Method
Algorithm-1	[18]	Smoke detection using Spatio-temporal analysis, motion modeling and dynamic texture recognition	HSV color segmentation for candidate smoke regions detection, Spatio-temporal energy analysis, and histograms of oriented gradients and optical flows (HOGHOFs)
Algorithm-2	[17]	Smoke detection approach using optical flow characteristics for alarm systems	Combined features from the Gabor filter-based edge orientation and the smoke energy components of Spatial-temporal frequencies. SVM is used for smoke classification.

Figure 16 presents the average CA of different videos of the proposed model and state-of-the-art algorithms. In Algorithm-1 [18], the author presented an efficient spatial-temporal analysis-based smoke detection model. In this approach, the smoke region of the video frame is segmented using HSV color analysis only, which, on occasion, shows a false smoke region for the complex and dynamic movement background. In our proposed model, this challenge is addressed by adaptive moving object detection and the color segmentation-based smoke region detection process. In paper [18], spatial-temporal smoke analysis is applied, which has proven to be effective. Conversely, histograms of oriented gradients and optical flow (HOGHOF) descriptors for smoke motion modeling are sub-optimal for a smoke-like moving object. In V_Bil_02, there is a background light and V_Bil_03 has very thin smoke. Thus, the performance of Algorithm-1 shows very low classification accuracy, which degrades the overall performance. In Algorithm-2 [17], the author presented a smoke flow pattern for smoke detection. By contrast, only color segmentation is used for smoke segmentation in this paper. The smoke flow pattern is detected by a multi-angle orientation-based Gabor filter applied to the temporal differenced frame. Furthermore, the statistical parameter is extracted as features. Temporal differenced smoke edges extracted by the Gabor filter indicate good performance in many cases. In the case of extremely dynamic moving smoke, however, our proposed model outperforms this model. During the experiment, the performances of the state-of-art model for video V_Bil_02 (Bilkent/sEmptyR) is below average because of static smoke-color light. Even so, the proposed model shows better performance. In the video V_Bil_03 (Bilkent/sParkingLot) the smoke transparency is extremely high,

which affects the classification performance. Figure 17 provides the average classification accuracy for the proposed model as well as the state-of-the-art algorithms.

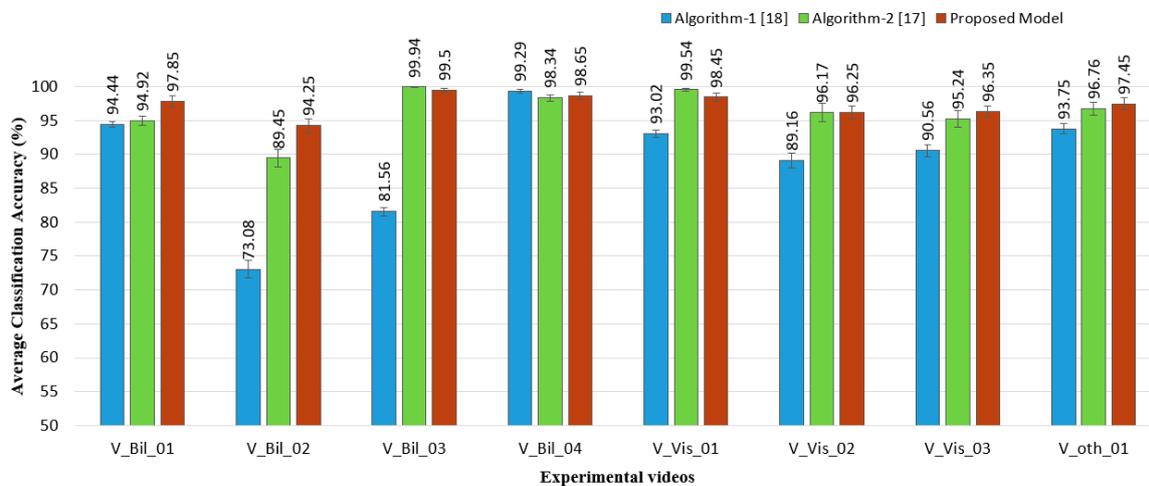


Figure 16. Average classification accuracy (CA) of different videos of the proposed model and state-of-the-art algorithms.

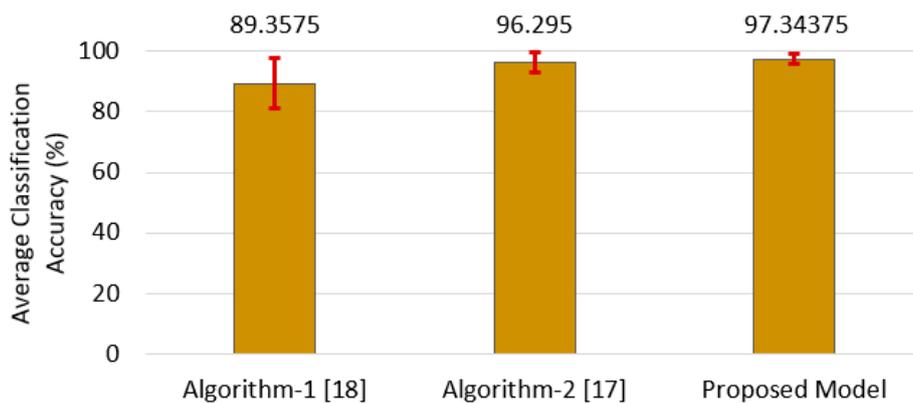


Figure 17. Average classification performances of smoke detection algorithms.

4. Conclusions

Nowadays, fire incidents are taking lives more than other accidents. Such incidents destroy properties, businesses, industries, and cause an imbalance to nature. Smoke is an early indicator of a fire accident. In this paper, an effective approach is proposed for the detection of smoke based on the smoke-growth analysis. The flow of the working process starts with the preprocessing. In preprocessing, the background is subtracted, and moving smoke objects are separated using GMM-based adaptive background subtraction and HSV color analysis-based segmentation. To reduce the burden of image processing, the temporal frame is selected. The presence of smoke in a target area is determined by the growth of smoke in the region based on novel smoke block segmentation-based smoke growth analysis. After analysis, the smoke growth features are extracted. Simultaneously, the spatial-temporal energy features are extracted, and a feature vector is generated for smoke classification. Thereafter, SVM-based decision making is applied to identify the smoke appearing in the video frames. The proposed model has experimented with standard videos from benchmarked datasets, which eventually showed improvement in classification. The proposed model outperforms state-of-the-art algorithms, yielding 97.34% average classification accuracy. Besides the benefits, this proposed model might not be suitable for a smoke accident at far from the camera. This is very much effective for close monitoring in the indoor and outdoor environment with other movable objects. Moreover, we did not evaluate our algorithm for a dense fog scenario. As future work in the realm of outdoor smoke and fire monitoring,

we might consider some abnormal environments. Moreover, we may consider an advanced deep learning model for better performance.

Author Contributions: Conceptualization, M.R.I.; Data curation, M.R.I., M.A. and S.N.; Formal analysis, M.R.I.; Funding acquisition, M.A.; Investigation, M.R.I.; Methodology, M.R.I. and S.N.; Project administration, J.S.; Resources, M.A., and J.S.; Software, M.R.I. and S.N.; Supervision, J.S.; Validation, M.R.I. and M.A.; Visualization, M.R.I., M.A. and S.N.; Writing—original draft, M.R.I. and S.N.; Writing—review & editing, M.A., and J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was in part supported by the Kent State University Tenure-Track faculty startup fund.

Acknowledgments: First, we would like to thanks to Pattern Processing Lab, University of Aizu, Japan for giving us the opportunity to do this research. Also, we want to say thanks to University of Asia Pacific for co-operation. Finally, thanks to Kent State University for supporting us the research funding.

Conflicts of Interest: There is no conflict of interest.

References

1. Xiong, Z.; Caballero, R.; Wang, H.; Finn, A.M.; Lelic, M.A.; Peng, P.-Y. Video-based smoke detection: Possibilities, techniques, and challenges. In Proceedings of the IFPA, Fire Suppression and Detection Research and Applications—A Technical Working Conference (SUPDET), Orlando, FL, USA, 5–8 March 2007.
2. Zikria, Y.B.; Afzal, M.K.; Kim, S.W. Internet of Multimedia Things (IoMT): Opportunities, Challenges and Solutions. *Sensors* **2020**, *20*, 2334. [[CrossRef](#)] [[PubMed](#)]
3. Fujiwara, N.; Terada, K. Extraction of a smoke region using fractal coding. In Proceedings of the IEEE International Symposium on Communications and Information Technology (ISCIT), Sapporo, Japan, 26–29 October 2004; Volume 2, pp. 659–662.
4. Ojo, J.A.; Oladosu, J.A. Application of panoramic annular lens for motion analysis tasks: Surveillance and smoke detection. In Proceedings of the 15th International Conference on Pattern Recognition, Barcelona, Spain, 3–7 September 2000; Volume 4, pp. 714–717.
5. Catrakis, H.J.; Aguirre, R.C.; Ruiz-Plancarte, J.; Thayne, R.D. Shape complexity of whole-field three-dimensional space-time fluid interfaces in turbulence. *Phys. Fluids* **2002**, *14*, 3891–3898. [[CrossRef](#)]
6. Cappellini, Y.; Mattii, L.; Mecocci, A. An intelligent System for automatic fire detection in forests. In Proceedings of the IEEE 3th International Conference on Image Processing and its Applications, Warwick, UK, 18–20 July 2002; pp. 563–570.
7. Yamagishi, H.; Yamaguchi, I. Fire flame detection algorithm using a color camera. In Proceedings of the 1999 International Symposium on Micromechanics and Human Science, Nagoya, Japan, 23–26 November 2002; pp. 255–260.
8. Foggia, P.; Saggese, A.; Vento, M. Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion. *IEEE Trans. Circuits Syst. Video Technol.* **2015**, *25*, 1545–1556. [[CrossRef](#)]
9. Kolesov, I.; Karasev, P.; Tannenbaum, A.; Haber, E. Fire and smoke detection in video with optimal mass transport based optical flow and neural networks. In Proceedings of the IEEE International Conference Image Process, Hong Kong, 26–29 September 2010; pp. 761–764.
10. Toreyin, B.U.; Dedeoglu, Y.; Cetin, A.E. Wavelet based real-time smoke detection in video. In Proceedings of the 13th European Signal Processing Conference, Antalya, Turkey, 4–8 September 2005; pp. 1–4.
11. Muller, M.; Karasev, P.; Tannenbaum, A.; Kolesov, I. Optical flow estimation for flame detection in videos. *IEEE Trans. Image Process.* **2013**, *22*, 2786–2797. [[CrossRef](#)] [[PubMed](#)]
12. Stauffer, C.; Grimson, W.E.L. Adaptive background mixture models for real-time tracking. In Proceedings of the IEEE Conference Computer Vision and Pattern Recognition, Fort Collins, CO, USA, 23–25 June 1999; pp. 246–252.
13. Healey, G.; Slatcr, D.; Lin, T.; Drda, B.; Goedeke, A.D. A system for real-time fire detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, NY, USA, 15–17 June 1993; pp. 605–606, date of published 06 August 2002.
14. Dimitropoulos, K.; Barmoutis, P.; Grammalidis, N. Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection. *IEEE Trans. Circuits Syst. Video technol.* **2015**, *25*, 339–351. [[CrossRef](#)]

15. Chunyu, Y.; Jun, F.; Jinjun, W.; Yongming, Z. Video fire smoke detection using motion and color features. *Fire Technol.* **2010**, *46*, 651–666. [[CrossRef](#)]
16. Vicente, J.; Guillemant, P. An image processing technique for automatically detecting forest fire. *Int. J. Therm. Sci.* **2002**, *41*, 1113–1120. [[CrossRef](#)]
17. Appana, D.K.; Islam, M.R.; Kim, J.M. Smoke detection approach using optical flow characteristics for alarm systems. *Inf. Sci.* **2017**, *2017*, 418–419.
18. Barmpoutis, P.; Dimitropoulos, K.; Grammalidis, N. Smoke detection using spatio-temporal analysis, motion modeling and dynamic texture recognition. In Proceedings of the 2014 22nd European Signal Processing Conference (EUSIPCO), Lisbon, Portugal, 1–5 September 2014; pp. 1078–1082.
19. Uddin, M.S.; Islam, M.R.; Khan, S.A.; Kim, J.; Kim, J.-M.; Sohn, S.-M.; Choi, B. Distance and density similarity based enhanced k-NN classifier for improving fault diagnosis performance of bearings. *Shock Vib.* **2016**, *2016*, 1–11. [[CrossRef](#)]
20. Rahim, M.A.; Islam, M.R.; Shin, J. Non-touch sign word recognition based on dynamic hand gesture using hybrid segmentation and CNN feature fusion. *Appl. Sci.* **2019**, *9*, 3790. [[CrossRef](#)]
21. Islam, M.R.; Uddin, J.; Kim, J.-M. Acoustic emission sensor network based fault diagnosis of induction motors using a gabor filter and multiclass Support Vector Machines. *Ad. Hoc. Sens. Wirel. Net. (AHSWN)* **2016**, *34*, 273–287.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).