Fire detection and 3D surface reconstruction based on stereoscopic pictures and probabilistic fuzzy logic

ByoungChul Ko*, Ji-Hun Jung, Jae-Yeal Nam

Shindang-dong Dalseo-gu, Department of Computer Engineering, Keimyung University, Daegu 704–701, Republic of Korea

ARTICLE INFO

Article history:
Received 29 January 2014
Received in revised form
17 April 2014
Accepted 31 May 2014

Keywords:
Fire detection
Gaussian membership functions
Fuzzy logic
Feature point
3D reconstruction

ABSTRACT

In this paper, a new fire detection method is proposed, which is based on using a stereo camera to calculate the distance between the camera and the fire region and to reconstruct the 3D surface of the fire front. For the purpose of fire detection, candidate fire regions are identified using generic color models and a simple background difference model. Gaussian membership functions (GMFs) for the shape, size, and motion variation of the fire are then generated, because fire regions in successive frames change constantly. These three GMFs are then applied to fuzzy logic for real-time fire verification. After segmentation of the fire regions from left and right images, feature points are extracted using a matching algorithm and their disparities are computed for distance estimation and 3D surface reconstruction. Our proposed algorithm was successfully applied to a fire video dataset and its detection performance was shown to be better than that of other methods. In addition, the distance estimation method yielded reasonable results when the fire was a short distance from the camera and the reconstruction of the 3D surface showed a shape that was almost the same as that of the real fire.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

In recent years, fire detection based on systems that use computer vision techniques has been widely researched because of the relative advantages of computer vision over conventional sensors, such as infrared, optical, or ion sensors, which depend on certain characteristics of fire. Unlike conventional sensors, which sense smoke, heat, or radiation, vision-based approaches using CCD cameras have the following advantages [1–3].

• The equipment cost is lower, as these approaches utilize surveillance cameras that are already installed in many public places.
• Vision sensors can monitor a larger area because the camera functions as a volume sensor, rather than as a traditional point sensor.
• Cameras can easily be used to gather additional information, such as the location, size, and degree of the fire.
• The system manager does not need to visit the location to check the surveillance monitor.

Although a vision-based approach can gather data on the exact location of a fire, exact information about its status, such as the volume and the distance from the fire to the camera, cannot be estimated using only the camera.

The purpose of this study is to estimate not only the exact location but also the volume of a fire, together with the distance between the fire region and the camera, as input for an automatic fire suppression system (AFSS). The main functions of AFSS include detecting fire regions using a camera and suppressing the fire by automatically activating a water cannon. The water cannon, equipped with a camera, is installed at remote sites and transmits image sequences to a monitoring server over a wired or wireless network as shown in Fig. 1. If a fire region is detected, the warning system sends a command to the water cannon at the remote site to suppress the fire. The water pressure is controlled in relation to the size of the fire region, and the nozzle of the water cannon oscillates in order to suppress the fire, because the system cannot calculate the exact distance between the water cannon and the fire. Therefore, the estimation of distance and real volume is essential for designing a more efficient AFSS that will suppress fires rapidly.

In our study, we use a pre-calibrated stereo camera instead of a normal one for detecting the fire region, tracking the fire spread in three-dimensional space, and calculating the distance from the fire to the camera. In addition, because a stereo camera is used to influence the estimate depth (distance) information according to the baseline between two cameras and the focus length, we show that the distance accuracy when a stereo camera is used is reasonable, particularly over a short distance, regardless of whether the environment is indoors or outdoors.

*Corresponding author. Tel.: +82 53 580 5235; fax: +82 53 580 6275.
E-mail addresses: niceko@kmu.ac.kr (B. Ko), jihunJung@kmu.ac.kr (J.-H. Jung), jynam@kmu.ac.kr (J.-Y. Nam).

http://dx.doi.org/10.1016/j.firesaf.2014.05.015
0379-7112/© 2014 Elsevier Ltd. All rights reserved.

Please cite this article as: B. Ko, et al., Fire detection and 3D surface reconstruction based on stereoscopic pictures and probabilistic fuzzy logic, Fire Safety J (2014), http://dx.doi.org/10.1016/j.firesaf.2014.05.015
1. Related works

A number of papers on the subject of computer vision-based fire detection has been published because it is a relatively new subject in computer vision research; some promising results have already been reported. However, many problems remain completely unsolved, as in most such systems, because the behavior of uncontrolled flames varies and cameras may produce images of the same scene that have different colors as a result of their different internal settings and different illumination balancing algorithms [4]. In several types of studies [1–13], researchers have attempted to solve the drawbacks of vision sensors so that dependable fire detection results can be achieved.

Töreyin et al. [1] proposed a real-time fire detection method in which the motion, color clues, and flame flickering are analyzed. A temporal wavelet transform is used for detecting quasi-periodic behavior of flame boundaries and a spatial wavelet transform of moving fire-colored regions is used for detecting color variation in flame regions. The irregularity of the boundary of the fire-colored region is also used for verifying fire regions.

Ko et al. [2] proposed a support vector machine (SVM)-based fire detection method. In this method, candidate fire regions are detected using motion and fire color. Then, since fire regions generally have a higher luminance contrast than neighboring regions, a luminance map is made and used to remove non-fire pixels. Finally, a temporal fire model with wavelet coefficients is applied to a two-class SVM classifier for the final fire-pixel verification. In addition, Ko et al. [5] used fuzzy logic for fire-flame detection, whereby the probability membership functions for each logic set are modeled using intensity, wavelet energy, and motion features.

Ho [6] proposed a video-based flame and smoke detection method in which a motion history detection algorithm is implemented to register a possible flame and then analyze the spectral, spatial, and temporal characteristics of the flame in the image sequences. Statistical distribution, temporal probability density, and the continuously adaptive mean shift-tracking algorithm are employed to detect the flame in real time.

Wang [7] used flame color probability based on a Gaussian color model and employed motion probability for updating the background image. The color and motion probabilities are integrated to obtain flame candidates and the successive feature vectors are then applied to the randomness test in order to obtain the flame probability.

For a video-based fire analysis, Pastor et al. [8] proposed a thermal image processing method for computing the rate of spread (ROS) of forest fires. To estimate the ROS, the correspondence between points in the thermal image obtained and the real plane is calculated using a direct linear transformation, and the position of the flame front is determined by applying a threshold-value-searching criterion within the temperature matrix of the target surface.

Martinez-de Dios et al. [9] presented computer vision techniques for forest fire detection involving the measurement of forest fires properties such as the fire front, flame height, flame inclination angle, and fire base width required for the implementation of advanced forest fire-fighting strategies. The system computes a 3D perception model of the fire and can be used for visualizing the fire evolution on a remote computer system.

Verstocket et al. [10] fused low-cost video fire-detection results from multiple cameras using a multi-view localization framework. This framework merges the single-view detection results of the multiple cameras using a homographic projection on multiple horizontal and vertical planes slicing the scene. To provide a more reliable fire analysis, this framework creates a 3D grid along with spatial and temporal 3D filters by extending the concept of existing 2D filters.

Although computer-vision based fire detection and analysis methods yield good detection results, they cannot provide the additional information required for extinguishing a fire, such as the fire’s volume and spreading direction, and the distance between the camera and the fire region, when only one camera is used. Unlike previous methods that focus on fire-flame detection using one camera, the instrumentation system proposed by Rossi et al. [11,12] for the visualization and quantitative characterization of fire fronts in outdoor conditions is based on stereovision. This paper introduces the modeling of 3D fire fronts and the extraction of geometric characteristics, such as volume, surface area, heading direction, and length.

In a previous study [13], we proposed a simple fire-region detection system that implements a 3D modeling algorithm and uses a stereo camera in outdoor settings. However, in the present study, to estimate the exact fire volume and the camera’s distance from the fire region as input for the AFSS, the candidate fire regions are first detected using fire-colored pixels and moving regions. Next, the Gaussian membership functions (GMFs) for fuzzy logic are learned from the image sequences based on the temporal variation in the fire regions, and GMFs are then applied to a pre-defined fuzzy logic model, which is used for the final verification of the fire region. For 3D reconstruction and a distance calculation, the feature points and their disparities are estimated from stereoscopic images of the fire region.

In brief, the main contributions and overall procedures of our work are as follows:

1. We introduce the AFSS for fire suppression by automatically activating a water cannon based on its distance error when using a single camera. To solve this problem and for the design
of a more efficient AFSS, we propose a distance estimation mechanism between the camera and fire, and construct a real 3D volume of the fire.

2. In this research, we propose a new temporal shape variation by changing the radius signature [14] because such a signature provides only spatial shape information.

3. In [2] and [5], statistical features are estimated from 100 frames, which is a major reason for the use of real-time processing and the memory requirement. In this research, to consider the vague and irregular characteristics of a flame, we propose a new temporal shape and size variation technique from only three corresponding video frames. Motion orientation proposed in [5] is also estimated from the corresponding three frames. Then, Gaussian membership functions are generated using the statistical features of these three visual characteristics for fuzzy logic.

4. We designed eight new fuzzy rules by modifying the work in [5] based on three input variables, i.e., Gaussian membership functions of the shape variation, size variation, and motion orientation.

5. To reduce the computation time, we apply a matching algorithm only to the corresponding points in two stereo images and estimate the disparity using the corresponding points for a depth calculation.

6. We present a method for converting pixels into millimeters to estimate the real distance between the fire and camera.

7. We prove experimentally that the real distance between the camera and fire may be erroneous according to length of the baseline and the image resolution of the stereo camera. Therefore, an adequate stereo camera is needed to calculate the exact distance of the fire by considering the baseline and image resolution of the stereo camera.

The remainder of this paper is organized as follows. In Section 2, we describe the detection of candidate fire regions and generation of Gaussian membership functions (GMFs) for application in fuzzy logic. Fire verification using fuzzy logic and GMFs is introduced in Section 3. In Section 4, we introduce the distance estimation and 3D surface reconstruction of a fire region. In Section 5, we discuss the accuracy and applicability of the proposed fire detection and distance estimation methods on the basis of our experimental results. Finally, our conclusions and the scope of our future work are presented in Section 6.

2. Detection of candidate fire region and extraction of Gaussian membership functions

2.1. Candidate fire detection using color model and temporal difference

Because one of the primary purposes of this research is to reconstruct 3D fire fronts, the exact segmentation of a fire region is very important. To segment fire pixels, we implement a rule-based generic color model [15] that uses YCbCr color space to separate the luminance from the chrominance, which is more effective than RGB or rgb color space. The rules are composed of four parts:

\[ Y(x, y, t) > Cb(x, y) \]  \hspace{1cm} (1)

\[ Cr(x, y, t) > Cb(x, y) \]  \hspace{1cm} (2)

\[ Y(x, y) > Y_{mean}, Cb(x, y) < Cb_{mean}, Cr(x, y) > Cr_{mean} \]  \hspace{1cm} (3)

\[ |Cb(x, y) - Cr(x, y)| \geq r, \]  \hspace{1cm} (4)

where \( Y(x, y), Cb(x, y), \) and \( Cr(x, y) \) are the luminance, chrominance-blue, and chrominance-red values at spatial location \( (x, y) \), respectively. Similarly, \( Y_{mean}, Cb_{mean}, \) and \( Cr_{mean} \) are the mean values of the same three color components. Any pixel that satisfies the conditions given from Eqs. (1) (through 4) is labeled as a fire pixel. In Eq. (4), \( r \) is a constant value for reducing the number of
false positive fire pixels; we set it at 40 according to the results presented in [15].

After the fire pixels have been segmented, neighboring fire pixels are merged into fire regions using morphological closing and region merging operations.

Uncontrolled fire regions tend to move continuously because of the airflow caused by the wind or burning material. False fire regions are easily removed using frame subtraction. This process is essential for improving fire detection performance and reducing detection time. If the average difference between one region of the previous frame, \( f_{t-1}(x,y) \), and one region of the current frame, \( f_t(x,y) \), at position \((x, y)\) is greater than the threshold \( t \), region \( r \) is defined as a candidate fire region using the formula:

\[
\begin{cases}
    r: \text{candidate region} & \text{if } \frac{1}{N_r} \sum_{(x,y) \in r} |f_{t-1}(x,y) - f_t(x,y)| > t \\
    \text{Remove} & \text{otherwise}
\end{cases}
\]

(5)

where \( N_r \) is the number of pixels included in a region \( r \). Fig. 2 show the segmentation results of fire-colored candidate regions when the rule-based generic color model and frame subtraction are applied after the morphological operation.

2.2. Gaussian membership functions for shape variation of fire region

After the candidate fire regions have been detected, false moving flame-colored regions should be removed. In this paper, we use a fuzzy logic model based on fuzzy rules and Gaussian membership functions (GMFs) inspired by [5] to verify that the candidate regions are real flames.

First, we estimate the temporal shape variation of the fire region, because flames exhibit irregular fluctuation patterns in the time domain [2] as a result of the airflow. The variance in shape is calculated using a modified radius-based signature [14]. After the boundary of the segmented region is extracted, we extract \( N \) uniformly sampled boundary points and estimate the starting point as the most distant point from the centroid to two symmetric boundary points aligned with the central axis. The \( N \) radius distances are computed from uniformly sampled boundary points to the centroid, as shown in Fig. 3.

Because the purpose of a shape feature is to analyze the changes in the irregular pattern of the radius distance between consecutive frames, we sum the corresponding radius distances \( (\text{Radi}_i) \) of three frames, i.e., the front \((t+1)\), current \((t)\), and rear \((t-1)\), in clockwise order using Eq. (6). To estimate the temporal variance, we compute the second derivative value between the front and rear frames centered at the current frame to measure any subtle changes in the radius distances.

\[
\frac{1}{N_{\text{radi}}} \sum_{i=1}^{N_{\text{radi}}} \left( |\text{Radi}_{i-1} - \text{Radi}_i| + |\text{Radi}_i - \text{Radi}_{i+1}| \right)
\]

(6)

where \( N_{\text{radi}} \) is the number of radius distances of each frame.

After estimating the variations in the shape feature using training data that include positive real fire regions and negative fire-colored rigid objects, we calculate the skewness to analyze shape patterns. Skewness is a parameter that describes asymmetry using a probability distribution, its value being zero when the distribution is symmetric, negative when the data are skewed left, and positive when the data are skewed right. The skewness value is then normalized to 0–1 for generating a probability density estimation [5]. After calculating the skewness for all the training data, we estimate the GMFs for fuzzification using the skewness
results instead of a heuristic triangular or trapezoidal membership function. The membership function that represents a positive fuzzy set for shape \( P_{\text{Shape}} \) is usually denoted by \( \mu_{P_{\text{Shape}}} \).

\[
\mu_{P_{\text{Shape}}}(x) = \exp \left( \frac{-(c_{P_{\text{Shape}}}-x)^2}{2\sigma_{P_{\text{Shape}}}} \right)
\]

where \( c_{P_{\text{Shape}}} \) and \( \sigma_{P_{\text{Shape}}} \) are the center and width, respectively, of the positive fuzzy set \( P_{\text{Shape}} \). The same method is used to estimate the membership function that represents a negative fuzzy set for shape \( N_{\text{Shape}} \); it is denoted by \( \mu_{N_{\text{Shape}}} \).

Fig. 4 shows the results of two GMFs, where each membership function shows a different distribution, with some overlap.

2.3. Gaussian membership functions for size variation of fire region

Because the size of fire regions also tends to change as time passes, for the same reason as the shape does, as shown in Fig. 5, we analyze the temporal size variation in the fire regions of three frames, \( t-1, t, \) and \( t+1 \), using Eq. (8). A description of the size properties of the fire region can be obtained by measuring the properties of the points belonging to the region.

\[
\frac{|\text{Size}_{t-1} - \text{Size}_t| + |\text{Size}_t - \text{Size}_{t+1}|}{2}
\]  

(8)

After estimating the variations in the size feature using the training data, we also calculate the skewness to analyze the size patterns. Then, the membership function \( \mu_{P_{\text{Size}}} \) for positive fuzzy set \( P_{\text{Size}} \) and another membership function \( \mu_{N_{\text{Size}}} \) for negative fuzzy set \( N_{\text{Size}} \) are modeled using GMFs, as shown in Fig. 6.

2.4. Gaussian membership functions for motion variation of fire region

Since fire regions tend to continuously move upward because of heat convection [3], the motion orientation is also estimated between three frames, \( t-1, t, \) and \( t+1 \), using Eq. (9). To obtain a precise and fast motion estimation, we estimate the motion orientation using eight discrete directions, regardless of the motion magnitude, applying an idea from our previous study [3]. In addition, the motion vector is estimated only from the outer region of the flame because the inner region tends to be static. After the motion for each positive and negative region has been estimated, the orientation of the motion is discretized into eight directions, and each discrete direction is coded as 1–8. The upward orientation ratio \( U \) for the \( i \)-th region in the three frames is then estimated by using Eq. (9).

\[
U = \sum_{i=1}^{3} \frac{P(M_i)}{3}
\]

(9)

where \( P(\cdot) \) is an indicator function that returns predefined motion scores according to the value of \( M_i \). Unlike the previous two features, the membership function \( \mu_{P_{\text{Motion}}} \) for the positive fuzzy set \( P_{\text{Motion}} \) and another membership function \( \mu_{N_{\text{Motion}}} \) for negative fuzzy set \( N_{\text{Motion}} \) are modeled using the upward orientation ratio and revising the original GMFs. \( U \) exhibits a higher value when it is greater than the average \( (m_U) \) of \( \mu_{P_{\text{Motion}}} \). Using the same method, \( U \) exhibits a higher value when it is smaller than the average \( (m_N) \) of \( \mu_{N_{\text{Motion}}} \). In Eqs. (10) and (11) [3], the GMFs for the motion orientation \( \mu_{P_{\text{Motion}}} \) and \( \mu_{N_{\text{Motion}}} \) are generated by

\[
\mu_{P_{\text{Motion}}}(U) = \begin{cases} 
  e^\left(-\frac{(U-m_U)^2}{\sigma_U^2}\right) & U < m_U \\
  1 & U \geq m_U
\end{cases}
\]

(10)

\[
\mu_{N_{\text{Motion}}}(U) = \begin{cases} 
  e^\left(-\frac{(U-m_U)^2}{\sigma_U^2}\right) & U \geq m_U \\
  1 & U < m_U
\end{cases}
\]

(11)

where \( m_U, \sigma_U, m_N, \) and \( \sigma_N \) are the mean and standard deviation values for the value of a positive fuzzy set, \( P_{\text{Motion}} \), and negative fuzzy set, \( N_{\text{Motion}} \), respectively. Therefore, the original GMFs of \( \mu_{P_{\text{Motion}}} \) and \( \mu_{N_{\text{Motion}}} \) are changed into the form shown in Fig. 7.

Fig. 8. Diagram of a stereo vision system and the processing of feature point selection. (a) Basic structure of the stereo vision system; (b) feature point matching between the left and right images based on cooperative stereo matching algorithm; and (c) selected corresponding feature points.
3. Fire verification using fuzzy logic and Gaussian membership functions

In this study, we use fuzzy logic to verify that candidate fire regions are in fact fire regions. Fuzzy logic is the part of artificial intelligence or machine learning that interprets human actions, such as the degree of truth or degree of falseness, and has been applied mostly to sophisticated control systems and decision-support expert systems [5]. In a fuzzy system, the input variables are mapped into fuzzy values using the appropriate membership functions. In the processing steps, fuzzy values are applied to each appropriate rule, and then the results of the rules are combined. In the final step, the combined result is converted back into a specific control output value [16].

Fuzzy logic consists of decision rules (if-then rules) and uses critical variables to interpolate the output between crisp boundaries. The fuzzy rules used in this study are based on three input variables, one output variable, and eight rules. The input variables are the shape variation (SV), size variation (SZV), and motion orientation (MO), all of which have two types of fuzzy value, Positive and Negative, whereas the one output variable Fire has five types of fuzzy values, VeryHigh (VH), High (H), Median (M), Low (L), and VeryLow (VL) as in [5]. The eight rules used for flame verification are as follows:

Rule 1: IF SV is Positive and SZV is Positive and MO is Positive, then Fire is VH, else
Rule 2: IF SV is Positive and SZV is Positive and MO is Positive, then Fire is H, else
Rule 3: IF SV is Negative and SZV is Positive and MO is Positive, then Fire is H, else
Rule 4: IF SV is Positive and SZV is Positive and MO is Negative, then Fire is M, else
Rule 5: IF SV is Negative and SZV is Negative and MO is Positive, then Fire is M, else
Rule 6: IF SV is Positive and SZV is Negative and MO is Negative, then Fire is L, else
Rule 7: IF SV is Negative and SZV is Positive and MO is Negative, then Fire is L, else
Rule 8: IF SV is Negative and SZV is Negative and MO is Negative, then Flame is VL.

To obtain a crisp output from the fuzzy system, the Larsen implication rule [17] is applied to the fuzzy rules. The composition process combines all the output membership functions assigned to each output variable to form a single fuzzy set. Here, “else” in each rule can be interpreted as OR (\(\lor\)) and a final consequent fuzzy set is constructed by the inference rule (fuzzy logic OR) using Eq. (12):

\[
\mu_{\text{final}}(\text{Fire}) = \phi[\text{Rule 1}] \lor \cdots \lor \phi[\text{Rule 8}]
\]

where \(\phi\) is the implication operator and the and connective is replaced with the product operator \(\phi[\mu_A(x), \mu_B(y)] = \mu_A(x)\mu_B(y)\). The fuzzy logic OR (\(\lor\)) connective is replaced with the max (union) operator. The result of the final consequent membership function is estimated after the max operator has been applied between the output function given by the eight rules. Finally, defuzzification is used to convert the fuzzy output set to a crisp number. In this study, we determine candidate fire regions to be real fire regions if the crisp output is over the 50%. A detailed explanation of the Larsen implication rule and defuzzification process is given in [5,17,18].

4. Distance estimation and 3D surface reconstruction of fire region

After segmentation of the fire regions from the left and right images, the feature points are extracted from the fire region and their disparities are also estimated for 3D surface reconstruction. One issue of 3D surface reconstruction is the computational runtime required for estimating the disparity for all pixels. However, not all pixels are necessary to construct a 3D mesh, and the depth information can be calculated from the representative feature points. Therefore, to reduce the runtime complexity, we extract only feature points satisfying the corner detection criteria. Since the inner areas of fire regions have weak edges and a low intensity gradient compared to the boundary of flame regions, we use the SUSAN corner detection algorithm [19] with a weak condition instead of the Harris corner detector to detect feature points inside the fire region. The advantages of the SUSAN operator are a fast computation time compared to a Harris operator, and robust results with unimpaired and unsmoothed images [19]. We first extract points from the left and right images using a SUSAN corner detector. However, not all of these points can be matched because stereo images contain occlusion edges and occluded regions [11] that are unique to a stereo pair, as

![Figure 9](https://example.com/fig9.png)

**Fig. 9.** Sample example of non-fire videos. (a) Video 6, (b) Video 7, (c) Video 8, (d) Video 9, and (e) Video 10.
Fig. 10. Comparison of the fire detection results using the proposed algorithm and two other algorithms. (a) True positive (TP), (b) False positive (FP), and (c) False negative (FN).
shown in Fig. 8(a). Therefore, we need a matching algorithm to find the corresponding points in the two images and a disparity estimation algorithm that uses the corresponding points for a depth calculation.

Matching algorithms can be classified as correlation-based and feature-based methods. In this study, we use cooperative stereo matching and occlusion detection methods that are based on the normalized correlation algorithm proposed by Zitnick and Kanade [20]. In Fig. 8(a), each element \((r, c, d)\) of the disparity space projects to the feature point \((r, c)\) in the left image and to the feature point \((r, c+d)\) in the right image. The cooperative algorithm is now summarized as follows [20].

1. Prepare a 3D array, \((r, c, d)\); \((r, c)\) for each feature pixel in the reference image, and \(d\) for the range of disparity.
2. Set the initial match values, \(Io\), using the function of image intensities that utilizes a simple normalizing correlation between the feature points of \((r, c)\) and \((r, c+d)\).
3. Iteratively update the matching values, \(ln\), using the updating function until the match values have no changes, which means the matching values have converged. The update function is constructed by combining the initial matching value with the inhibiting function. An iterative algorithm updates the matching values by diffusing the support among neighboring values, and inhibiting other values along similar lines of sight.
4. For each feature point \((r, c)\), find the element \((r, c, d)\) with the maximum matching value, as illustrated in Fig. 8(b).
5. If the maximum match value is higher than a threshold, a feature point is selected, as shown in Fig. 8(c), and disparity \(d\) for the feature point is calculated, otherwise, it is classified as an occlusion.

After the corresponding points have been selected and their disparities computed, we use the disparity value to calculate depth information, that is, the distance between the real fire and the stereo vision system, using

\[
Depth = \frac{b \times f}{x_l - x_R} \tag{13}
\]

where \(b\) is the baseline between the two cameras, \(f\) is the focal length of a camera, and \(x_l\) and \(x_R\) are the corresponding feature points of the left and right image separately.

Since the unit of \(b\) and \(f\) is millimeters (mm) and the unit of \(Depth\) is a pixel, we have to convert the pixel unit to the mm unit to estimate the real distance between the fire and the camera. For example, if \(b\) is 120 mm, \(f\) is 2.5 mm, and the disparity is 1 pixel, the \(Depth\) pixel can be converted to millimeters according to one pixel equaling 0.264583 mm.

\[
Depth = \frac{120 \text{ mm} \times 2.5 \text{ mm}}{1 \text{ pixel} \times 0.264583 \text{ mm}} = 1133.8597 \text{ mm} \approx 1.133 \text{ m} \tag{14}
\]

The real distance between the fire and the camera is computed by arithmetic averaging of the \(Depth\) of \(N\) feature points.

After the depth information has been calculated from the corresponding feature points, their 3D coordinates are computed using a 3D Voronoi diagram and Delaunay triangulation. Fig. 11 (d) shows the results of the 3D surface reconstruction using the 3D coordinates of the feature points and the Delaunay triangulation method. As shown in Fig. 11(d), the reconstructed 3D surface has a shape similar to that of the front view of the real fire.

### 5. Experimental results

The proposed fire detection system based on a stereo camera is designed to detect fire and estimate the distance from the burning point to the camera in real time, in order to facilitate fire suppression. We used a pre-calibrated stereo camera, the Bumblebee2 of Point Grey Inc. This camera has a focal length of 2.5 mm with a 100-degree HFOV, a baseline of 120 mm and a 1/3” Sony ICX 204 CCD sensor. The image resolution is 648 x 480; we resized the image to produce a 320 x 240-pixel image to reduce the computational time. The proposed fire detection system was implemented using an Intel Core 2-Quad processor PC with Visual C++ 2010 version. As the proposed system was designed to detect fire in real time, the average frame rate for fire detection was 12 fps, including 3D surface reconstruction.

To evaluate the performance of the proposed algorithm, we captured 13 types of stereoscopic videos, while varying the distance between the camera and the fire. The frame rate of the video data varied from 15 Hz to 30 Hz and the size of the input images was 320 x 240 pixels. Five videos (Videos 1–5) were captured from a 5-m distance and five videos (Video 6–10) from a 10-m distance as described in Table 1. The remaining three videos were captured from a distance of 15 m, 20 m, and 30 m from the camera.

Supplementary material related to this article can be found online at [http://dx.doi.org/10.1016/j.firesaf.2014.05.015](http://dx.doi.org/10.1016/j.firesaf.2014.05.015).

#### 5.1. Performance evaluation with related works

Because the images of the fire region were too small when the distance of the camera from the fire was greater than 10 m, we used the five videos captured from a distance of less than 10 m for the detection performance test. Sample images of the fire videos are shown in Fig. 11(a). Five non-fire videos that included fire-colored moving objects were used as test data, as shown in Fig. 9. To examine the proposed fire detection algorithm, we used three parameters that are commonly used to evaluate fire detection performance: True positive (TP), False positive (FP), and False Negative (FN). FP was obtained by testing all the detection boxes in the test data that overlapped by over 50% with any ground truth fire region. In contrast, TP was obtained by testing all the detection boxes in the test data that overlapped by less than 50% with any ground truth fire region. To evaluate the performance of the proposed algorithm, it was compared with that of Töreyin’s [1] and Ko’s [5] algorithms, which perform well among existing algorithms.

The comparative results are presented in Fig. 10. It can be seen that our proposed algorithm generally yielded a better fire

---

**Table 2** Fire distance estimation according to the change of real distance.

<table>
<thead>
<tr>
<th>Real distance (m)</th>
<th>Video</th>
<th>Frame no.</th>
<th>Estimated error (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Video 1</td>
<td>333</td>
<td>0.0089</td>
</tr>
<tr>
<td></td>
<td>Video 2</td>
<td>193</td>
<td>0.0627</td>
</tr>
<tr>
<td></td>
<td>Video 3</td>
<td>242</td>
<td>0.0089</td>
</tr>
<tr>
<td></td>
<td>Video 4</td>
<td>497</td>
<td>0.0119</td>
</tr>
<tr>
<td></td>
<td>Video 5</td>
<td>374</td>
<td>0.0089</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td></td>
<td>0.02026</td>
</tr>
<tr>
<td>10</td>
<td>Video 6</td>
<td>421</td>
<td>3.2243</td>
</tr>
<tr>
<td></td>
<td>Video 7</td>
<td>368</td>
<td>3.1974</td>
</tr>
<tr>
<td></td>
<td>Video 8</td>
<td>293</td>
<td>2.6149</td>
</tr>
<tr>
<td></td>
<td>Video 9</td>
<td>389</td>
<td>3.2485</td>
</tr>
<tr>
<td></td>
<td>Video 10</td>
<td>363</td>
<td>3.0254</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td></td>
<td>3.0621</td>
</tr>
<tr>
<td>15</td>
<td>Video 11</td>
<td>200</td>
<td>6.9857</td>
</tr>
<tr>
<td>20</td>
<td>Video 12</td>
<td>200</td>
<td>11.9857</td>
</tr>
<tr>
<td>30</td>
<td>Video 13</td>
<td>200</td>
<td>21.4514</td>
</tr>
</tbody>
</table>

Please cite this article as: B. Ko et al., Fire detection and 3D surface reconstruction based on stereoscopic pictures and probabilistic fuzzy logic, Fire Safety J (2014), [http://dx.doi.org/10.1016/j.firesaf.2014.05.015](http://dx.doi.org/10.1016/j.firesaf.2014.05.015)
detection performance than the other two methods. In terms of the average TP, our method achieved 91.2%, which was 7.5% higher than that achieved by Ko’s algorithm [5] and 22.7% higher than that achieved by Töreyin’s algorithm [1]. The best detection performance, 93%, was obtained for Video 1 and the worst detection performance, 74%, was obtained for Video 5, because it contained relatively small fire regions. These small regions were considered noise and were thus removed from all three

Fig. 11. Fire detection results for fire videos using proposed the method and processing of 3D surface reconstruction. (a) Fire region detection, (b) fire region segmentation, (c) feature point extraction, and (d) 3D surface reconstruction.
The average FP achieved using our proposed approach was 1.9%, which is similar to that achieved by Ko's algorithm [5] and 0.7% lower than that achieved by Töreyin's algorithm [1], as shown in Fig. 10(b). Because Videos 2 and 3 included flames that were moving rapidly because of the wind, all three algorithms had higher FP rates for these two videos.

The proposed method also yielded a significantly lower average FN rate, 9.0%, which is 7.3% lower than that yielded by Ko's algorithm [5] and 22.5% lower than that yielded by Töreyin's algorithm [1], as shown in Fig. 10(c). The main reason for the good performance as reflected by the TP, FP, and FN rates is that the proposed algorithm used GMFs that consider the spatial and temporal variation of three features, instead of heuristic thresholds. Moreover, from the detected candidate fire regions, the proposed method was able to remove false fire regions effectively by using probabilistic fuzzy logic. The three methods showed almost the same performance for five non-fire videos that included fire-colored moving objects. The average percentages of the TP, FP, and FN rates for these videos were also similar, being approximately 100%, 0%, and 0%, respectively. This means that the three algorithms can correctly distinguish fire-colored moving objects from real fire.

5.2. Analysis of distance estimation

The second purpose of this study was to estimate the distance between the camera and fire burning point so that the strength and amount of the water emitted by the water cannon to suppress a fire can be varied accordingly. Therefore, we compute the average distance using Eq. (1) using two parameters of the stereo camera, the baseline (120 mm) and the focus length (2.5 mm), as well as the estimated disparities of the feature points.

As shown in Table 2, the average distance error is about 0.02 m when the real distance is 5 m. However, the average distance error is increased when the distance between the camera and fire is greater than 5 m. This means that the distance estimate is influenced by the two parameters, the baseline and focus length. Therefore, an advanced stereo camera, for example, one having a longer baseline and high image resolution, is needed to calculate the exact distance when the fire is a great distance away.

Fig. 11 shows the fire detection results when the proposed method was applied to the test fire videos. Fig. 11(a) and (b) shows the fire detection and segmentation results, respectively, and Fig. 11(c) shows the feature points after the matching algorithm was implemented. Fig. 11(d) shows the results of the 3D surface reconstruction for which the feature points and their 3D coordinate information were used. The demonstration videos can be accessed at our web page: http://cvpr.kmu.ac.kr.

6. Conclusion

This study focused on a new fire detection method based on using a stereo camera for estimating the distance between the camera and the fire region and providing a 3D surface reconstruction of the fire. For fire detection, the Gaussian membership functions were learned from positive and negative image sequences based on the temporal variation of fire regions. Then, Gaussian membership functions were applied to a pre-defined fuzzy logic model, which was used for the final fire-region verification. From stereoscopic images of fire regions, feature points and their disparities were estimated for distance calculation and 3D surface reconstruction.

The experimental results showed that the proposed approach yields a more robust fire detection result than the other methods. In addition, when the current stereo system was used, the distance estimation was accurate when the real fire region was located within a 5-m distance from the camera. Therefore, it is suitable for application in an AFSS located in an indoor environment or when the fire is a short distance away from the camera in an outdoor environment.

7. Future works

Because we used a 320 pixel × 240 pixel image to reduce the computational time, the maximum distance for fire detection is limited to 30 m. Therefore, it is necessary to develop a rapid-fire verification and disparity estimation algorithm to detect fire and calculate its distance using a higher-resolution stereo image. In addition, although the 3D surface model for a fire region was reconstructed correctly, additional research is required to analyze the exact volume, spreading direction, and tilt angle of the fire to facilitate its rapid suppression using an AFSS.

References