IEEE Third International Conference on Data Stream Mining & Processing
August 21-25, 2020, Lviv, Ukraine

3D Fire Front Reconstruction in UAV-Based Forest-Fire Monitoring System

Vladimir Sherstjuk
Department of Software Tools and Technologies
Kherson National Technical University
Kherson, Ukraine
vgsherstyuk@gmail.com

Maryna Zharikova
Department of Software Tools and Technologies
Kherson National Technical University
Kherson, Ukraine
marina.jarikova@gmail.com

Irina Dorovskaja
Department of Technologies
European University
Kryvyi Rih, Ukraine
irina.dora07@gmail.com

***Abstract*—This work presents a new method of 3D reconstruction of the forest-fire front based on uncertain observations captured by remote sensing from UAVs within the forest-fire monitoring system. The use of multiple cameras simultaneously to capture the scene and recognize its geometry including depth is proposed. Multi-directional observation allows perceiving and representing a volumetric nature of the fire front as well as the dynamics of the fire process. The novelty of the proposed approach lies in the use of soft rough set to represent forest fire model within the discretized hierarchical model of the terrain and the use of 3D convolutional neural network to classify voxels within the reconstructed scene. The developed method provides sufficient performance and good visual representation to fulfill the requirements of fire response decision makers.**

Keywords— forest fire monitoring; fire front; remote sensing; process reconstruction; neural network; voxel; soft rough set.

I. Introduction

Due to global climate change, industrialization, urbanization, population growth, and the other outstanding features of new century, intensive forest fires have become a whole-planetary problem. They grow year by year, so forest fire response operations become challenging and expensive. Traditionally, these operations continue to be based on visual observations and decision-maker's estimations. However, smoke and flame substantially distort observations, while high temperature does not allow to approach the fire closely. That is why, inaccurate and incomplete observations cannot be a reliable basis for planning response operations. Although a forest fire can be considered as a poorly modeled and unpredictable process, it is well known that its fast development requires always a high responsibility of decision-maker with acute lack of time. It is also known that the efficiency of response operations depends mainly on the availability and usability of real-time forest fire monitoring tools. Therefore, today a lot of attention is paid to unmanned vehicles, remote sensing, image processing, and a range of other modern tools and technologies that can be synergistically used to forest fire monitoring. Since unmanned aerial vehicles (UAV) can fly closely enough to the fire carrying optical and infrared cameras as remoted sensors, they are the most suitable tools for the real-time forest fire monitoring [1].

However, obtaining a real and reliable picture of the ongoing fire spreading process is not an easy task due to effects of wind, smoke, and fire. There are many other factors such as multiplicity of fire spots, segmentation of the fire front, spatial and temporal variability of weather, fuels, and topography conditions, which make also their contribution. All of this complicates obtaining credible grounds for decision-making on planning and executing fire response operations. Decision-maker always requires a clear picture of the ongoing processes in order to understand a direction and a rate of fire spreading. At the same time, existing implementations of real-time forest fire monitoring systems offer a result in the form of a flat two-dimensional image of a burning area on a map [2]. Thus, the topic of our interest is a study of the ways of 3D reconstruction of the forest fire process during its real-time monitoring, which can provide decision-maker with a clear model of the fire front spreading to make in-time decisions on the fire response.

II. Literature Analysis

The considered problem is very close to three-dimensional (3D) object reconstruction that is a computer vision task having several applications including robotics, object tracking, etc. The object reconstruction is always reduced to taking a set of scans of the object's surface with a certain sensor or scanner, which must be located in different points due to both a limited field of view and occlusions of the objects [3]. A 3D model of the object can be generated by a physical sensing its surface from several views. Since objects have various shapes and sizes, and sensor locations have certain observation constraints, it is necessary to choose different views carefully. Usually, the set of views around the object can be chosen manually by operator [4] (in process or predefined), or automatically by a view planning algorithm finding proper locations for sensors [5].

Today, there are many well studied online and offline object reconstruction algorithms [6]. Online Simultaneous Localization and Mapping (SLAM) algorithms are based on an incremental appearance-based loop closure detector and usage of range sensors such as laser or stereo cameras. They are quite robust, have revisiting capability and work in real time. However, in general they provide insufficient opportunities for the detailed object reconstruction due to sufficiently sparce observations, although the shapes, features and colors can be recognized as good as possible [7]. The detailed review of SLAM odometry-, depth sensor-based, and other methods including their advantages and performance are presented in [8]. Offline algorithms such as Multi-View Stereo (MVS) are aimed at finding pairwise stereo correspondences to estimate dense and accurate

978-1-7281-3214-3/20/$31.00 ©2020 IEEE

reconstructions [6]. Sparse feature matching and patch growing method with photometric and visibility constraints is proposed in [9]. In [10], a Next-Best View (NBV) approach is proposed to select feasible stereo pairs views. This approach is expanded in [11] to use multi-views. However, the computational complexity of such methods is too high to use them in the real-time. A volumetric method based on hierarchical octree reconstruction was also proposed in [12]. Such methods are divided into real 3D methods, which describe the modeled object by the 3D point clouds returned by sensors, and 2.5D methods, which describe only measured height (depth) for each cell within a certain 2D grid. Point clouds can not distinguish free and unknown areas, so they are mainly suitable in static environments [13]. Besides that, they are sensitive overmuch to sensor noise. Height (depth) measurements represented as elevation map are mainly used to model the outdoor environment having a single surface [14]. Thus, they only discretize the environment vertically but do not provide its volumetric representation [15]. To overcome this problem, a hybrid approach is proposed in [16], where each cell in a 2D grid stores a list of vertically ordered voxels. Another important area of improvement is concerned with the use of Next-Best View algorithms to enhance the efficiency of the preformed reconstruction. NBV methods can improve quality and coverage using a minimum amount of data [17].

It should be noted that forest fire is not an object but a dynamic process, within which the presence of a flame is a certain “eigenfeature”. From the image recognition point of view, such eigenfeature is significantly variable with respect to its shape and color. There are also effects of smoke, flares, and flickers, which significantly complicate the recognition and distort the picture of the process. Moreover, forest fires are always spreading, and a rate of fire spreading is also variable under the influence of wind, type of vegetation, and other factors. Despite the overall progress in the field of object reconstruction, the issues of 3D process reconstruction based on multiple observations still remain open and have a little reflection in the literature. We can reduce the problem of the fire process reconstruction to the problem of the reconstruction of the fire front. However, this process has such decisive dynamics that all above-considered methods are poorly applicable to solve the reconstruction problem. The solution of the process reconstruction problem must be three-dimensional (volumetric), must take into account the inaccuracy and uncertainty of the observations and provide adequate representation of the dynamics of the real process. The images captured by the cameras mounted on the UAVs can provide information necessary for such reconstruction. However, the geometry of the fire front does not allow observing the process completely from one viewpoint, even in the case of the continuous movement of the camera. Since the viewpoint position cannot get to the fire front closer than a certain safe distance, a simultaneous observation from several different points is required to solve the reconstruction problem. Since several issues related to the forest fire front reconstruction still remain insufficiently investigated, this needs further research.

III. Problem Statement

A forest-fire monitoring system must provide real-time information to the decision maker during response operation. The most important information regarding the forest fire process is information about the dynamics of the fire front. Understanding this dynamic enables decision-maker to assess fire intensity based on the rate of fire front spreading and the estimated length/height of flame. Thus, reconstruction of the fire front should mainly be aimed at the determining its geometric parameters within the spatial and temporal scale. Actually, forest fire monitoring systems involve UAVs equipped with optical and infrared cameras to remote sensing that allows identifying the features of flame and smoke on the base of the processing of the captured images [18]. Although both optical and infrared sensors are sensitive to noise and interferences so the observations are often ambiguous, imprecise, and inconsistent, nonetheless, they can be used for obtaining the feasible estimates of the fire front parameters. The main questions will be how accurate the reconstruction of the fire front is and how much the degree of its credibility is. Fortunately, in most cases, decision-maker does not require precise estimates of fire front parameters, but he needs to know how the fire front evolves over time and understand the course of the process and its drivers. Clearly, a visual representation of the fire front dynamics can be a valuable information for making adequate decisions.

Therefore, the aim of this work is to develop a method of 3D reconstruction of the fire front based on uncertain observations captured by remote sensing from UAVs within the forest fire monitoring system. We propose to use multiple cameras simultaneously to capture the scene and recognize its geometry. We assume that multi-directional views of the process can be used to estimate volumetric nature of the fire front as well as the fire processes.

IV. Remote Sensing and Image Processing in Forest-Fire Monitoring System

A. *Observable attributes of the forest fiore*

The main attributes of the forest fire are heat, smoke, and flame, including such its manifestations as light, flicker, and motion [19]. All of them are observable by sensors.

Flame emits their own visible light and can be considered as “eigenfeature” of the forest fire. Usually, flame has also such visible properties as flickers, flares, movement, and transparency [19]. Depending on the temperature, the flame color can vary from dark red to light yellow and even up to white at the developed stage of the fire [20]. Despite of the variability of its color, flames are usually distinctive to recognition on images with respect to the background. Another visual feature of flame is its shape. However, it can essentially vary depending on fuel consuming and composition, wind variation, etc. Thus, the flame shape recognition is more challenging. Obviously, the image recognition can be complicated by interferences (i.e., sunlight) and distortions (shadow, smoke, flares, and flickers) that affect images. Moreover, smoke and precipitation must be considered as noise.

Smoke is always a reliable visual attribute for the detection of forest fires but during the fire monitoring it is rather an occlusion that masks eigenfeature and obscures the visibility of flame. Smoke has also dynamic properties changing fire’s shape, size and color as well as rising in plumes. Usually, smoke can be recognized through low values of chrominance in captured images [19].

Heat is usually transported from the fire by convection, conduction, and radiation but only radiated heat can be remotely sensed and measured by infrared cameras.

Unfortunately, such measurements do not have a depth effect, so they cannot be directly used to reconstruct a fire process. However, images captured by infrared camera can be used to unmask a flame covered by smoke.

B. Sensors

Modern electro-optical cameras have a high enough resolution and wide field of view, but the quality of their images depends heavily on lighting conditions. Thus, they are sensitive to darkness (nighttime), smoke and precipitations such as rain, haze, mist, fog, etc. Although the image processing algorithms are mainly based on color properties, they should also analyze flickering of pixels over time as well as intermittency and irregular oscillations of the edges of the flame regions since the latter usually vary in height, size, and brightness. Smoke can be identified by low values of chrominance and variations in color and density.

The broadband thermal infrared camera measures energy release within the combustion reaction but it has a limited dynamic range. Images captured by infrared (IR) camera are usually affected by such interferences as saturation, reflected sunlight, energy radiated from non-fire sources, etc. In daytime, its images have too small contrast. However, IR camera can be used overnight due to good flame-to-background contrast and under the smoke conditions because smoke is quite transparent on the corresponding wavelength.

Thus, both optical and infrared cameras have their drawbacks and can provide imprecise, uncertain, or ambiguous information in captured images because of interferences and distortions. Turbulence and vibrations of UAV distort captured images additionally. Therefore, we cannot build a clear and accurate model of the fire front based on remote sensing information, but we can reconstruct a model of the fire front approximately.

C. Remote Sensing

Monitoring UAVs need to have hovering capabilities providing a respectively long hovering time. Usually, multi-rotary-wing UAVs equipped with a gimbal carrying both infrared and high-resolution electrooptical camera with pan and tilt units are used for monitoring purposes. In [2], the monitoring UAVs are equipped with the 16-Mp optical camera (5376x3024 pixels) providing scalability of images, the thermal infrared camera, weather sensors, GPS receiver, and inertial measurement unit for self-localization and navigation purpose.

Monitoring UAVs capture images flying towards the fire and hovering around it. Obviously, such UAVs should perform onboard only stabilization, geo-localization, and geo-rectification of images (Fig. 1). Their sensors capture and transmit images to the ground control station. The digital elevation model represents cartographic dataset describing terrain surface, so the captured images are transformed to a stream of geo-mapped frames where each frame is complemented with the coordinates of the upper left corner of the frame and scale value. This makes it possible to merge images taken from different positions of observation.

D. Image processing

It is assumed that that fire spreads in three-dimensional space C above the terrain discretized by a grid $D=\{d_{xyz}\}$ of isometric cubic cells d_{xyz} with the size being $\delta\times\delta\times\delta$.

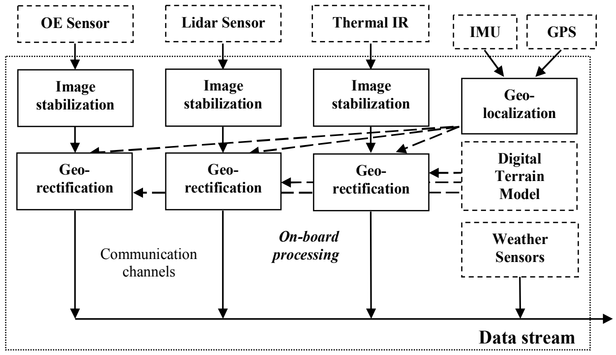

Fig. 1. Onboard image processing

There are three channels to process image from each UAV at the ground center: one for image frames captured by infrared camera and two separate channels providing flame and smoke recognition based on the image frames captured by optical cameras (Fig. 2). In the infrared channel, image pixels represent a heat radiation by colors ranged from black to white, so the image analysis is performed in three stages: image mapping, image averaging, and gray color evaluation. Burning areas are always represented as white areas within the image, while non-burning areas are black. Clearly, a plenty of pixels is greyed due to uncertainty of observations. As the result of the image analysis, each cell $d_{ijk}\in D$ is associated with a certain “degree of grayness” μ_{ijk} ranged in the interval $[0..1]$ and based on the average brightness B_{ijk} of this cell.

In the optical channels, the process of image analysis is performed in such sequential stages as mapping, transformation (only in optical camera channels), averaging, filtering, generalization, and conversion of images. This analysis is quite similar in smoke and flame channels and performed in HSI color space, but the differences are in the intervals of processed hue and saturation values. In the flame channel, color can vary from dark red to light yellow with the high saturation while in the smoke channel, color can be in gray-and-silver interval with low saturation but high intensity. As the result of image processing, each cell d_{ijk} is associated

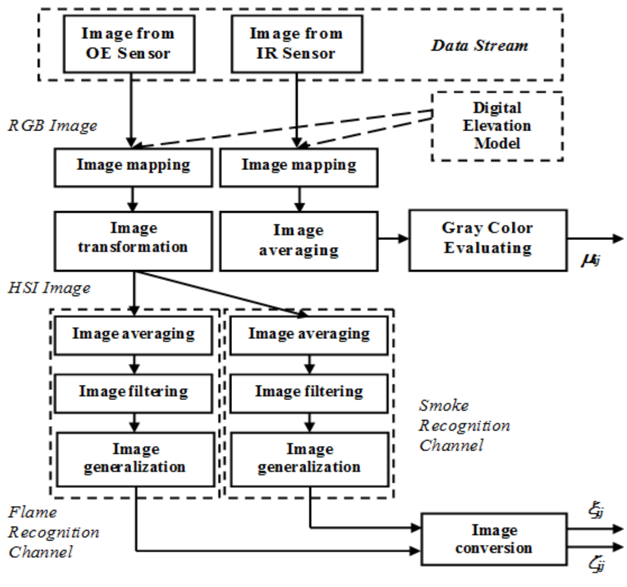

Fig. 2. The image analysis at the ground center

with certain degrees of flame (ξ_{ijk}) and smoke (ζ_{ijk}) recognized within this cell and normalized in the range [0, 1].

Finally, the values ξ_{ijk} , ζ_{ijk} , and μ_{ijk} are processed into merged value η_{ijk} with respect to each cell d_{ijk} . The final stage of the image analysis consists of image filtering, cleaning and merging. Such multi-stage processing algorithm allows reducing interferences, distortions, and noise during remote sensing. However, the implementation of this method in [2] implies averaging images captured from different points of view (different UAVs) and projecting the resulting grayscale image frame based on values η_{ijk} for the cells $d_{ijk} \in D$ into a two-dimensional plane $d_{ij} \in D'$ lying "at the ground" to recognize the flat forest fire front.

V. METHOD OF 3D FIRE FRONT RECONSTRUCTION

To obtain the volumetric fire front representation, we must reconstruct it from uncertain observations based on the results of the image analysis described in the previous section. However, such reconstruction is quite challenging because of fire spreading. Fig. 3 shows the forest fire observed by UAVs' sensors from two distinct viewpoints. Obviously, a part of the fire process is masked with smoke, which spreads out mainly in the wind direction. This makes it impossible to directly observe the combustion process covered by smoke. Therefore, the viewpoints for observation are often selected at the opposite sides (windward A vs leeward B in Fig. 3) of the fire front. Nevertheless, a significant part of the combustion process remains hidden due to presence of hidden areas and occlusions in the observed scene such as shown in Fig. 3. Furthermore, fire fronts can cover a sufficiently deep area that is poorly visible from two opposite viewpoints. Often, they are segmented and even broken up into several parts.

The fire front spreads in such a way that the fuel burns out inside of the fire perimeter remaining burned-out areas, but due to the environmental effects, new areas of vegetation are covered by burning outside the fire perimeter. Accordingly, monitoring UAVs should move avoiding unsafe areas and choosing positions that can maximize gaining of observable information overcoming uncertainty.

A. Scene geometry

The forest fire scene is represented in Fig. 4. A configuration of viewpoints can be defined as a set of poses (Pose A, B, C in Fig. 4). The pose is a description representing the UAV location and the sensor orientation

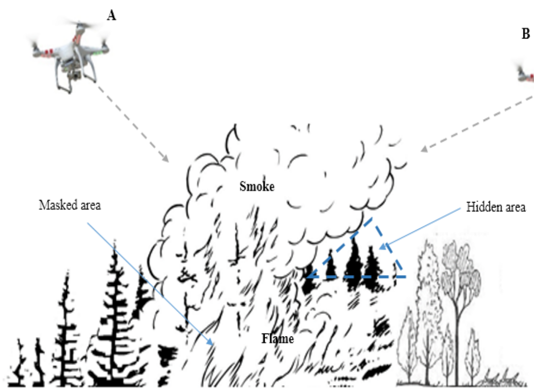

Fig. 3. Multi-view fire front monitoring

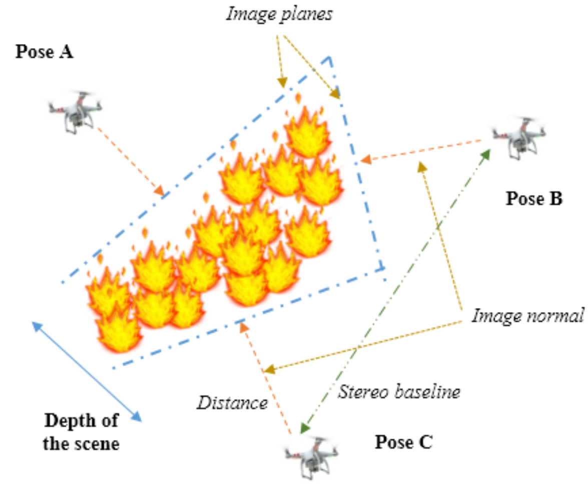

Fig. 4. The reconstructed scene

within three-dimensional space. Usually, the sensor orientation is chosen in accordance with the normal to the image plane. Each pair of UAVs can be involved in a stereo-pair view; therefore, they must be located in accordance with stereo baseline determined through the crossing of their image normals at the certain point at the maximal depth of the scene. Thus, images from UAV cameras can be used to scene depth perception both on its own (through stereo-view) and with the help of additional neural network processing. Determining a set of optimal poses is a separate independent research task, which will not be considered in this paper.

B. Spatial Representation

The grid of cells D used in the image processing at the ground center is not suitable for the information presentation during the process reconstruction because of the spreading of the fire front. Thus, the initial grid needs to be extended with new spaces dynamically along with fire spreading process while burned-out areas lose their value. Consequently, we must initially build a space such big as the maximal bounds of the considered forest fire, but they are not known. Therefore, we cannot use the grid D due to performance requirements and memory limitations.

We create a new 3D structure to reconstruct fire front based on a grid of voxels, which are considered as certain cubic volumes of equal size. Although voxels are a bit like cells that discretize a space in a grid D , nonetheless, they are organized in a completely different way. Each voxel is considered as a node in a certain tree-like structure called octree. Each node of the octree can be recursively divided into 8 sub-nodes, each of which is also a voxel but has a smaller size. Such recursive process can be both top-down and bottom-up. Initially, we can start from the voxel that covers the overall forest fire area and subdivide it recursively down to the minimum voxel size reachable with respect to the sensor resolution. After the fire front achieves a bound of the discretized area, we can just move a step up and create a voxel containing the original voxel as a sub-node. As the result, we describe a hierarchical data structure based on octrees that can be cutted at any level of hierarchy. Such hierarchical data structure known as OctoMap can be used to representation of the reconstruction scene, where voxels constitute a certain 3D vector where each unit is a voxel. We use a revised version of an open source library OctoMap [21].

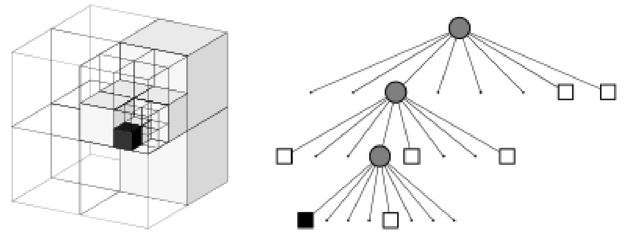

Fig. 5. OctoMap data structure [21]

C. Scene Representation

Within the scene, each voxel belongs to a specific class w_i, which describes the type of its inner content within the considered reconstruction scene:

- Unmarked ($i = 0$). It is a type of voxels that have not been seen by sensors. Initially, the reconstruction scene is represented by a 3D vector filled by unmarked voxels.
- Empty ($i = 1$). It is a type of voxels that lie between the sensor position and sensed “surface” of the fire process represented by flame and smoke. Such voxels usually represent a “free of fire” space.
- Flame ($i = 2$). These voxels represent a remotely sensored “surface” of the fire. Such voxels correspond to a burning kernel of the fire process.
- Smoke ($i = 3$). This type of voxels represents the areas within the scene shrouded in smoke, which prevents sensoring the inner voxels. Such voxels can mask “surface” of the fire.
- Uncertain ($i = 4$). It is a type of voxels that can not be precisely assessed as “burning”, i.e. voxels that are possibly involved in the combustion process due to the uncertain observation but there is no certainty about it.
- Burnt ($i = 5$). Such voxels represent the areas that are already burnt and, therefore, cannot be involved in the combustion processes.
- Fuel ($i = 6$). This type of voxels corresponds to the vegetation areas that do not participate in combustion processes but can be ignited due to readiness of the fuel. It should be noted that voxels, which do not contain flammable vegetation, belong to the “empty” class instead of the “fuel” class.
- Occluded ($i = 7$). These voxels represent the areas occluded by other voxels that prevent perception of the fire front depth.

D. Classification of Voxels

A 3D convolutional neural network (3D-CNN) can be used to classify voxels within the reconstructed scene. It is a special kind of CNNs using the 3D volumes in the kernels instead of 2D maps. We propose to use VoxNet [22] architecture (Fig. 6). The first convolutional layer has 3 channels with 128 features, a kernel of size 5×5×5, stride 2×2×2, and max pooling operation of stride 2×2×2; the second one has 32 features, a kernel of size 4×4×4, stride 2×2×2, and max pooling operation of stride 2×2×2; and the third one - 8 features, a kernel of size 3×3×3, stride 2×2×2, and max pooling operation of stride 2×2×2. Three fully connected layers have 1200, 400, and 50 parameters as output.

E. Soft Rough Fire Front Model

The model of the fire front can be represented using soft sets of cells [23]. Considering a set of voxel classes $W = \{w_0, ..., w_7\}$ and a 3D vector of voxels V, we can define a *soft set of voxels* over V as a pair (Υ, W) where Υ is a mapping of W into the set of all subsets of the set V. This set can be defined as $\Upsilon(t) = \{(w_i, \Upsilon(w_i, t)) : w_i \in 2^W, \Upsilon(w_i, t) \in 2^V\}_{i=0}^{7}$, where each $\Upsilon(w_i, t)$ is an w_i-element of the soft set (a set of voxels of a class $w_i \in W$ at the reconstruction moment $t \in T$). Using the voxels of a class w_2, we can define a lower approximation containing the cells, which definitely belong to the w_2-element of the soft set $\underline{\Upsilon}(t) = \Upsilon(w_2, t)$, while using the voxels of classes w_3 and w_4, we can define an upper approximation containing the cells, which possibly belong to the w_3- and w_4- elements of the soft set, $\overline{\Upsilon}(t) = \Upsilon(w_3, t) \cup \Upsilon(w_4, t)$. As the result, both lower and upper approximations constitute a soft rough set of voxels $\hat{\Upsilon}(t) = \{\underline{\Upsilon}(t), \overline{\Upsilon}(t)\}$ that represent a 3D fire front model at the reconstruction time. Other voxels belong to the negative area of the rough set $NEG(\hat{\Upsilon}(t)) = V - \overline{\Upsilon}(t)$. The boundary area of the fire process is a subset of the set of voxels, which belong to the upper approximation, but don’t belong to the lower approximation, $BND(\hat{\Upsilon}(t)) = \overline{\Upsilon}(t) - \underline{\Upsilon}(t)$.

VI. IMPLEMENTATION

The proposed method of 3D reconstruction of the forest fire front has been implemented using Visual C++, OctoMap framework, ConvNet and Fast Artificial Neural Network (FANN) libraries. The software prototype has been tested on PC based on the Pentium i5-7400 3-3,5 GHz processor and 16 GB RAM, the reconstruction rate has been evaluated during the simulation. The initial weights of neural networks have been randomly generated. The simulation results show that the method can achieve an accuracy of reconstruction up to 98% (Fig. 7), which shows the rate of correctly classified voxels (true positive) from the prepared test set. Thus, the method is effective for the forest fire front reconstruction, it has

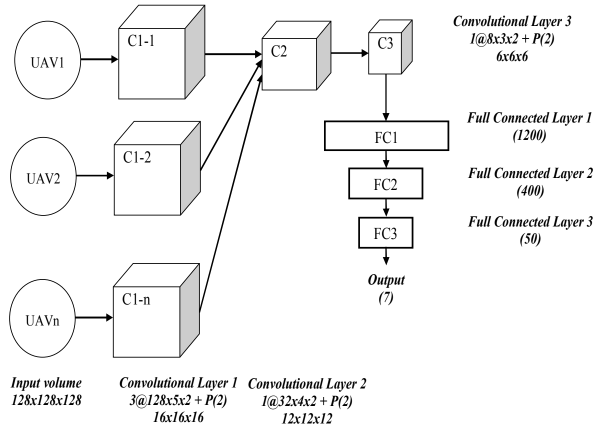

Fig. 6. 3D-CNN Architecture

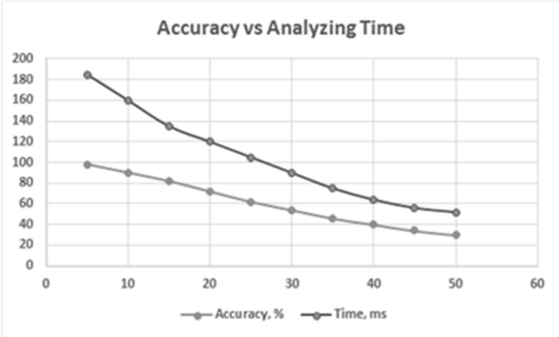

Fig. 7. Simulation results

relatively the same performance as the method of the fire front recognition proposed in [24]. At the same time, reconstructed 3D model of the fire front gives to decision-maker a significantly better visual representation of the forest fire dynamics (Fig. 8), so the credibility of fire response decision making is respectively increased, which makes the proposed method of practical use.

VII. Conclusion

The proposed in the paper method of 3D reconstruction of the forest fire front is based on uncertain observations captured by remote sensing from UAVs within the forest fire monitoring system. We suggest to use multiple cameras simultaneously to capture the scene and recognize its geometry including depth. Multi-directional observation allows us to perceive and represent a volumetric nature of the fire front as well as the dynamics of the fire processes. The novelty of our approach lies in the use of soft rough set to represent forest fire model within the discretized hierarchical model of the terrain as well as in the use of 3D convolutional neural network to classify voxels within the reconstructed scene. The developed method provides sufficient performance and good visual representation to fulfill the requirements of fire response operations.

References

[1] L. Merino, J. Martínez de Dios, and A. Ollero, “Cooperative Unmanned Aerial Systems for Fire Detection, Monitoring, and Extinguishing,” In Handbook of Unmanned Aerial Vehicles, Springer, Netherlands, 2015, pp. 2693–2722.

[2] V. Sherstjuk, M. Zharikova, and I. Sokol, “Forest Fire-Fighting Monitoring System Based on UAV team and Remote Sensing,” In Proc. of 2018 IEEE 38th Int. Conf. on Electronics and Nanotechnology. IEEE, Kyiv, 2018, pp. 663–668.

[3] J. I. Vasquez-Gomez, L. E. Sucar, R. Murrieta-Cid. and E. Lopez-Damian, “Volumetric Next-Best-View Planning for 3D Object Reconstruction with Positioning Error,” Int. J. of Advanced Robotic Syst., vol. 11, no. 10, pp. 1–13, 2014.

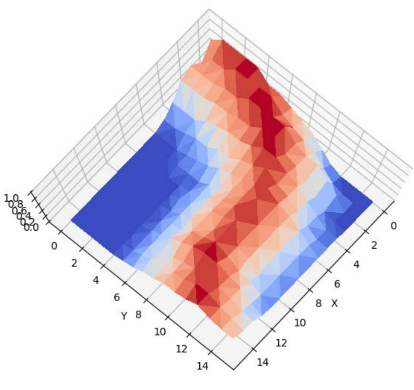

Fig. 8. Representation of the forest fire front

[4] V. Lippiello, F. Ruggiero, B. Siciliano, and L. Villani, “Visual grasp planning for unknown objects using a multifingered robotic hand,” in IEEE Trans. on Mechatronics, vol. 18, no. 3, pp. 1050–1059, 2013.

[5] W. R. Scott, G. Roth, and J.-F. Rivest, “View planning for automated three-dimensional object reconstruction and inspection,” ACM Comput. Surv., vol. 35, no. 1, 2003, pp. 64–96.

[6] O. Mendez, S. Hadfield, N. Pugeault and R. Bowden, "Taking the Scenic Route to 3D: Optimising Reconstruction from Moving Cameras," 2017 IEEE Int. Conf. on Computer Vision (ICCV), Venice, 2017, pp. 4687–4695.

[7] R. Mur-Artal, J. M. M. Montiel and J. D. Tardós, "ORB-SLAM: A Versatile and Accurate Monocular SLAM System," in IEEE Transactions on Robotics, vol. 31, no. 5, pp. 1147–1163, Oct. 2015.

[8] J. Delmerico, S. Isler, R. Sabzevari, D. Scaramuzza, “A comparison of volumetric information gain metrics for active 3d object reconstruction,” Autonomous Robots, vol. 42, pp. 197–208, 2018.

[9] Y. Furukawa and J. Ponce, "Accurate, Dense, and Robust Multiview Stereopsis," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 8, pp. 1362–1376, 2010.

[10] M. Jancosek, A. Shekhovtsov and T. Pajdla, "Scalable multi-view stereo," 2009 IEEE 12th Int. Conf. on Computer Vision Workshops, ICCV Workshops, Kyoto, 2009, pp. 1526–1533.

[11] S. Galliani, K. Lasinger and K. Schindler, "Massively Parallel Multiview Stereopsis by Surface Normal Diffusion," 2015 IEEE Int. Conf. on Computer Vision (ICCV), Santiago, 2015, pp. 873-881.

[12] S. Isler, R. Sabzevari, J. Delmerico and D. Scaramuzza, "An information gain formulation for active volumetric 3D reconstruction," 2016 IEEE Int. Conf. on Robotics and Automation (ICRA), Stockholm, 2016, pp. 3477–3484.

[13] A. Nüchter, K. Lingemann, J.Hertzberg, and H. Surmann, “6D SLAM – 3D mapping outdoor environments,” J. Field Robot., vol. 24, no. 8–9, pp. 699–722, 2007.

[14] R. Triebel, P. Pfaff and W. Burgard, "Multi-Level Surface Maps for Outdoor Terrain Mapping and Loop Closing," 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, 2006, pp. 2276-2282.

[15] J. S. Gutmann, M. Fukuchi, M. Fujita, “3D perception and environment map generation for humanoid robot navigation. Int. J. Rob. Res. Vol. 27, no. 10, pp. 1117–1134, 2008.

[16] J. Ryde, H. Hu, “3D mapping with multi-resolution occupied voxel lists,“ Autonomous Robots, vol. 28, no. 2, pp. 169–185, 2010.

[17] E. Marchand and F. Chaumette, "Active vision for complete scene reconstruction and exploration," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 1, pp. 65–72, 1999.

[18] V. Sherstjuk, M. Zharikova and I. Sokol, "Forest Fire Monitoring System Based on UAV Team, Remote Sensing, and Image Processing," 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2018, pp. 590–594.

[19] R. Allison, J. Johnston, G. Craig, and S. Jennings, “Airborne Optical and Thermal Remote Sensing for Wildfire Detection and Monitoring,” Sensors, vol. 16, no. 8, 1310, 2016.

[20] Y. Qiang, B. Pei, and J. Zhao, “Forest Fire Image Intelligent Recognition based on the Neural Network,” Journal of Multimedia, vol. 9, no. 3, pp. 449–455, 2014.

[21] A. Hornung, K.M. Wurm, M. Bennewitz, C. Stachniss, W. Burgard, “Octomap: An efficient probabilistic 3d mapping framework based on octrees,” Autonomous Robots, vol. 34, no. 3, pp. 189–206, 2013.

[22] D. Maturana and S. Scherer, "VoxNet: A 3D Convolutional Neural Network for real-time object recognition," 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, 2015, pp. 922–928.

[23] D. Molodtsov, “Soft Set Theory – first results,” Comput. and Math. with Appl., vol. 37, pp. 19–31, 1999.

[24] V. Sherstjuk and M. Zharikova, "Fire-Front Recognition in UAV-Based Forest-Fire Monitoring System Using Fuzzy Rough Soft Sets," 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), Lviv, Ukraine, 2019, pp. 1091–1096.