# Measurements of wildland fires by drone

Vito Ciullo

▶ **To cite this version:**

Vito Ciullo. Measurements of wildland fires by drone. Modeling and Simulation. Université Pascal Paoli, 2020. English. NNT : 2020CORT0005 . tel-03721162

**HAL Id: tel-03721162**
**https://theses.hal.science/tel-03721162**

Submitted on 12 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**UNIVERSITE DE CORSE-PASCAL PAOLI**
**ECOLE DOCTORALE ENVIRONNEMENT ET SOCIETE**
**UMR CNRS 6134 (SPE)**

**Thèse présentée pour l'obtention du grade de**

**DOCTEUR EN AUTOMATIQUE, SIGNAL, PRODUCTIQUE, ROBOTIQUE**

**Mention : Génie Informatique, Automatique et Traitement du Signal**

**Soutenue publiquement par**

**VITO CIULLO**

le 17/07/20

---

# Measurements of wildland fires by drone

---

**Directeur :**
Mme Lucile ROSSI, Dr-HDR, Université de Corse

**Rapporteurs :**
M Simon LACROIX, Dr, Centre de Toulouse LAAS/CNRS
M Yvon VOISIN, Professeur, Université de Bourgogne

## Jury

M Paul-Antoine BISGAMBIGLIA, Dr-HDR, Université de Corse
M Paul-Antoine BISGAMBIGLIA, Professeur, Université de Corse
M Simon LACROIX, Dr, Centre de Toulouse LAAS/CNRS
Mme Lucile ROSSI, Dr-HDR, Université de Corse
M Franck RUFFIER, Dr, Institut des Sciences du Mouvement UMR 7287
M Yvon VOISIN, Professeur, Université de Bourgogne


# THESIS


**Submitted for the degree of**

## DOCTOR OF PHILOSOPHY

**Thesis at** *University of Corsica (Pasquale Paoli)*

*2020/07/17*

---

**A dissertation presented by:**

**Vito CIULLO**

**Measurements of wildland fires by drone**

---

**GRADUATE SUPERVISORY COMMITTEE:**

| | | |
|---|---|---|
| Lucile ROSSI | Associate Professor | University of Corsica |
| Simon LACROIX | Research Director | LAAS/CNRS Robotics and InteractionS team (Toulouse) |
| Yvon VOISIN | Professor | University of Bourgogne |
| Franck RUFFIER | Research Director | Institute of Movement Science (Marseille) |
| Paul-Antoine BISGAMBIGLIA | Professor | University of Corsica |
| Paul-Antoine BISGAMBIGLIA | Associate Professor | University of Corsica |

---

**Discipline:**
*Automatic and Signal Processing, Productive, Robotics, Computer Engineering*

**Doctoral School and Specialty:**
*ED 377: Environnement et Société*

**Research Unit:**
*UMR CNRS 6134 SPE*

**Thesis Advisor:**
*Lucile ROSSI*

**Examiners:**
*Yvon VOISIN*
*Simon LACROIX*

UMR 6134 - Laboratoire Sciences Pour l'Environnement SPE, CNRS - Université de Corse Pascal Paoli, Campus Grimaldi, Bâtiment PPDB 3eme étage, Avenue du 9 Septembre, Corte - 20250, Haute-Corse, France

**Riassunto** — Questo lavoro di tesi presenta l'estimazione delle misure delle caratteristice geometriche della propagazione di un incendio di vegetazione, usando un sistema di stereovisione multimodale equipaggiato su un drone (UAV). A partire dalle immagini stereoscopiche nel campo del visibile e dell'infrarosso, i punti 3D sono calcolati e le caratteristiche geometriche dell'incendio quali posizione, velocità di propagazione, altezza, lunghezza, larghezza, angolo d'inclinazione della fiamma e superficie sono estimate.

Il primo importante contributo di questa tesi é lo sviluppo di un sistema di stereovisione multimodale portabile da un drone. Questo dispositivo integra telecamere che lavorano nel campo visibile e infrarosso, un computer Rasperry PI, batterie elettriche, GPS e un'unità di misura inerziale. Questo sistema permette di ottenere immagini stereoscopiche multimodali e georeferziate.

Il secondo importante contributo di questa tesi é il metodo per l'estimazione delle caratteristiche geometriche di un incendio a partire da immagini stereoscopiche aeree.

Il framework per l' estimazione delle caratteristiche geometriche é stato validato su un'auto di dimensioni conosciute e i risultati ottenuti confermano la buona accuratezza del dispositivo. È stato anche testato con successo un fuoco in esterno e senza propagazione. I risultati ottenuti da un incendio di vegetazione in propagazione su un terreno con differenti inclinazioni sono presentati in dettaglio.

**Parole chiave :** Incendi boschivi, misura, caratteristiche geometriche, stereovisione, multimodale, UAV (veicolo senza pilota).

**Abstract** — This thesis presents the measurement of geometrical characteristics of spreading vegetation fires with a multimodal stereovision system carried by an Unmanned Aircraft Vehicle. From visible and infrared stereoscopic images, 3D fire points are computed and fire geometrical characteristics like position on the ground, rate of spread, height, length, width, flame tilt angle and surface are estimated.

The first important contribution of this thesis is the development of a multimodal stereovision portable drone system. This device integrates cameras working in the visible and infrared domains, a Raspberry Pi computer, electric batteries, GPS receptors and an Inertial Measurement Unit. It allows to obtain georeferenced stereoscopic multimodal images.

The second important contribution of this thesis is the method for the estimation of the fire geometrical characteristics from aerial stereoscopic images.

The geometrical characteristics estimation framework have been evaluated on a car of known dimensions and the results obtained confirm the good accuracy of the device. It was also successfully tested on an outdoor non propagating fire. The results obtained from vegetation fires propagating on terrain with slope changes are presented in detail.

**Keywords:** Wildland fire, measurement, geometrical characteristics, stereovision, multimodal, UAV (unmanned aerial vehicles).

**Résumé** — Les travaux menés dans cette thèse concernent le développement d'un dispositif de mesure des caractéristiques géométriques de feux de végétation en utilisant un système de stéréovision multimodale porté par drone. A partir d'images stéréoscopiques aériennes acquises dans les domaines du visible et de l'infrarouge, les points 3D de feux sont calculés et les caractéristiques géométriques tels que sa position au sol, vitesse de propagation, hauteur, longueur, largeur, inclinaison de flamme et surface sont estimées.

La première contribution importante de cette thèse est le développement d'un système de stéréovision multimodale portable par drone. Ce dispositif intègre des caméras travaillant dans les domaines du visible et de l'infrarouge, un ordinateur Raspberry Pi, des batteries éléctriques, des récépteurs GPS et une carte inertielle. Il permet d'obtenir des images stéréoscopiques multimodales géoréférencées.

La seconde contribution importante de cette thèse est la méthode d'estimation des caractéristiques géométriques de feux à partir d'images stéréoscopiques aériennes.

Le dispositif d'estimation de caractéristiques géométriques a été évalué en utilisant une voiture de dimensions connues et les résultats obtenus confirment la bonne précision du système. Il a été également testé avec succès sur un feu extérieur ne se propageant pas. Les résultats obtenus pour des feux de végétation se propageant sur un terrain de pente variable sont présentés en détail.

**Mots clés:** Feux de végétation, mesure, caractéristiques géométriques, stéréovision, multimodale, drone volant

**Riassuntu** — Questu travagliu presenta l'usu di un drone arricatu di a sistema di visione stereoscopica multimodale (RGB è IR) per misurà e caratteristiche geometriche di u focu è geolocate questi dati in una mappa. Le caratteristiche geometriche stimate sò a posizione, larghezza, lunghezza, altezza, profondità, angulu d'inclinazione di a fiamma, a superficie di a basa, i fili di u fronti, di l'angule lucale è di u a velocità locale trà duie righe di u frontale, a velocità media, l'inclinazione di a terra è u volumu di a forma frontali.

U drone hà una carta IMU è sensori cume GPS è compass. Tutte e immagini è l'infurmazioni correlate di l'IMU, di GPS è di bussole sò conservate localmente in un computer di bordu. Quest'informazione pò esse poi trasferita à un computer di terra per algoritmi post-elaborazione.

U posttrattamentu hè fattu in un computer cù Matlab per fà una forma di u focu è u su misure chì ponu esse integrate in una mappa GIS (o una carta DEM ottenuta da u drone).

U miglioramentu di più in questu travagliu hè di creà una piattaforma di sensori in aereo per monitorà è ottene dati di furesta in a so propagazione, aduprendu un sistema stereoscopicu multimodale (1 metru di basa), chì pò esse equipagiatu dinò cù telecamere miniaturizzate è economiche, per monitorà e spazii relativamente grandi (stu tipu di drone ùn esiste micca ancora nant'à u mercatu), è l'altra questione impurtante è di misurà e caratteristiche geometriche di u focu cù l'immagini scattate da sta piattaforma. U scopu era ancu di sviluppà un carregu di prototipu per u drone chì cattura immagini di quattru telecamere per creà un sistema di acquisizione cun una banda di 400 à 900 nm (nano-metru), cù tutti immagini sincronizate nantu à un sistema di posizionamentu per posizzioni u focu cù una precisione di 1 metru in una mappa cù un sistema KML (Keyhole Markup Language), per esempio Google Map.

Hè pussibule di migliorà stu prughjettu implementendu un modem per mandà un flussu di dati in una stazione di centrali in terra, in modo da ch'elli operiscenu in una modalità in temps reale, chì puderebbe aduprà cume parametri di input di un mudellu di predicazione di cumportamentu di u focu.

Stu dispositiu hè statu valutatu in duie fasi: prima, l'acquisizione di una vittura di dimensioni cunnoscute è cun marcu specificu hè stata fatta per avè una confermazione di a precisione di l'acquistamentu, dopu à un test di focu staticu (senza propagazione) hè stata fatta per confermà tuttu u prucessu automatica. Infine, i risultati riguardanti un focu di propagazione cù una inclinazione differente di a terra sò presentati.

**Parole chiave:** Misurazione di incendii forestali, veiculi aerei senza equipaggiu (UAVs), sistema stereoscopicu, trattamentu di l'immagine infraredu è visibile, robotica.

I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be.

---

Lord William Thomson *Lecture on "Electrical Units of Measurement" (3 May 1883), published in Popular Lectures Vol. I, p. 73*

# Contents

# List of Figures

# List of Tables

# Introduction

Forest fires represents a major risk to many countries around the world [1]. They cause environmental damages (deforestation, desertification, air pollution $CO_2$ emissions, heavy metals and radionuclide recirculation) and human and financial losses (destruction of housing and other buildings, destruction of equipment of struggle). Globally, more than 350 million hectares (ha) of vegetation are estimated to be affected by fires each year [2].

In Canada 9% of the total amount of fire in the world destroy 2.5 million of hectares each year with a cost about $500 million to $1 billion of dollars annually [3]. In Europe, up to 1 thousand ha of vegetation are burned annually [4] and in the Mediterranean area between 700 000 and 1 million ha of vegetation are destroyed each year [2], numerous studies have been conducted to identify the conditions that lead to these "dangerous" phenomena [5].

France is extremely impacted by this phenomenon and particularly the region of Corsica: large fires occur and in recent years an average area of 2 625 ha is burned annually [6].

The events of 2018 and 2019 provide an example of the extend of this scourge. During the year 2018, in the United States, fires destroyed 1.8 million ha of vegetation and 25% of them was in California during November. The fire of the 8$^{th}$ November was the biggest fire since 2003 [7, 8, 9, 10, 11]. California is often hit by fires and from long since some procedures have been implemented [12], but in 2018, 6 841 fires caused more than one hundred deaths, about 22 000 structures were destroyed and 62 053 hectares were burned. The year 2018 was also a dramatic year for Europe and shown that the fires now affect also the northern regions. Sweden had 5 times the usual number of fires and 25 000 hectares of forest destroyed [13, 14, 15]. The largest forest fire ever recorded in the modern history occurred in Norway destroying 3 000 ha of vegetation [16]. In August in Portugal, 1 150 firefighters battled the blaze over one thousand hectares of vegetation, 25 people were injured and other one hundred were evacuated from villages [17]. Greece had the worst fire disaster from a decade causing more than 74 deaths and more than 300 homes and shops were destroyed or seriously damaged [18]. In January, a fire destroyed in Corsica 2 000 ha of vegetation, it caused injuries in at least three people, caused the destruction of dwellings within a village and a sheepfold, leaving no chance for the herd of goats found calcined. In 2019, a record-setting heatwave occurred in parts of Europe. Record temperatures for the month of June were reached in Germany, the Czech Republic and Poland. France set an all-time record on Friday 28$^{th}$ of 45.9 °C in the town of Gallargues-le-Montueux near the Mediterranean Sea [19]. In the Gard department of southern France, fires fueled by wind and dry brush destroyed 11 houses and scorched 620 hectares. In Corsica, fires appeared also during the winter and in February 33 fire departures were recorded in two days and 1 500 ha of vegetation were burned mobilizing more than 150 firefighters and 3 Canadair planes [20].

With global warming, there will be an extension of wildfire-sensitive areas [21, 22, 23] and

these numbers will increase if efficient fire-fighting and land-use tools are not developed and used.

In order to fight effectively a wildfire, it is important to anticipate its behavior in particular the positioning of firefighters and the quantity of water to drop in such a way that people are in a safe place and their actions reduce the fire propagation.

For more than twenty years, researchers of the University of Corsica have been working on the problem of forest fires to understand the phenomena involved during their spread, to model them and to be able to predict and integrate them into simulators.

A behavior model has been developed by members of the "Fire" project of the joint unit 6134 Sciences For Environment CNRS - University of Corsica [24, 25, 26, 27, 28].

There are other models such as those presented in the publications [29, 30, 31, 32, 33, 34, 35, 36, 37].

The experimental studies of the fire spread across vegetable fuels are of great interest for understanding and modelling fire behaviour. Parameters such as the fire front geometrical properties are of particular interest during a spreading fire experiment, because they influence the propagation and the heat transfer of the fire [38, 39]. However, their estimation is difficult given the distances travelled, the dangerousness of the phenomenon and the impossibility to predict the trajectory of the fire for outdoor experiments with wind.

In the last ten years, frameworks using visual and infrared cameras have been developed in order to be used as complementary metrological instruments in fire spread experiments. The first frameworks were proposed for fires at laboratory scale on different kind of fires [40, 41, 42, 43, 34, 44, 45, 46, 47]. However, these methods can't be used for outdoor fires mainly due to their dimensions, the range of distances travelled, the presence of smoke and the uncontrolled background. More recently, frameworks were developed in order to measure fire geometrical characteristics during outdoor experiments [48, 49, 50, 51, 52, 53, 54].

Measuring fires in outdoor scenarios is a challenging task and it is important that the designed framework and associated systems are easily deployable on an unstructured environment which is unknown for each new experiment of fire propagation. For this reason it will be only presented works about outdoor fires tests, as this thesis is focused mostly on large fires. This thesis presents the work that was conducted for the measurement of wildland fires by drone.

This report is organized into 8 chapters whose details are as follow:

The first chapter is dedicated to the presentation of the state of the art of the measurement by vision of wildfires. It describes, in a first part, the techniques using ground vision systems. In a second part, methods employing hybrid frameworks composed by ground and aerial vision

devices are presented. A third part is dedicated to the use of Unmanned Aircraft Vehicles (UAV) for the measurement of some fire geometrical characteristics. It concludes with the statement of scientific difficulties to be solved in order to achieve the objective of this thesis.

The second chapter describes the analysis that was conducted to determine the most appropriate vision equipment and the method for the measurement of the geometric characteristics of the vegetation fires by drone.

The third chapter focuses on the theoretical principles of stereovision, in particular this chapter presents the geometrical model of a camera and of a stereovision system, the intrinsic and estrinsec parameter estimation, and the reconstruction of 3D points of matching points.

The fourth chapter explains the procedure used for fire pixel from visible and long wave infrared images, and the transformations applied to the 3D fire points in order to calculate the fire geometrical characteristics.

The fifth chapter describes two experiences carried out in order to evaluate the measure uncertainly of the proposed framework. One was done using a car as target, the other was carried out with a pseudo-static fire

The sixth chapter is dedicated to the presentation of the results of a fire propagation of 10 m long on a slope with variable inclination. The temporal and spatial evolution of the geometric characteristics of a fire during its propagation are presented.

The seventh chapter presents the conclusions and improvements applicable to this thesis work.

The eighth chapter presents a summary of this document in French language.

Chapter 1

# State of the art of vision for wildland fire measurement

Contents



## 1.1 Introduction

During these last decades, around the world there has been an increased activity in the research community to measure wildfires in order to better understand, model and predict them. Help by the development of electronics, computer science, and digital camera technologies, several methods has been developed, based on computer vision in order to measure the geometrical characteristics of a wildfire in propagation.

Three families of vision frameworks for detection and measurement of wildfires has been developed: systems composed only by cameras positioned on the ground, hybrid system composed by ground systems and aerial platforms, and frameworks composed only by manned or unmanned aerial systems (UAS).

### 1.1.1 Fire geometrical characteristics

A fire in propagation is described by several geometrical data such as the height, the length, the flame inclination angle, the position (front line, back line), the width and the depth of the fire front. These characteristics are presented in Fig. 1.1.

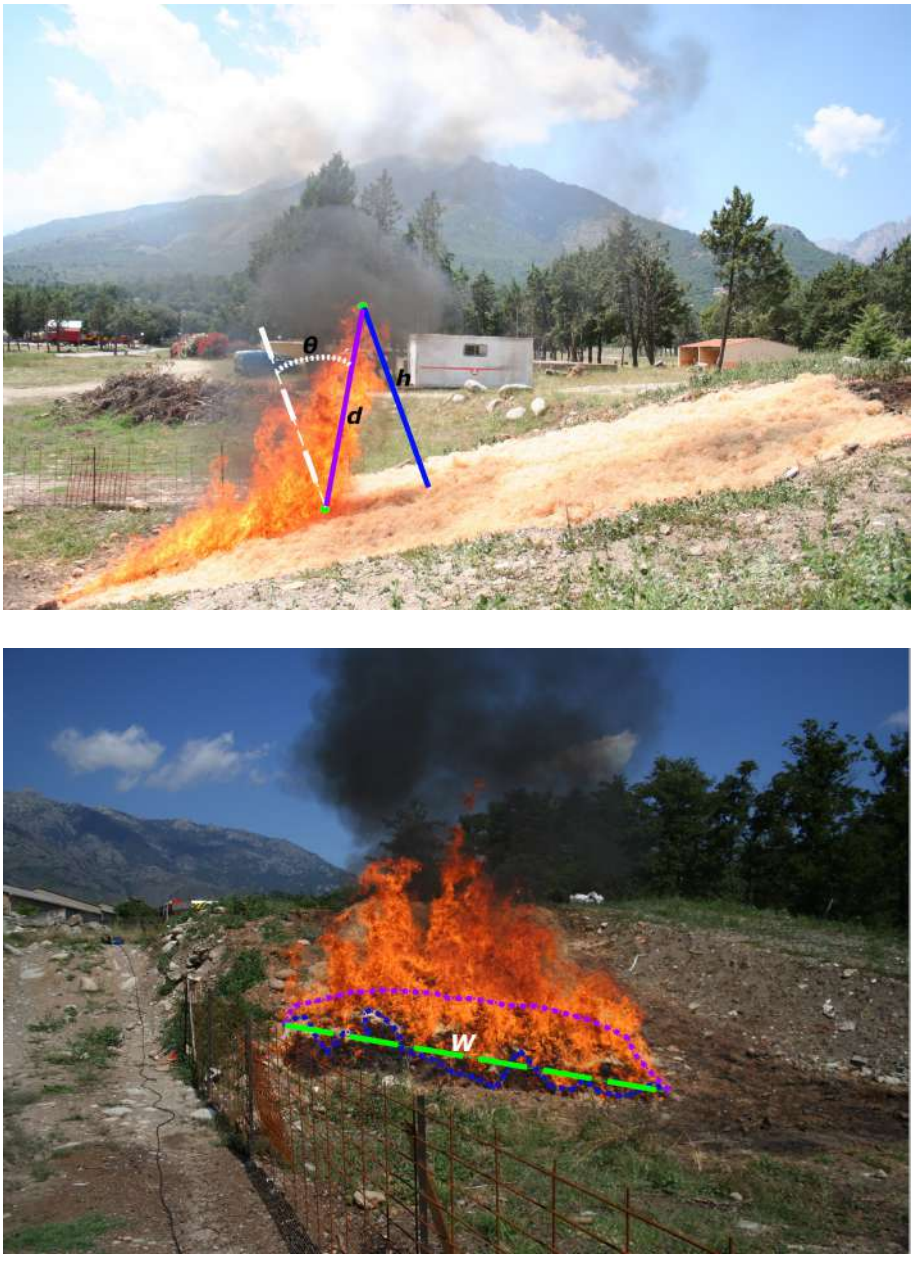



Figure 1.1: Principal fire geometrical characteristics: flame inclination angle ($\Theta$), flame length ($d$), flame height ($h$), fire front width ($w$), fire front line (*dotted blue*), fire back line (*dotted purple*).

Flame height corresponds to the length of the segment between the top of the flame and the orthogonal projection of this point on the top of the fuel layer [55]. Flame length can be defined in several ways [56]. This thesis work use the one that defines it as the distance between the top of the flame and the most advanced point of the fire front line. The front line and the back line are composed respectively of the most forward and backward ground points of the fire. The union of these lines delimits the contour of the fire base from which its perimeter and the area of fuel which is burning are computed. Width is the dimension of the fire base computed with the two horizontal extreme points and depth is the dimension of the fire base computed with the two vertical extreme points. Flame inclination angle is the angle between the length segment and the normal of the ground.

There are also others important fire parameters that are not represented in Fig. 1.1.

Rate of spread (ROS) defines the propagation velocity of the fire and allows to compute the distance traveled by fire for a specific period of time. It expresses the behaviour of fire regarding external condition as wind, fuel and topography of the terrain. 3D information are as complete as can be expected. Thus, for example, the ROS of all the points of a fire front reflects the deformation of the front lines.

In addition to the geometrical characteristics specific to the front of the fire, the researchers working on the modeling of the behavior of fire are interested by: the speed of each point of the front associated with its local angle, the position, the flame surface and volume.

The energy transferred from a fire to a target depends on the distance between the target and the fire and the orientation of the fire surface in relation with the position of the target [57]. Obtaining the three-dimensional shape of a fire and having at each instant the distance fire-target and the surface of the flame that is oriented toward the target allows to estimate this energy.

Fire surface and volume because they intervene into the behavior of a spreading fire [58].

## 1.2 Fire measurements with ground vision systems

In this section, it is presented frameworks using images obtained from several ground points of view to measure fire geometric characteristics. 3D information are computed from visible and infrared images and geometrical characteristics are estimated. Sensor fusion techniques involving telemetry sensors and GPS are applied.

### 1.2.1 Computer vision techniques for forest fires using simple cameras

Martinez-de Dios *et al.* [49] developed a camera system composed by two cameras working in the visible spectrum positioned frontally and laterally in relation to the assumed direction of fire propagation and a medium infrared (3-5 $\mu$m) camera situated in front of the fire as shown in Fig. 1.2.

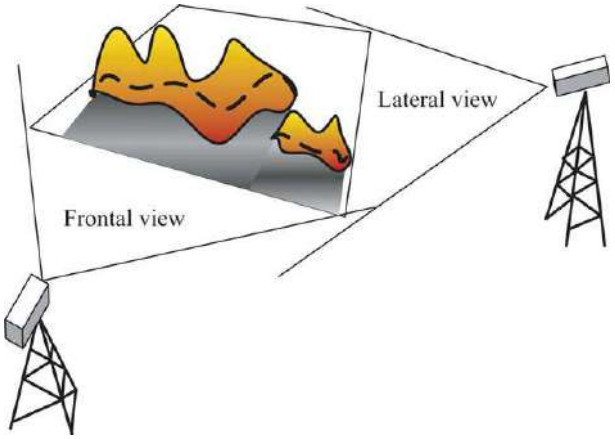



Figure 1.2: Cameras displacement [49].

Visual cameras are used to estimate measure such as flame length, flame height, flame inclination angle, and fire front line; although they present many problems in case of smoke presence that occludes fire areas. With infrared camera the smoke is transparent [59], and is a good solution for the estimation of the parameters of the base: position, ROS, and fire width.

The methodology to measure characteristics like fire front location, fire height, and flame inclination angle proposed by the authors is composed of several steps. The first operation is the segmentation of fire pixels in the images. The second step is to take measure in pixel over the segmented image. The third step is the geo localization on the terrain, and finally, in the fourth step measure like fire front location, flame height, flame inclination angle and fire width in meters are computed.

This methodology obtains a perception model of fire and not a real 3D model, but a study of the techniques used were useful to compute some of the parameters as presented in Sec. 2.1.

The segmentation algorithms chosen by the authors are different for visual and IR images. For visual images two algorithms are tested, one based on a iterative threshold on red component [60], and a second is a training-based described in [61]. This second method uses a look-up table for the RGB colour space. For each triple (R,G,B) the look-up table is built creating a Boolean mask of the same size of the image where for each position of the mask the TRUE value corresponding to the presence of fire in the corresponding pixel of the

image, and FALSE value otherwise. The learning phase builds the look-up table from several pairs of image, where each pair is formed by the RGB image andthe Boolean mask. Thus an RGB histogram is created accumulating Gaussian distribution when the triple correspond to a fire pixel and substrating Gaussian distribution otherwise, finally the RGB histogram is thresholded. The training algorithm proved a more robustness and it is the algorithm chosen by [49] for the final framework.

The IR camera can see the intensity of the heat radiation in a scene and it is then possible to discriminate a fire from the scene background analyzing the intensity values. Knowing that in medium infrared band the radiation intensity of the flame is lower than the radiation intensity of the base, it is possible to use this kind of cameras to differentiate flame and fire base. In [62, 63] a fuzzy multi resolution algorithm is used to segment fires into IR images, this method finds a threshold intensity with a coarse-to-fine search analysis.

The results of this process are binaries images where 1 represent fire pixels and 0 background pixel. When the fire is segmented into the IR and visible images, a contour detection function is applied to find the edges of the segmented regions.

To identify the fire base, given the supposition that the base fire pixels change slower than the fire pixels of the flame, it is considered a set of $n$ consecutive frames taken at 1 frame per second and an image containing the pixels that have no changed over these frames is obtained by using the binary AND operator $\otimes$:

$$\overline{I}_i = I_{i-n} \otimes I_{i-n+1} \otimes \bullet\bullet\bullet \otimes I_i. \tag{1.1}$$

where $I_i$ is the binary image at the frame $i$. The edges of the fire region in $\overline{I}_i$ represents the fire contour of the base that includes pixels of backward fire line and pixels of forward fire line:

$$C_i^b \cup C_i^f = \{\mathbf{p}_{ik}^b, k = 1, ..., N_i\} \cup \{\mathbf{p}_{ik}^f, k = 1, ..., N_i\} \tag{1.2}$$

being $C_i^b$ the set of the pixels for the backward base contour and $C_i^f$ the set of the pixels for the forward base contour.

To differentiate between pixels of the backward and pixels of the forward the algorithm considers the velocity vector $\overline{v}_i^b$ of the centroid of $C_i^f \cup C_i^b$, and to get a robust estimation of $\overline{v}_i^b$ it considers an average of 3 frames. Tracing the normal line of the vector $\overline{v}_i^b$ passing through the centroind of $C_i^f \cup C_i^b$, it is possible to discriminate the points of the backward line and the points of the forward line. For each point of the backward line, a line with the same direction of $\overline{v}_i^b$ identifies an intersection point on the line of the forward points. Thus, for each point of the backward line, it is possible to compute the euclidean distance with its correspondence point on the forward line, and the median of these distances gives a depth value in pixel coordinates, as shown in Fig. 1.3.

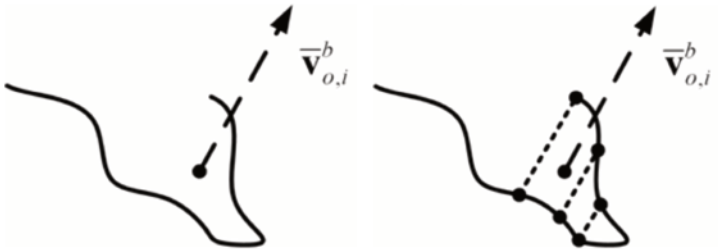



Figure 1.3: Measured fire depth. ($a$) fire base contour and direction vector of its centroid. ($b$) pixels of backward and forward contour to compute the fire depth [49].

The method to estimate the flame inclination angle needs first to identify the set of edge pixels of the flame $C^l_i$, thus an approximation of the flame can be expressed by the vector

$$\overline{v}^l_i = c^l_i - c^b_i \tag{1.3}$$

with $c^l_i$ the centroid of $C^l_i$ and $c^b_i$ the centroid of $C^b_i$.

Once computed the vector $\overline{v}^l_i$, the flame inclination angle and the flame length can be easily estimated, Fig. 1.4.

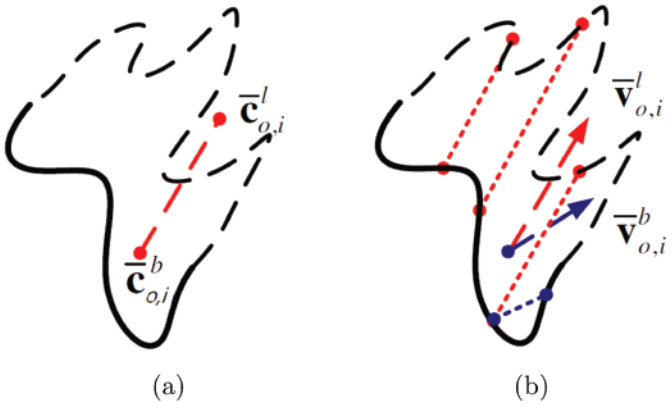



Figure 1.4: Flame inclination and flame lenght estimation. ($a$) backward base contour (solid), top of the flames (dashed) and their centroids. ($b$) correspondence among pixels of the backward base contour and pixels of the top of the flames. [49].

The intensity, or vector length, of $\overline{v}^l_i$ is the estimation of the flame length, while the angle between the vectors $\overline{v}^b_i$ and $\overline{v}^l_i$ corresponds to the flame inclination angle. These parameters are depicted in Fig. 1.5.

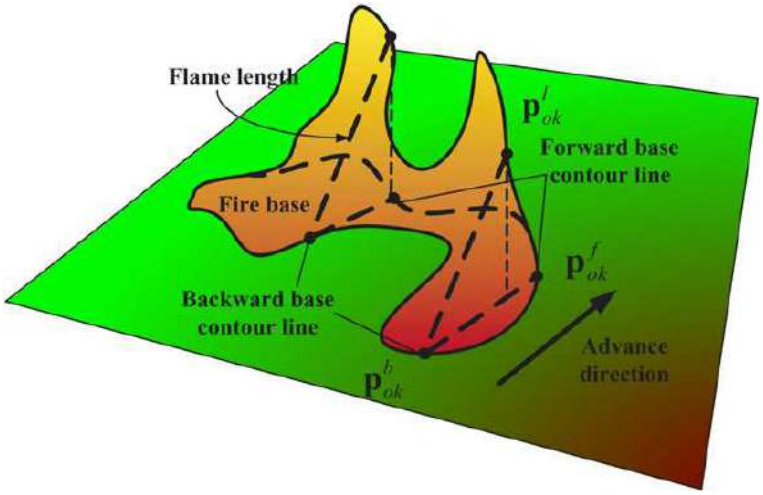



Figure 1.5: Scheme of the parameters used by Martinez *et al.* [49].

The geolocalization of the information is made using data from GPS and IMU sensors. This operation is necessary when several images are acquired, because in this step all the measures for each fire can be merged to create an evolution.

Two field configurations are considered: one with a planar field and the other with a full 3D field.

If fire is situated on planar field, it is possible to approximate the terrain to a plane $Z = 0$; in this way the pixels coordinate $\mathbf{p}$ and the world coordinate $\mathbf{P}$ are related by:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{H}_j \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{1.4}$$

where $s$ is the scale factor, $\mathbf{p} = [u\ v\ 1]^T$ are the pixel homogeneous coordinates, $\mathbf{P} = [X\ Y\ Z\ 1]^T$ are the world homogeneous coordinates, and $\mathbf{H}_j$ is the homography matrix of the image $j$ computed as

$$\mathbf{H}_j = \mathbf{A}_j \begin{bmatrix} \mathbf{r}_{1,j} & \mathbf{r}_{2,j} & \mathbf{t}_j \end{bmatrix} \tag{1.5}$$

with $\mathbf{A}_j$ is the intrinsic matrix composed by the physical parameters of the lens: focal length, distortion factor, and coordinates of the central pixel in the image. $[\mathbf{r}_1,\mathbf{r}_{2,j},\mathbf{t}_j]$ contains the rotation and translation vectors to relate the world coordinate system to the camera system, as shown in Fig. 1.6.

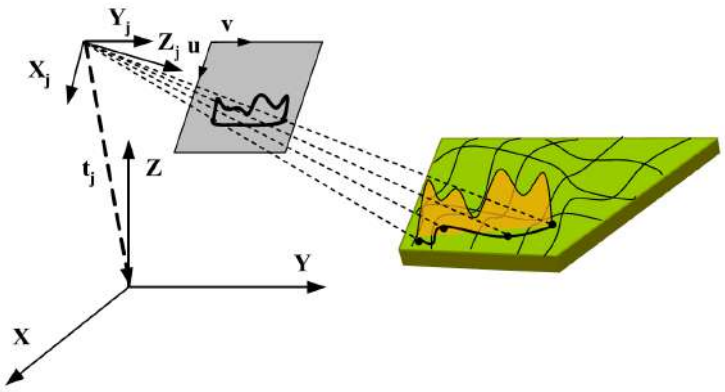



Figure 1.6: Relation between real world and camera reference [49].

In planar propagation tests only four correspondences between the pixels coordinates and theirs world coordinates are needed to calibrate the cameras. The homography matrix $\mathbf{H}$ is invertible, so it is possible to compute any world coordinates given the pixel coordinates in the image. However this kind of field configuration is not very interesting to study because is a condition too limiting.

If the terrain is a full 3D field more correspondences between pixels on the image plane and points on the 3D world coordinates are needed to calibrate the camera.

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A}_j \begin{bmatrix} \mathbf{r}_{1,j} & \mathbf{r}_{2,j} & \mathbf{r}_{3,j} & \mathbf{t}_j \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \tag{1.6}$$

The translation $\mathbf{t}$ is computed using the position of the camera read from GPS, and for the rotation an IMU board that gives roll, pitch and yaw measures is used. To compute the homography matrix firstly the values of roll, pitch and focal length are set to initial values, and the initial matrix $\mathbf{H}_j^0$ is computed. This matrix is used to compute the 3D coordinates of some landmarks positioned over the terrain. If the real GPS position of the landmark is known, it is possible to use the inverse matrix $\mathbf{H}_j^{0^{-1}}$ to project the landmark position on the image plane. The error between the projection of the landmark position on the image plane and the real pixel coordinate on the image is computed, and a non-linear least square interior-reflective Newton method is applied to minimize this error to obtain a new homography matrix. This method will be iterate for $k$ step and the last matrix $\mathbf{H}_j^k$ at step $k$ is considered [64, 65].

Once the cameras are calibrated it is possible to obtain measures of fires in geographical coordinates. Each camera computes a measure $m$ with a confidence index $\omega$ related to the camera resolution (pixel/meter), and an heuristic information related to the camera $j$. The

final measure $M_i$ is a combination of all the single measures:

$$M_i = \frac{\sum_j (m_{ij}\omega_{ij})}{\sum_j \omega_{ij}} \tag{1.7}$$

$$\omega_{ij} = F_j\left(1/r_{ij}^2\right) \tag{1.8}$$

where $i$ is the frame index and $j$ is the camera index. $F$ will be set with a value in [0,1] where high values indicates more measure confidence, for example if the camera IR is in front of fire it will have an high value of F and visual camera in the same position a small F value (because in front of fire there is too smog that occludes flame in the visible band). However, not all the cameras can estimate with accuracy all the parameters because some cameras could not see some parts of the flame, or as with the medium infrared bands some parameters are not accurate as the top of the flame that is not clear because it is merged to the flame heat radiation, and infrared camera can't distinguish with single fire pixel inside the fire countour shape.

The approach proposed in [49] uses a sensor fusion technique to represent some parameters visible in one camera into a reference system of another cameras (Fig. 1.7). Given two cameras $i$ and $j$, a point $\mathbf{p}^i$ in camera $i$ can be projected on the image plane of the camera $j$ with the following relation:

$$\mathbf{p}^j = \mathbf{H}_j \mathbf{H}_i^{-1} \mathbf{p}^i; \tag{1.9}$$

this relation is a combination of the Eq. 1.6 applied to the two cameras.

Fig. 1.7 shows this operation: with frontal infrared view it is not possible to detect with accuracy the position of the point **D** that corresponds to the top of the flame, but in lateral visual view this point is clearly visible, so the transformation allows to bring the point **D** from the lateral camera reference to the frontal camera reference.

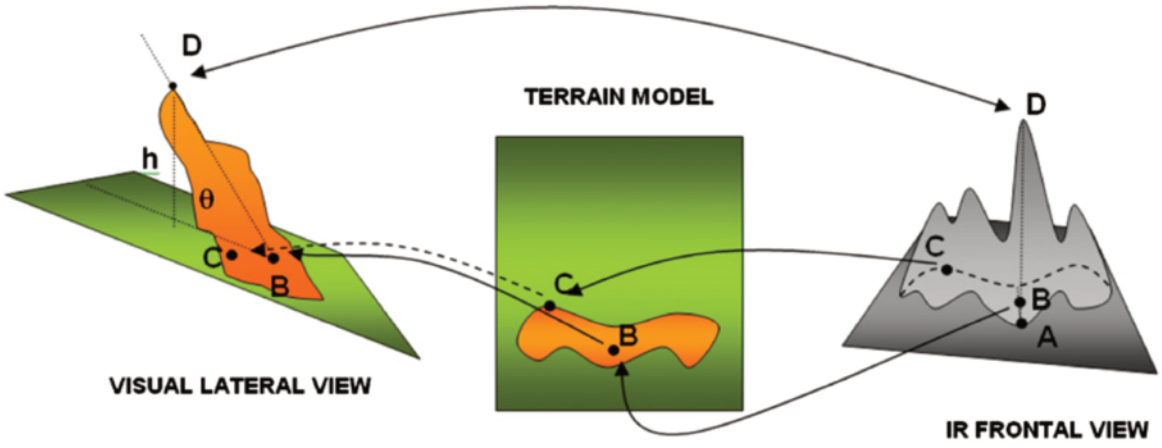



Figure 1.7: Example of transformation function between lateral and frontal view [49].

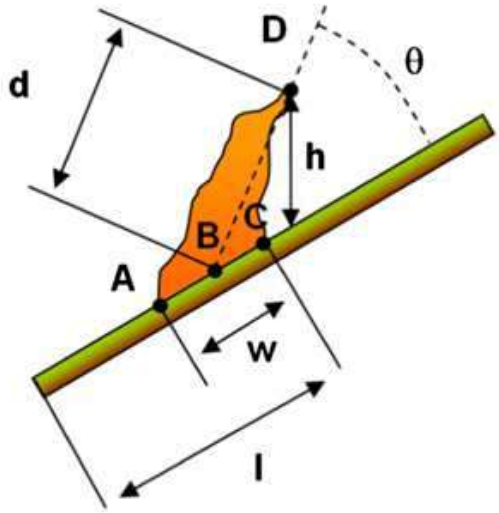



Figure 1.8: Fire geometrical characteristic computed by Martinez *et al.* [49].

Point **B** corresponds to the projection of point **D** on the fire base, and a transformation to bring the point **B** to the lateral camera view allows to compute the vector of the flame. A similar process is applied to the point **C**. Using this process to all the key points **A**,**B**,**C**,**D** the flame height $h$, flame length $d$, flame tilt angle $\Theta$, depth $w$, and distance of the most advanced point $l$ can be estimate. Doing this process for all the images, the evolution of the parameters is obtained, in Fig. 1.9 and Fig. 1.10 are presented the fire front lines, the most advanced points, and the fire width.

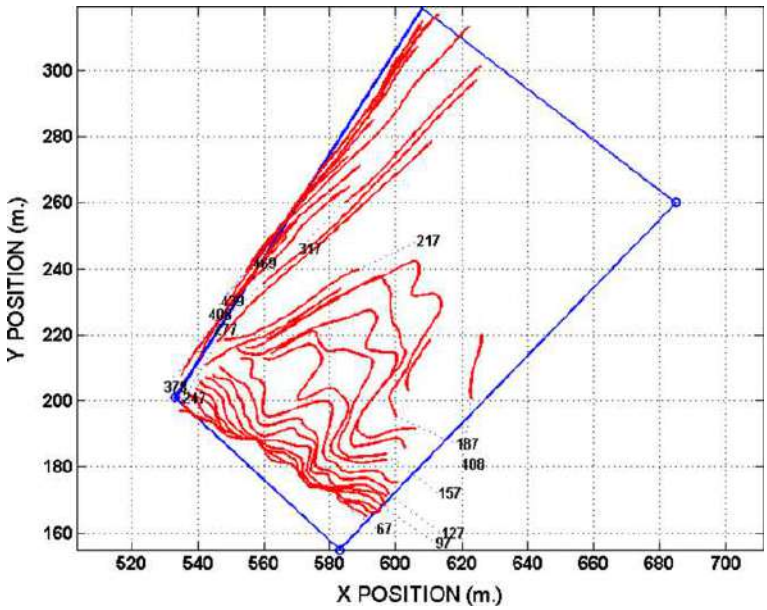
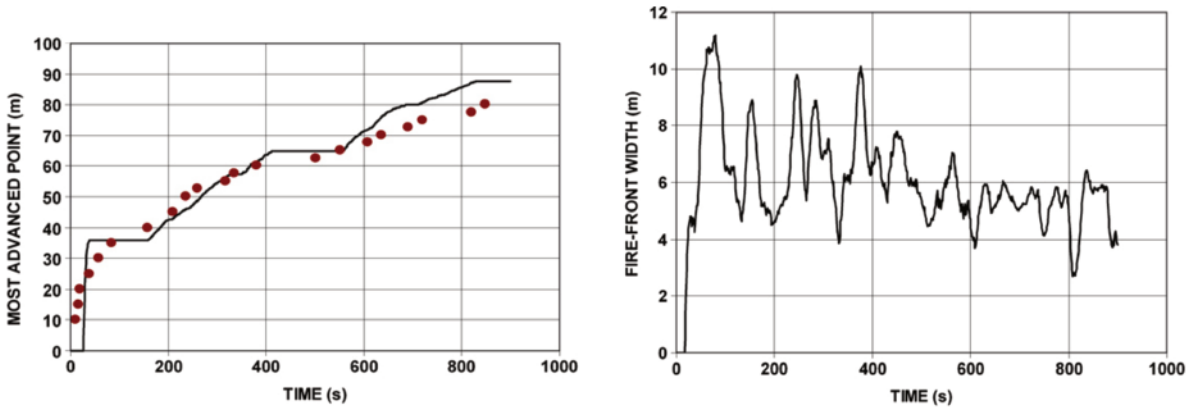



Figure 1.9: Fire front line evolution in geographical coordinates [49].




(a) (b)

Figure 1.10: Most advanced point and width evolution: (*a*) most advanced point. (*b*) fire width [49].

The measures obtained in this way can create an approximation of fire characteristics, but with this methodology there is no information about the depth for each fire points and it is not possible to have all the geometric characteristics needed by researchers working on propagation modeling and fire behaviour modelling. In addition the fixed position of the cameras makes this method unusable for fields not easy to access, so it is difficult to position the cameras at the place that would be the best and then if the fire does not follow the direction expected there is a problem. It is necessary to predict the direction of fire spread to put the cameras in the best position (front, behind or lateral of the fire). This is difficult because the direction of fire depends on wind that can be changing. Moreover, this method requires placing beacons or to have landmarks on the field to register the images from the different vision systems and this can be a difficulty. Finally, it is difficult to evaluate the efficiency of this method for the measurement of geometrical characteristics of experimental fires because the article [49] doesn't mention the distance between the cameras and the fire and the results precision.

Given a 3D point $P$, a camera is a device which transforms any point $P$ of 3D space into a point in 2D space of the image. This transformation therefore removes the third dimension and is, therefore, irreversible. Figure 1.11 show two 3D points that are projected onto the image plane at the same 2D point, because they are on the same projective line. This means that given a 2D image point, there are infinitely many 3D points that can be the projection.

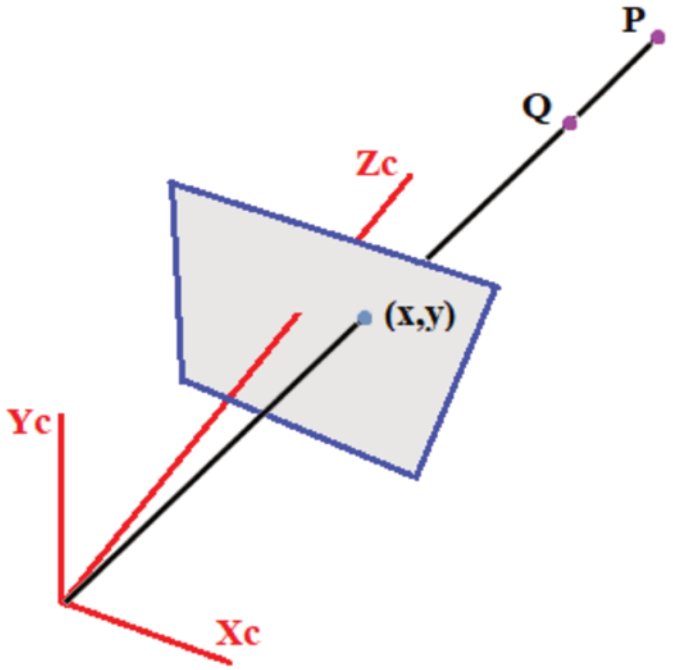



Figure 1.11: $P$ and $Q$ are projected in the same 2D coordinates ($u$, $v$).

Using two camera, it is possible to determine the three-dimensional position of the point by triangulation. If for each camera is selected a single 2D projection point, for example ($x_1$, $y_1$) for the first camera and ($x_2$, $y_2$) for the second camera, then exist a single 3D point in the space which is projected in the first camera in ($x_1$, $y_1$) and in the second camera in ($x_2$, $y_2$) (Fig. 1.12).

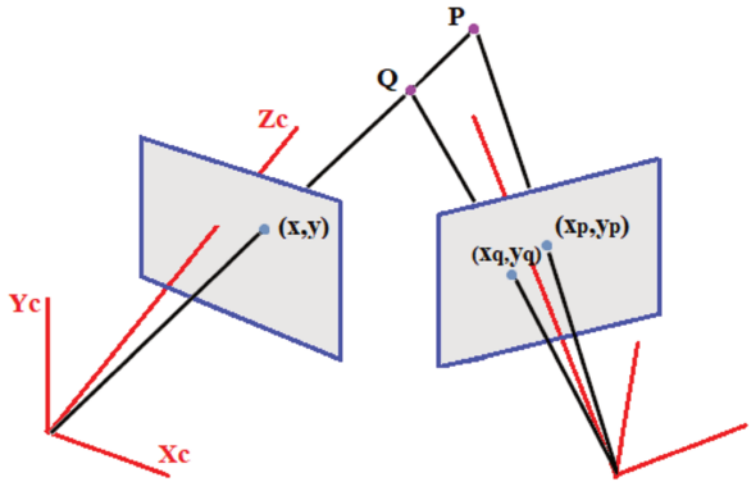



Figure 1.12: $P$ and $Q$ are projected in the same 2D coordinates for the first camera but in two different 2D coordinates in the second camera.

### 1.2.2 Computer vision techniques for outdoor experimental fires using stereovision

The use of stereoscopic vision techniques makes possible to estimate depth information of a target with two or one moving camera. The principles of the stereovision will be discussed in details in the Chap. 3.

**Methodology to estimate fire geometrical characteristics using stereovision system**

A first solution was proposed by Rossi *et al.* [53, 66] for small outdoor fires. The methodology is composed by five steps. The first step is the detection in the images of fire pixels; the authors [67] uses information from YUV and RGB color spaces. In the second step, points of interest are detected on the contours of the fire zones in the images. A third step makes it possible to match the pixels of fire between the stereoscopic images. In a fourth step, the 3D points of fire are obtained by triangulation of the paired points. Finally, from the three-dimensional points, the geometric characteristics of the fire are calculated. From 3D fire points, parameters like front position, rate of spread, flame length, flame height, flame tilt angle, surface and volume are estimated, on a field with with no a priori knowledge about the slope of the ground. The detection algorithm of fire pixels proposed by the authors [67] uses a combination of the information from the RGB and YUV color spaces. Once the zones of fire have been detected, features are searched on the fire area contour, then the matching of these points is carried out and finally the obtaining of the point cloud of 3D coordinates by a triangulation algorithm [68].

The authors in [69] describe the estimation of fire geometrical characteristics by using two devices positioned in the rear and lateral places of the assumed fire direction. In this research, the authors didn't use GPS and IMU sensors to find position and orientation of each camera. A procedure especially created for the registration of cameras at field scale has been developed. It is based on the use of a 1 m diameter balloon put successively at several positions on the field and visible by the different stereovision systems. For each stereovision system, the 3D position of the centre of gravity of the feature put at different places is compared with a theoretical network of positions generated by computer and put at a given place in space. The place which minimizes the distance between the 3D positions of the feature and the theoretical network of positions gives the real pose (translation and rotation) of the features in the camera frame

An example of 3D points is shown in Fig. 1.13; from this points all the characteristics previously discussed can be computed.

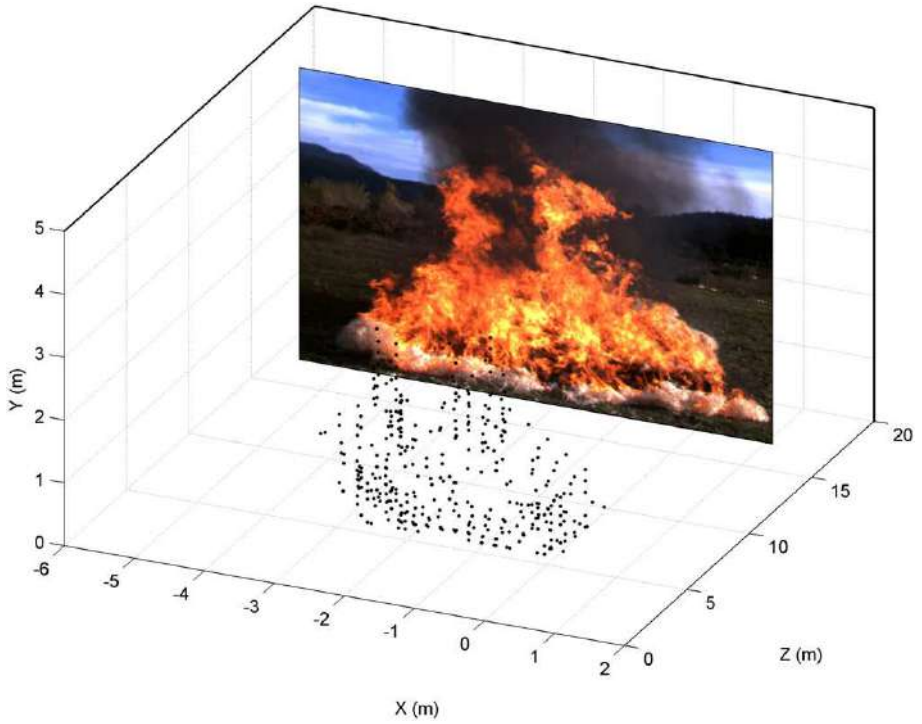



Figure 1.13: 3D point cloud of Fire [53].

The method needs at each instant to know the equation of the plane where the fire is in order to compute geometrical characteristics such as height and length without including in the estimated values the elevation difference given by the topology of the ground. To compute this plan, the lower 3D points of successive fire fronts (corresponding to an advance of 3 m) which are at ground level are considered. From them, a plane equation is computed using a least square techniques [64, 65] (Fig. 1.14).

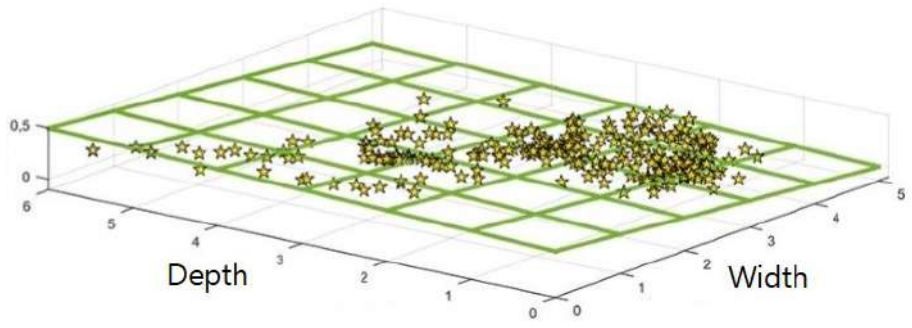



Figure 1.14: Estimation of a propagation plan from the low 3D points of a fire [70].

All the 3D fire points situated on the propagation plane are transformed in such a way that the fire lower 3D points have a height equal to zero. From the 3D fire points on the

ground, the line of the most advanced points (front line) and the line of the most backward points (back line) are determined interpolating by a Bezier-spline interpolation [71] a linked list of lowest points containing the point at maximun depth and the two extreme points along the width axis (Fig. 1.15).

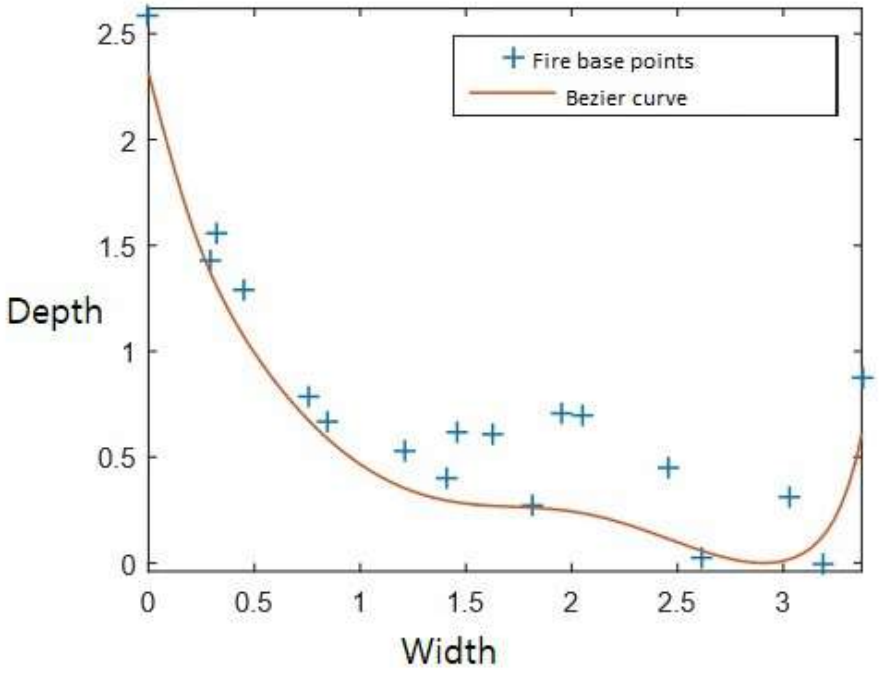



Figure 1.15: Fire line points with its related Bezier curve [70]

To compute the fire perimeter a convex hull algorithm only on the $XZ$ coordinates of the 3D points is applied [72], where $X$ and $Z$ correspond to the width and depth axes, respectively. For ROS computation several two pair-temporal front lines previously computed are taken into account. Firstly the algorithm selects for the first line a set of equidistant points and for each point the corresponding point in the second line is chosen as the intersection between the orthogonal line that passes by the first point and the second line (Fig. 1.16).

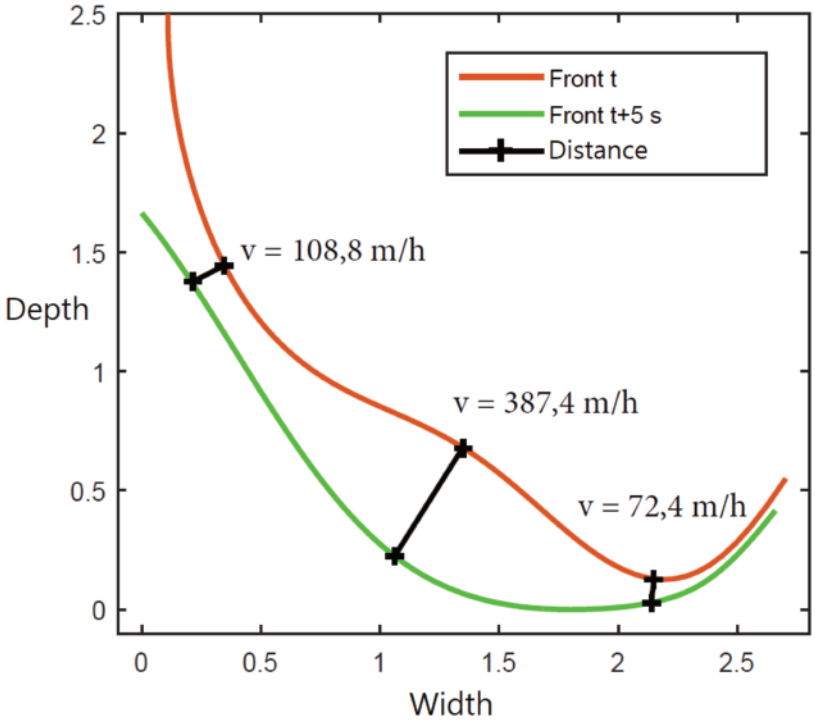



Figure 1.16: Corresponding points of two fire lines and ROS estimation of 3 points [70].

ROS is computed as the ratio between the length of the segment delimited by this two points by the lapse of time between the two image acquisitions corresponding to the two fire fronts.

The distance between the base plane and points of the upper part of fire is the fire height. Upper points of fire are interpolated by B-spline. For each point, a vector between this point and the point that has the same abscissa on the front line is computed. The length of the vector is the fire length and the angle between this vector and the normal of the base plane is the flame tilt inclination. The $Z$ coordinate of the upper point is the height (Fig. 1.17).

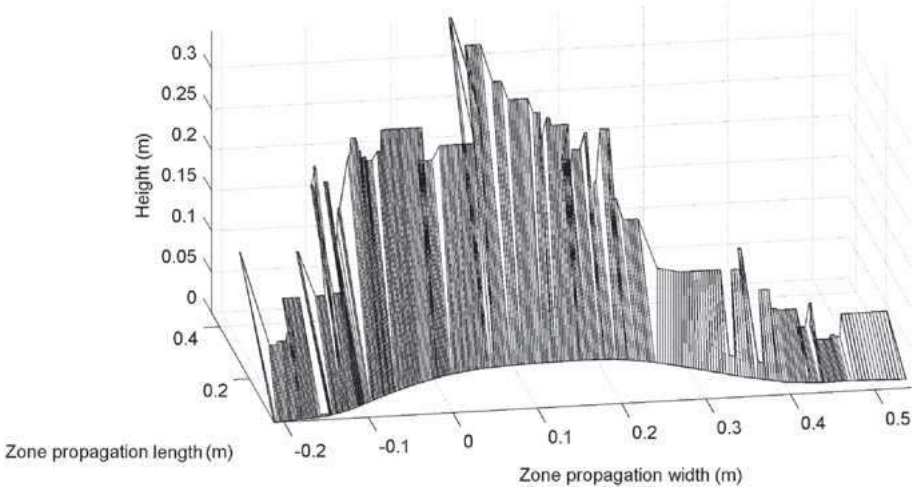



Figure 1.17: Height estimation [66].

The volume estimation is made from the convex hull shape of the 3D points described in Fig. 1.13 [53]. The framework is based on the use of a visible pre-calibrated stereo camera in order to compute 3D information without the need for artificial beacons or natural marks.

Toulouse *et al.* [54] improve the use of stereovision to estimate outdoors fire characteristics including the volume by developing a multimodal (visible and near infrared) stereovision framework. In Fig. 1.18 the general scheme is presented. The first step is a fire pixel extraction conducted in both the visible and NIR spectrum. It is followed by a feature detection procedure used in a stereo matching. From the corresponding points, a triangulation procedure permits their 3D reconstruction. An ultimate step is carried out from the 3D points in order to estimate the fire geometrical characteristics.

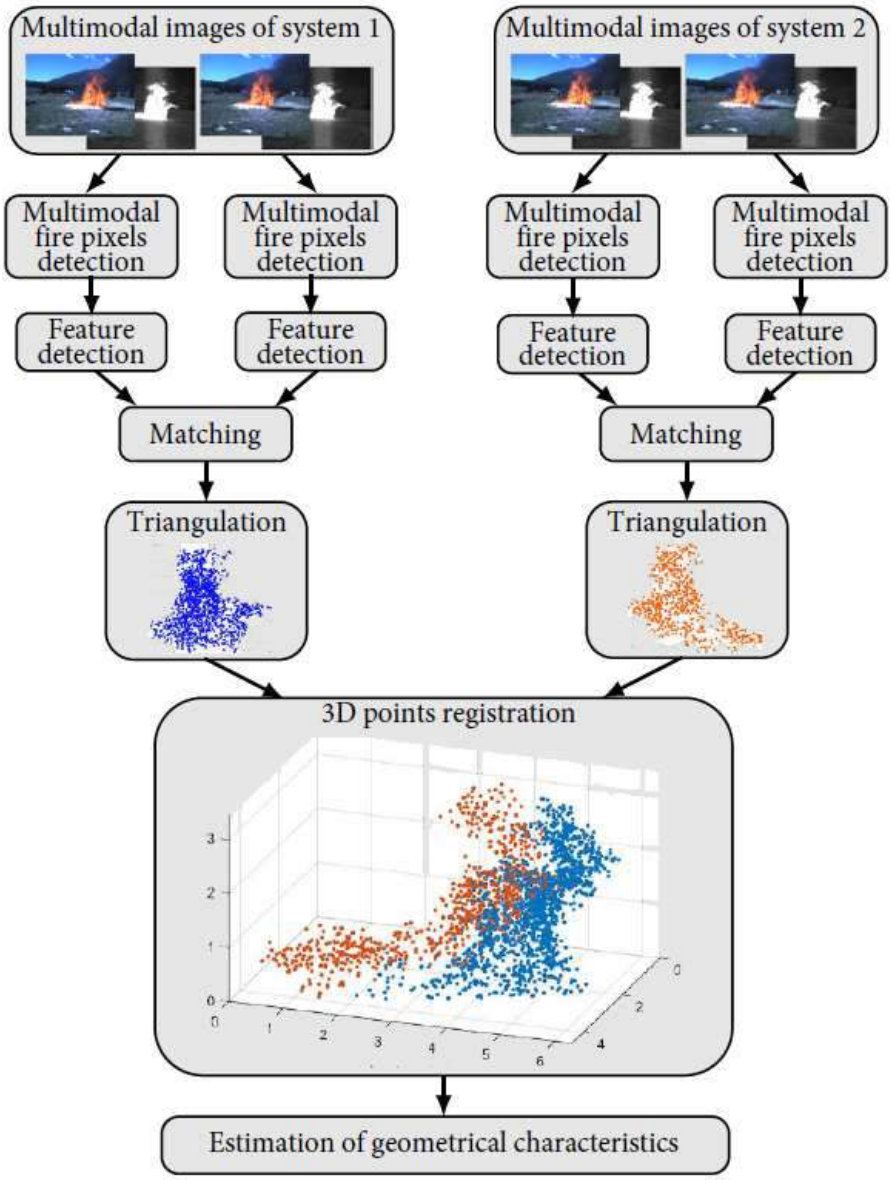



Figure 1.18: Diagram of the Toulouse *et al.* framework [54].

Two multispectral cameras JAI AD-080GE [73] that work simultaneously in near infrared band (750-900 nm) and visible band (1024x758 resolution pixel) are used. These cameras are positioned to 1 m apart on a rigid axis. A camera inter-distance (baseline) of 1 m was chosen in such a way to produce with the camera focal length a theoretical depth error lower of 30 cm at 12 m. The target considered is a fire with a supposed width of 10 m and 2 m of height. The left camera of each stereovision system was equipped with an IMU sensor to obtain the orientation of the devices and their position was obtained with a centimeter-precision GPS system. A synchronization system remotely controlled allowed to take simultaneous pictures from all the cameras; the acquisition time was fixed by the authors to 1 image per

second. Each stereovision system was calibrated using a checkboard of known dimensions positioned in several points on the terrain at about 12 m. In Fig. 1.19 it is possible to see a stereovision system composed by two cameras, the box containing the IMU sensor and the remotely activated trigger with its antenna.

Figure 1.19: System developed by Toulouse *et al.*: two cameras that work in multimodal bands (near infrared and visible), on the left there are the IMU sensors, and in the middle there is the receiver to start the cameras [54].

In [54] the authors gets a new approach to detect fire pixels. Two method are used, one is an histogram based method descripted in [61], to obtain a detection of fire pixel in visual images, and a threshold based method on the near infrared image. For each pixel a probability to be a fire pixel is computed as:

$$\mathbf{P}_x = \frac{p_1(x_1) + p_2(x_2)}{2} \tag{1.10}$$

where $x$ is the pixel in the multimodal image, $x_1$ the intensity value in the infrared image, $x_2$ the RGB triplet in the visual image. $p_1$ is the probability that the pixel $x$ correspond to a fire pixel into the infrared image:

$$p_1(x_1) = \begin{cases} \frac{x_1 - t}{2t} + \frac{1}{2} & \text{if } x_1 \leq t \\ \frac{x_1 - t}{2(255 - t)} + \frac{1}{2} & \text{otherwise} \end{cases} \tag{1.11}$$

$$t = \begin{cases} \mu + 2\sigma & \mu + 2\sigma \leq t \\ 255 & \text{otherwise} \end{cases} \tag{1.12}$$

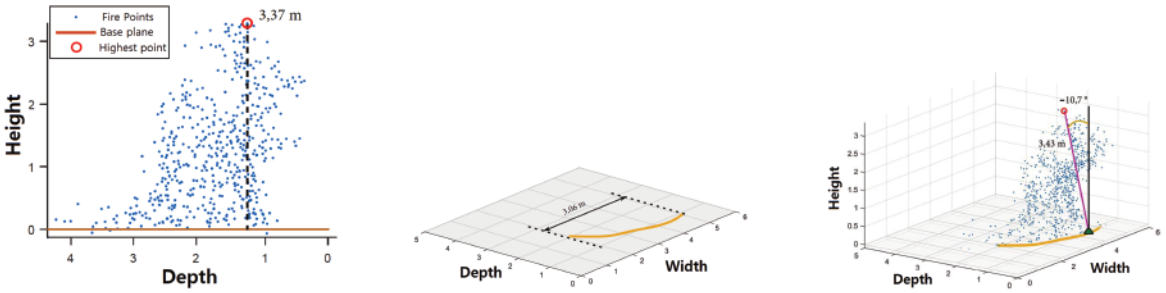
$\mu$, $\sigma$ are the intensity average, the standard deviation of the infrared image, and $t$ is the threshold used for $p_1$. $p_2$ is the probability that the pixel $x$ correspond to a fire in the visual image. The RGB triplet $x_2$ associated to the pixel $x$ which correspond $p_1(x_2) > 0.5$ are used as learning pixel. The learning pixels computed are used as input for an histogram based method and $p_2$ is computed. Finally $x$ is a fire pixel if $\mathbf{P}_x > 0.5$.

To extract points of interest an algorithm based on a multi-scale oriented patch technique

is used [74] applied on visual images. From each system a set of 3D points is computed through triangulation method.

The techniques described previously produce two sets of 3D points, relative to their local reference frame (stereovision system frame). The registration allows them to be merged into one model. There is no obligation on the positions of the two stereovision devices, but to get the complete shape of a fire the best positions are in front and behind the fire front. For each new configuration of stereovision systems, the position given by the GPS and the IMU values associated to the devices. Knowing these measures from each system a matrix transformation is computed and used to project all the points in a unique reference frame.

Once a merged point cloud is obtained, the framework computes all the geometrical characteristics of fire [54, 75] described in Sec. 1.1. The geometric characteristics of fire are expressed in relation to a basic plane which is the surface of the ground; this base plane is determined from the low 3D points of fire obtained from the rear view of the fire. The base plan is calculated from successive images (the images are taken with 1 fps). The authors used 300 images (corresponding to 5 minutes of propagation) and renewed by slice of 120 images (corresponding to 2 minutes of propagation). From each image a low pixels searching algorithm is performed divided the image in column. For each column selects the detected pixel which higher row index. To refine this selection, for example because some pixel points can be the lowest of their column but are not situated on the base plane, the average $\mu_p$ and the standard deviation $\sigma_p$ of the row indices $v_p$ are computed and the points located over a certain distance from the average are eliminated. The 3D fire points identified as being on the ground are accumulate and used to estimate the equation of the plane that better fit them. Once the plane is known a new refinement is computed eliminating the 3D fire points located more than 30 cm from the ground. The lowest points of the fire are those on the ground and therefore correspond to the base of the fire. Depending on the direction of fire spread, the base has a front line and a rear line. A fire front line and a rear line are computed using the device positioned in front and behind the fire, respectively. From two fire lines, the rate of spread computed with the method described in the methodology presented by Rossi *et al.* previously discussed. The principal direction taken by a fire during its propagation corresponds to the average axis according to which it propagates. This direction is estimated by the average direction of the segments which connect each point of the first line and its equivalent in the second line. In particular, the Bezier curves are considered two by two and on each of them one hundred equidistributed points are numbered. All the vectors connecting the points of the same number of the curve pairs are then calculated and the average of these vectors gives the main direction of propagation. The fire direction previously computed is important to estimate the height, depth, width and flame inclination angle. All the 3D points are rotated to align the Z-axis to the fire direction vector. Fire height is defined as the y-coordinates of the highest point. Fire width is the difference of X-coordinates of the two extreme points of the fire line. Fire length is the euclidean distance between the most advanced point of the front line and the highest point. Fire inclination angle is computed as the angle between the normal of the base plane and the line used for the fire length estimation. All of these estimations are shown in Fig. 1.20.




(a) (b) (c)

Figure 1.20: Fire geometrical characteristics estimation. (*a*) height. (*b*) width. (*c*) length and inclination angle [70].

A Delaunay triangulation applied to the 3D points allows to obtain the 3D shape of fire, for each 3D point the 3 nearest points are chosen to build a tetrahedron (Fig. 1.21). First, all the tetrahedrons for which the inscribed sphere is greater than 60 cm are deleted. The value of 60 cm was chosen empirically using experimental data and remains the same regardless of the distance between fire and the stereovision systems. Second, the tetrahedrons for which the projection of their centers on the segmented images do not correspond to fire pixels are eliminated. Thus the volume is given by the sum of each tetrahedrons volume.

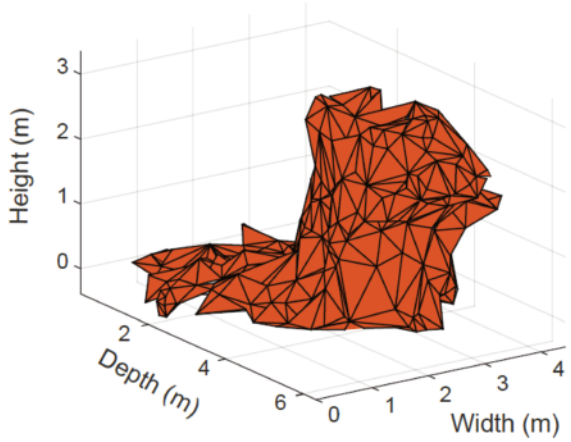



Figure 1.21: 3D reconstruction of a fire with Delaunay triangulation [54].

## 1.3 Fire measurements with hybrid device composed by ground vision systems and aerial vision systems

Martinez de Dios [51] proposed a framework for automatic fire measuring using an hybrid system composed by cameras on the ground and cameras mounted on unmanned aerial vehicles (UAV).

The obtained measures are the fire location, maximal fire height, fire width, flame tilt inclination, flame length, and rate of spread. The proposed framework is an extension of [49] previously discussed, with an improvement of the fire location.

Knowing that the change of direction of fires creates problem with fixed system, Martinez *et al.* propose to use in addition to cameras on the ground, drones equipped with cameras in order to follow the evolution of the fire and to have at every moment the best image possible. (Fig. 1.22).

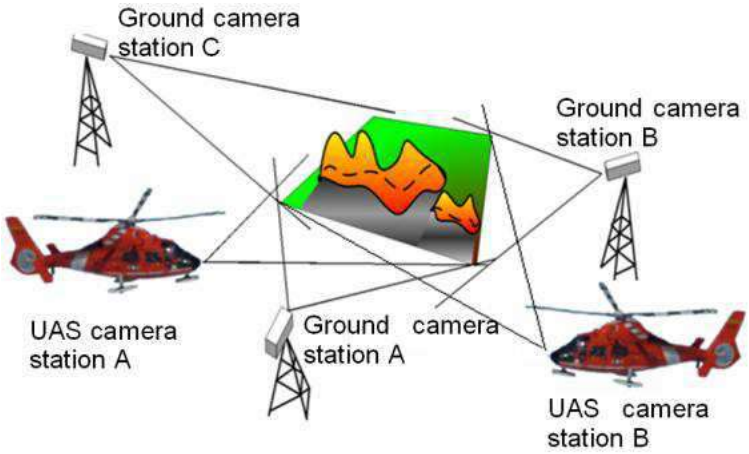



Figure 1.22: Deployment of the forest-fire measuring system proposed by Martinez *et al.* [51].

The entire deployment is composed by three different combinations of infrared-visual cameras on the ground, and two pairs infrared-visual cameras mounted on UAV. In particular the cameras used are a mid-infrared ground camera with 24° of HFOV, a mid-infrared ground camera with two lenses and HFOV of 10° and 20°, a far-infrared ground camera with three lenses and HFOV of 12°, 24°, 45°, a micro far-infrared drone camera with 24° of HFOV, and a camera on the ground and mounted on UAV with $752X582$ of pixel resolution and a focal length adjustable from 15 to 90 nm. Aerial images were used to compute the fire shape, fire position, and fire width. The proposed method estimates in a first step the fire characteristics in the same way that the one described in [49] and improves the fire location using a filter associated with a propagation model. The filter used is a Recursive Bayesian Filter associated to a one-update fire model, this model do a short-term prediction of the fire. The state of the

filter will be the input of the fire model to compute the short-term prediction. The input of the filter is the set of measurements obtained from each camera, and the output is the state of the fire front, that includes all the fire parameters. The authors assume that the fire in very short-time is similar to a linear dynamic system with noise. To reduce the noise a Kalman filter is used. In the Fig. 1.23 it is shown the fire front lines obtained from a visual camera carried by an UAV (a), from a frontal visual ground camera (b), from a frontal infrared ground camera (c), and finally the merge from all the measurements (d). In this figure the authors have highlighted an area with an high percentage of smoke, where it is possible to see that only the infrared camera is able to work in this condition.

This method could be relevant for the measurement of wildfires, but the lack of information does not allow us to judge their effectiveness, in particular the distance drone-target, the focal length of the camera and a ground truth comparison is missing.

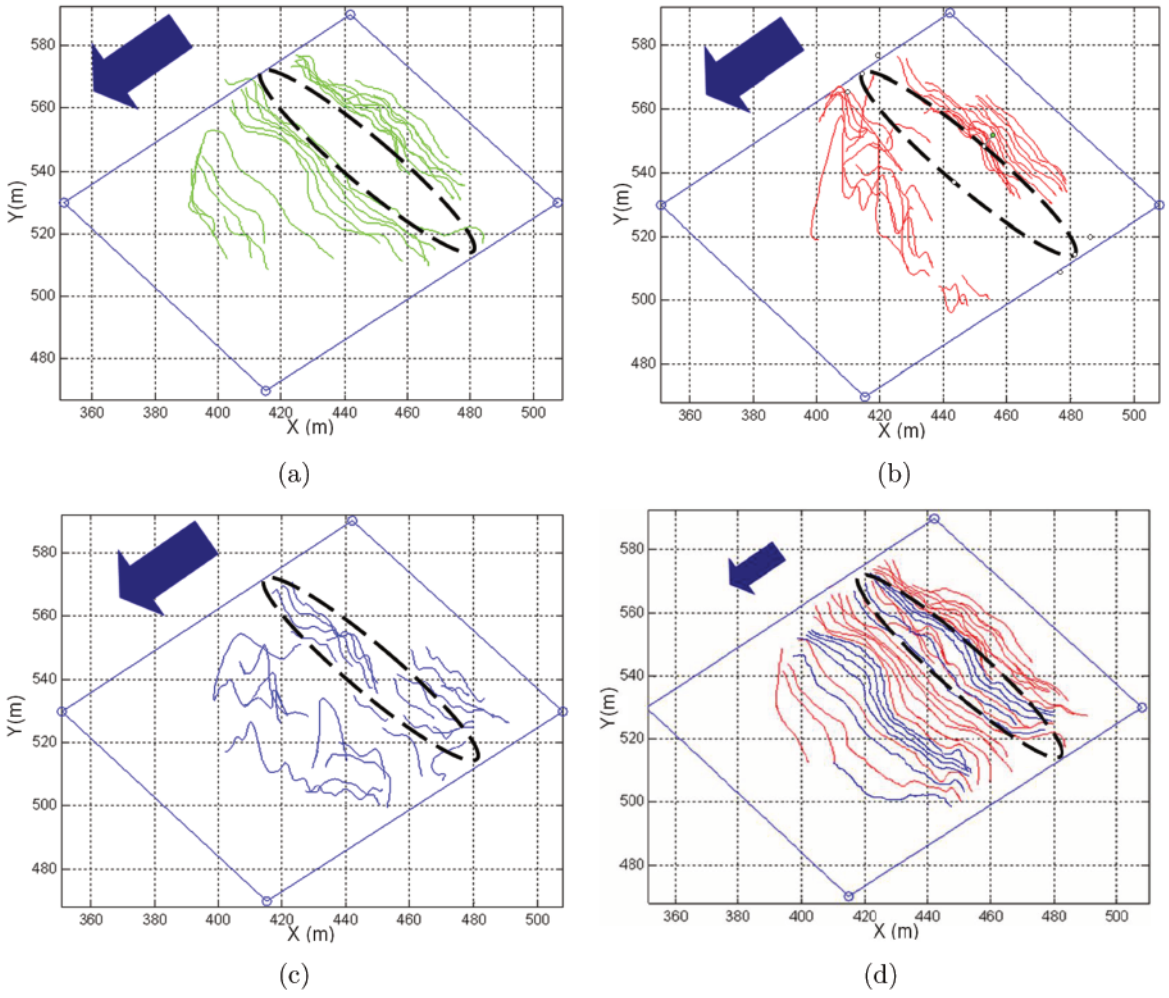



Figure 1.23: Time evolution of fire front location, the dotted ellipse represents a fire zone characterized by high presence of smoke. (*a*) visual camera on UAV. (*b*) frontal ground visual camera. (*c*) frontal ground infrared camera. (*d*) fires front merged [51].

## 1.4 Fire measurements with aerial systems

### 1.4.1 Airborne systems

Ononye *et al.* presented in [76] a method to determine fire perimeter, fire line, and fire propagation direction using as input multi and hyper spectral images from the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), MODIS Airborne Simulator (MAS) and Wildfire Airborne Sensor Program(WASP).

AVIRIS is an optical sensor developed by NASA that contains calibrated images of the spectral radiance in 224 contiguous spectral bands between 400 to 2500 nm. It was tested by NASA on two altitudes, 20 km and 4 km over the ground level. MAS is an airborne scanning spectrometer that acquires high spatial resolution imagery of surface features. WASP is a multispectral camera system that can discriminate fire from solar reflection by measuring relative brightness into the band 0,4-15 $\mu$m. It is associated with an IMU sensor and it was tested by NASA with an altitude of 1,5 km.

To estimate the line of the fire front, the authors uses a gradient operation on each spectral band and calculate the edge curve from the contours obtained. The fire direction is computed for several points of the front of fire. To estimate the orientation at certain points of the front line, the normal of the fire front curve at these points is calculated. The obtained fire front line separates the burning vegetation zone from the area that has not yet been burned, and this information can help to determine the propagation direction (Fig. 1.24). Only the information about fire line are estimated.

(a) (b) (c)

Figure 1.24: Image of big fire. (*a*) multi spectral image. (*b*) segmented burned area (*black zone*). (*c*) fire front direction at certain points [76].

Paugam *et al.* presented in [77] works on the use of "Handheld" thermal imaging cameras

to study vegetation fire behavior and collect data from airbone platforms. The authors propose an approach based on the automated detection of a set of fixed thermal “ground control points,” coupled with the use of a linear transformation matrix for warping the raw IR imagery to a fixed coordinate system. In this way, it is possible to georeference the thermal images obtained from a camera pointing at nadir or off-nadir station positions (tripod mounted, high platform, portable mode from a helicopter). In this study, the front lines of a fire propagating on a plot of nine hundred and forty-five square meters are referenced in the same reference system. From a set of front fire lines and their distances the rate of spread and the direction of propagation are estimated using methods similar to those presented in [76]. This method does not make it possible to estimate geometrical characteristics such as the flame height or the surface of the front face.

### 1.4.2 Unmanned Aerial Vehicle systems

Unmanned Aerial Vehicles have less autonomy and are much less resistant to wind and turbulence, but are cheaper and faster than traditional airborne vehicles. It is possible to fly to monitor and detect forest fires passing over fire many times compared to a classical air system. These are the reasons that today UAVs equipped with optical systems have a great potential for forest fires detection and monitoring. In the article [78] it is presented a review of different architectures of UAVs and its capacities: fixed wing, rotatory wing, airship. Each of these systems presents a variety of technological issues and practices. Methods to detect and monitor fires using single or cooperative UAVs are briefly discussed. A part entitled "Vision-based technologies for automatic forest fire diagnosis and prognosis" presents all the points to consider when using cameras mounted on drones.

The use of multiple UAVs for surveillance, detection, localization, and measurement of forest fires is described in several articles [79, 80, 52]. The basic elements of a general UAV-based system of surveillance of forest fires are a team of UAVs, with different kind of sensors, and a ground control station, as it show in Fig. 1.25.

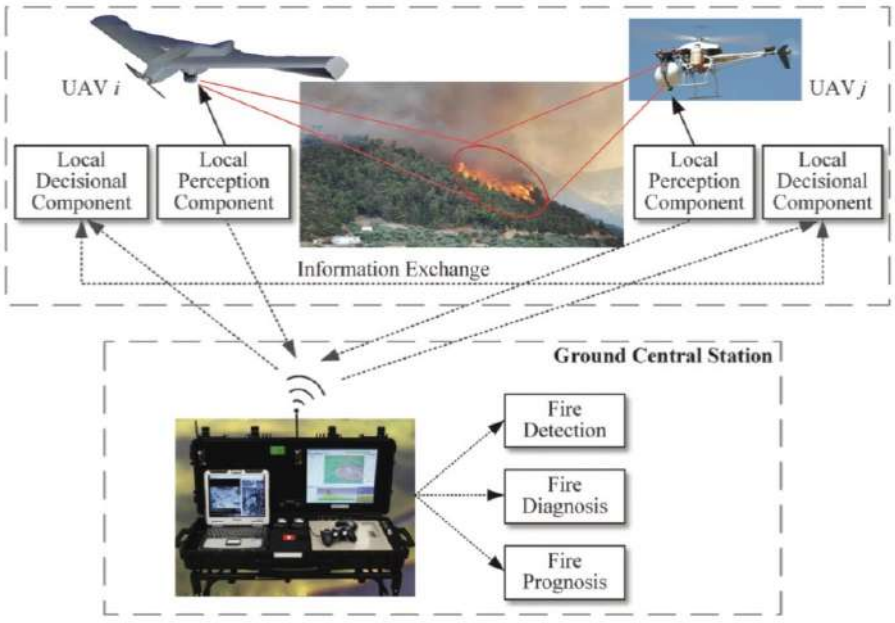
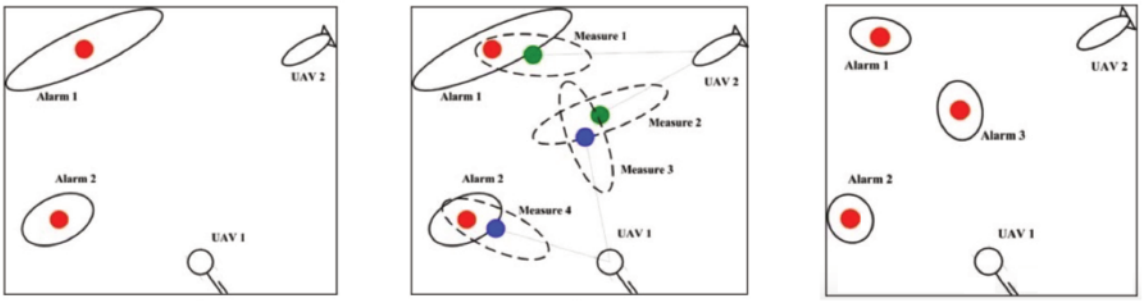



Figure 1.25: Conceptual UAV-based forest fire detection and diagnosis [78].

Martinez-de Dios *et al.* in [79] use 2 helicopter UAVs and 1 airship UAV for a precise estimate of the location of the fires. With the use of several drones it is possible to increase the position accuracy of the fire. In step 1, each UAV computes the fire position using its sensors. This position is affected by noise sensors and an uncertainty is computed taking into account the sensors sensibility. In step 2, the entire fleet is sent to the GPS position of the fire previously computed (that is also the center of gravity of the uncertainty region) and each UAV makes its measure of fire position with its uncertainty region. In step 3, the regions are intersected and the barycenter of the intersection corresponds to the new position of the fire. Being an intersection, the new region is smaller than the previous region, and this mean that the accuracy is increased. It is possible to iterate this procedure using the final position computed as input for the step 2. In Fig. 1.26 are represented the positions and uncertainties obtained at each step of the procedure.




(a) (b) (c)

Figure 1.26: Positions and uncertainties at each step of the procedure in [79]. (*a*) fire position detected during the previous iteration and their uncertainties, presented as ellipses. (*b*) new measures with their uncertainty. (*c*) intersection of all the uncertainty regions.

The mounted sensors are the following: a micro far-infrared drone camera and a visual camera with 752 × 582 of pixel resolution and a focal length of 6 nm. These two kinds of cameras are mounted together to create a multi spectral image. The relation between the two images is:

$$s\mathbf{m}_{IR} = H_{\infty}\mathbf{m}_{VIS} \tag{1.13}$$

where $\mathbf{m}_IR$=$[u\ v\ 1]^T$ is the point $\mathbf{X}$ in pixel homogeneous coordinates in infrared image, $\mathbf{m}_VIS$=$[u'\ v'\ 1]^T$ is the same point in pixel coordinates in visual image (Fig. 1.27), $s$ is the scale factor, and $H_{\infty}$ is the infinity homography matrix computed knowing at least four points correspondences in both images.

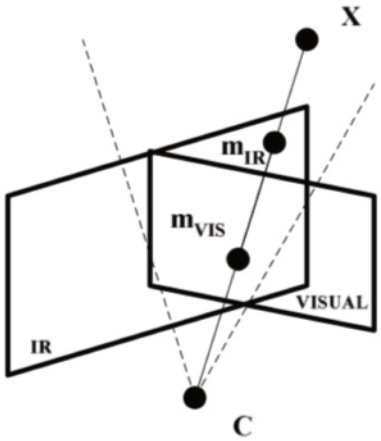



Figure 1.27: Geometry configuration of the cameras [79].

The Eq. 1.13 is valid if the distance of the center of projection of the two sensors is very small compared to the camera-target distance. $H_{\infty}$ is computed using a pattern calibration for the two cameras.

Fire segmentation and fire contours detection are obtained with the procedure described in Sec. 1.2.1, in order to compute the fire front detection as discussed in [49]. In Fig. 1.28 are shown fire fronts detected into UAV images.

Figure 1.28: Fire observation using UAV images: fire front line (*green*) and flame contour (*red*) [79].

The method presented in the article [79] is interesting because it shows the use of several drones to improve the accuracy of position measurement of a fire. However, as no information concerning the position of the drones and their distance from the fire and experimental data are given, it is difficult to evaluate the real effectiveness of the method.

## 1.5 Conclusion

In this chapter was presented the work of literature that is close to the problematic of this thesis. Methods for the estimation of fire geometrical characteristics like fire front location, fire height, flame inclination angle and fire width using image processing and computer vision with 2D and 3D data were described. The images are acquired from ground stations, hybrid systems composed by ground stations and UAVs, and aerial vehicles. The various solutions use visible and/or infrared cameras (in different spectral IR bands).

In the first part, were presented the geometrical characteristics needed by the researchers working on fire behaviour. In the second part, were discussed the methods using ground systems, composed by monocameras or sterevision devices. In the third part, was presented a solution using ground cameras and on-board cameras carried by drone. Finally in the last part were presented methods using only cameras carried by manned or unmanned aerial vehicle.

Concerning the solutions using ground-based cameras methods using monocameras on the ground from multiple viewpoints (in general frontal and lateral views) were presented. These methods, that synthesize the data in subsequent steps, fail to have fire depth information and are limited in estimating geometric characteristics. This kind of results is not complete to characterize the fire in all of its parts. Methods that do not use stereovision require bacons and landmarks which make these techniques difficult to use on natural terrain. A framework based on the use of ground stereovision devices has also been described. This system that couples pair of cameras and that is capable to compute the depth of each point of a fire was presented and discussed. Frameworks based on the used of stereovision solutions provides three-dimensional fire information and methods have been successfully developed for small outdoor fires. Solutions using ground-based cameras only can not be used on hard-to-reach terrain and there is a decrease of accuracy of measurements when the distance camera-fire increases what could be the case when the fire is moving away.

Manned air vehicles can fly with an altitude that provides good point of view and they are able to follow the the possible changes of direction of fire. Usually their dimension doesn't allow to go close to the fire. Moreover the planes are expensive to use and can not be supported by all research teams during their many experimental burnings carried out in unstructured environments. The use of drones allows to stay close to fire throughout its evolution. Drones and cameras have now accessible prices and it is now possible to develop platforms for the detection, monitoring and measurement of fires. Finally with UAVs there aren't any risks for humans.

This chapter has presented the basics of the works in relation to the topic of this thesis. It allowed to show the main ideas that have been published to overcome the problem of the measurement of fires by vision as well as the limits of the existing methods. It shown the principal works on the using of drone for fire position measurements. This work will serve to justify the solutions chosen for the development of new framework for fire geometrical characteristics estimation.

Chapter 2

# System overview

**Contents**



## 2.1 Introduction

For this thesis, a work on a multimodal stereovision system was developed. This device can be fixed on a drone and it is capable to acquire and store georeferenced images.

This chapter presents the study that was conducted to determine the architecture of the proposed solution to measure experimental fires geometrical characteristics (position, rate of spread, height, length, flame tilt inclination and surface) by drone. This solution is composed by an hardware architecture and a framework software.

This chapter is dedicated, in a first part, to a bibliographic study on wavelength emissions of fires and spectral bands suitable for their observation. In a second part, the developed stereovision system fixable to a drone is presented. In a third part, the UAV platform used to carry the vision system is discussed. Finally, the main principle of the framework for the 3D point computation is presented. The method developed in order to compute the geometrical characteristics from the fire 3D points is independent of the developed stereovision system and will be presented in the next chapter.

## 2.2 Spectral bands selected for the observation of wildfire

Each object with a temperature higher than 0 K emits electromagnetic waves which are its characteristic and which constitute its spectral signature. The emission spectrum of an element is the set of electromagnetic radiation wavelengths emitted by the electrons of its atoms when they make a transition from a higher energy state to a lower energy one.

### 2.2.1 Electromagnetic radiations from wildfire fires

A forest fire is a combustion process that can generate temperatures from 400°C to 1500°C, the glowing combustion of coal is between 500°C and 600°C and the flames reach temperatures of 1000°C [81]. In a combustion process the species that irradiate the most are soot and gases (for example CO, $CO_2$ and $H_2O$), with the contribution of soot that can exceed the combustion gases [82]. These solids emit in all wavelengths [83, 84]. In case of vegetation fires more than 90% of the gas mass emitted is composed by $H_2O$ and $CO_2$. In general the wildfire spectra are always composed of emission lines linked to carbon dioxide (4,3 $\mu$m) and to water vapor (2,7 $\mu$m and between 5 $\mu$m and 10 $\mu$m) more or less distinguishable according to soot emissions. The carbon dioxide emissions are particularly visible in the case of dry vegetation fires, while the water vapor emission are more important for wet vegetation. In Fig. 2.1, it is shown the spectral emission of fire with several flame widths (from 0,5 m to 4 m), where it is possible to note that there are different signatures mainly due to soot emissions which increase with the area of the fuel [82].

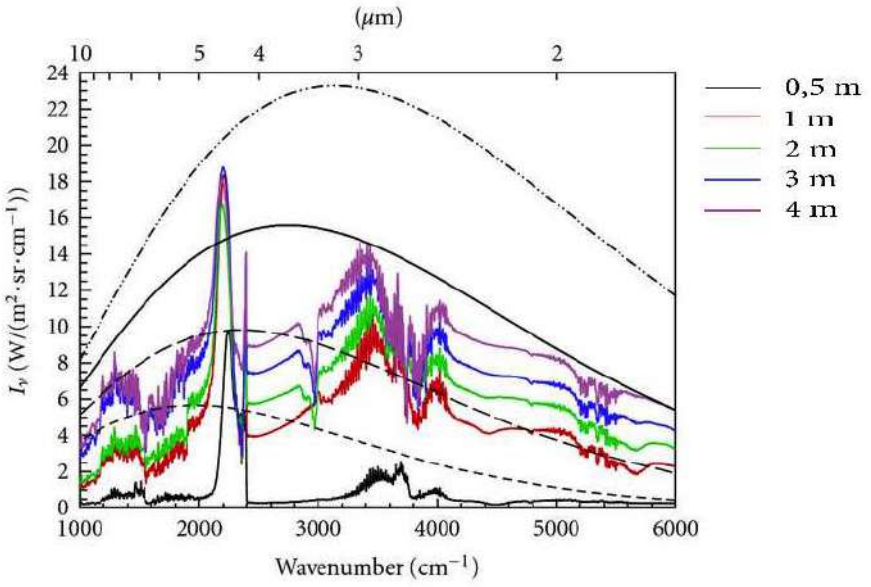



Figure 2.1: Maximum emission spectra from vegetation fires with a width varying from 0,5 m to 4 m. [82]

### 2.2.2 Spectral bands for fire observation

The atmospheric transmittance of a wave is the ratio between its incident intensity flux and its transmitted intensity flux, and it is a dimensionless quantity. In general, the atmosphere absorbs all the waves whose wavelengths are those of the gases that compose it. This consists of several gases such as nitrogen ($N_2$), oxygen ($O_2$), argon (Ar), water vapor ($H_2O$), carbon dioxide ($CO_2$), methane ($CH_4$), carbon monoxide (CO), nitrous oxide ($N_2O$), chlorofluorocarbons (CFCs) or ozone ($O_3$). Each of these gases absorbs the radiation passing through it corresponding to that of its emission, which produces many absorption bands. The broadest absorption bands are in the infrared (from 0,75 $\mu$m to 1 nm) and are due to the molecules of carbon dioxide and water. Fig. 2.2 shows the atmospheric transmittance in the different spectral bands of the infrared. Infrared wavelengths are classified in near infrared or NIR (from 0,75 $\mu m$ to 1,4 $\mu$m), short-wave infrared or SWIR (from 1,4 $\mu$m to 3 $\mu$m), medium-wave infrared or MWIR (from 3 $\mu$m to 8 $\mu$m), long-wave infrared or LWIR (from 8 $\mu$m to 15 $\mu$m), and far infrared or FIR (from 15 $\mu$m to 1 mm) [85].

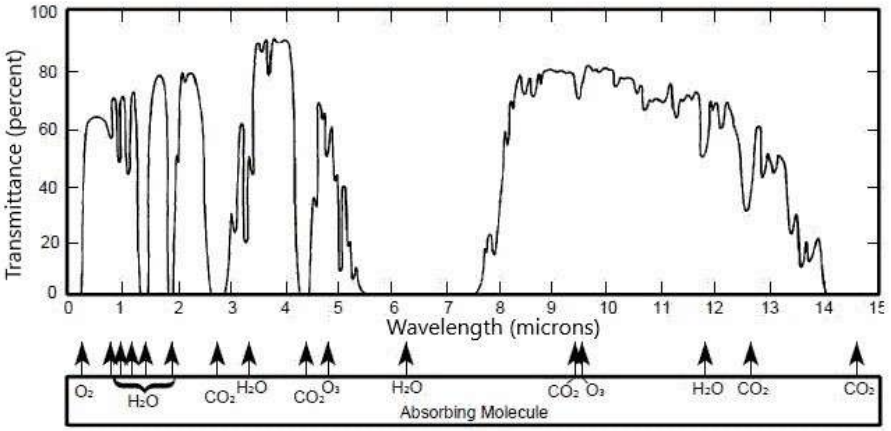



Figure 2.2: Transmittance of Atmosphere in the infrared spectral band [82]

.

In [83], Le Maoult *et al.* have conducted a study of the transmittance parameter for radiation emission by fire, using the complete IR band from near infrared to long-wave infrared. In Tab. 2.1 it is presented the atmospheric transmittance in each sub-band of the infrared spectrum for an object at 1500 K and situated at a distance of 100 m. Values close to unity mean that almost all the flux of radiations transmitted by the object crosses the atmospheric space.

| | Near infrared | Short-wave infrared | Medium-wave infrared | Long-wave infrared |
|---|---|---|---|---|
| $\tau_{atm}$ | 0,95 | 0,75 | 0,78 | 0,93 |

Table 2.1: Atmospheric transmittance each sub-band of the infrared spectrum for an object at 1500 K and situated at a distance of 100 m [83].

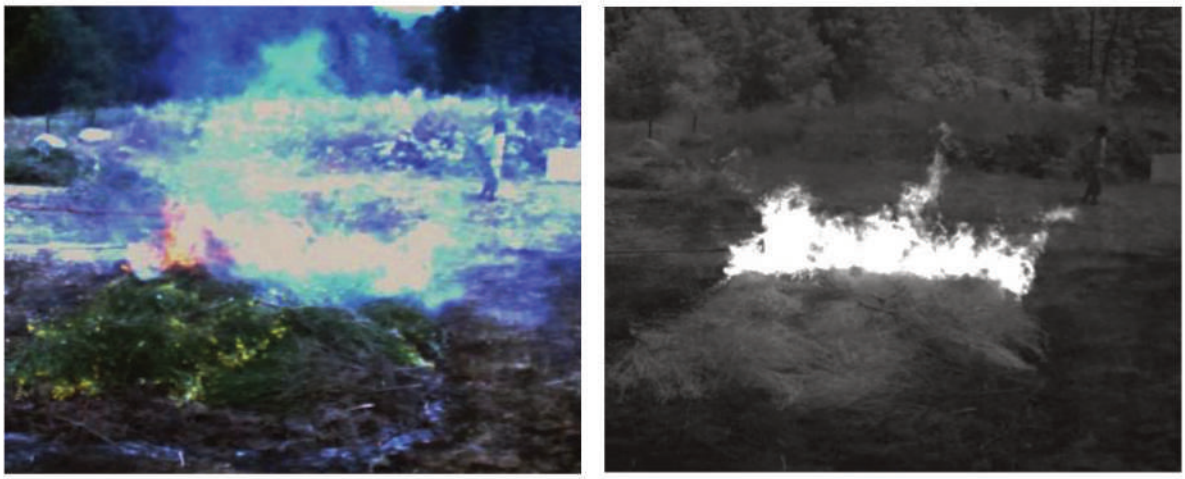
### 2.2.3 Comparison of fire images in visible and infrared spectral bands

A fire emits radiations more than the human eye can see; in human eyes only the visible spectrum can be detected; for this reason and given the very affordable price of visible cameras and their simple use, the community of researchers working on the problem of wildfire keeps the field of visible as a reference domain. Fire areas can be difficult to segment in visible images due to the fire texture and colors that can be very varied. Smoke generated by fire can also make difficult the segmentation of fire zones in image by masking them (Fig. 2.3).

Figure 2.3: Image of fire acquired in the visible domain.

The infrared band can be used to overcome this difficulty. Thermal images (NIR, MWIR, LWIR or FIR) can make appear flame areas with a high contrast of intensity with the environment and whose shapes are an over-envelope of those obtained in visible images.

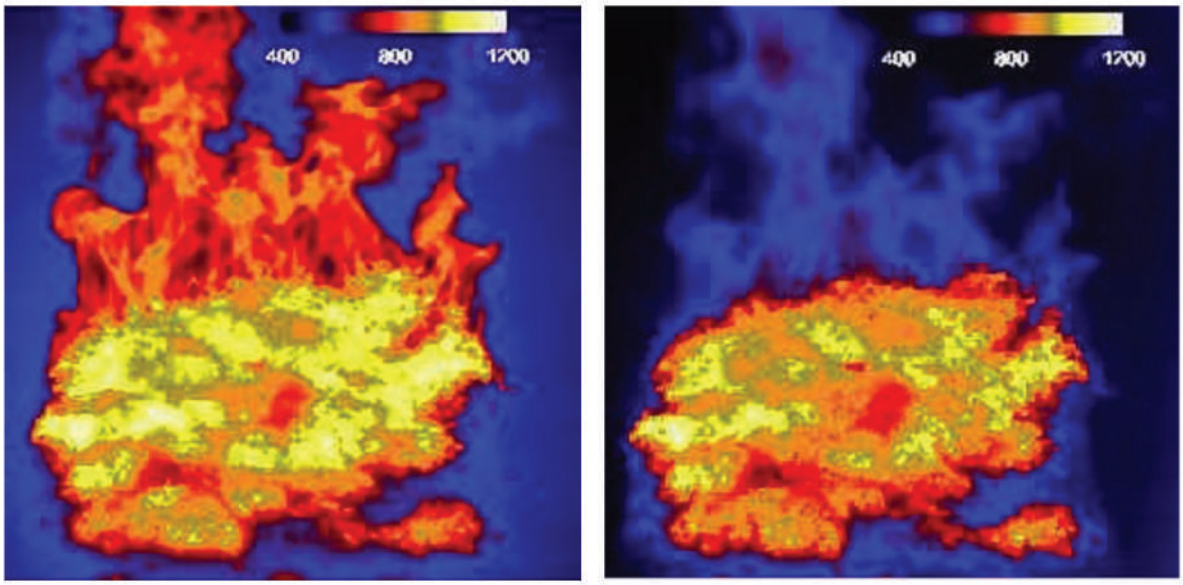
Toulouse *et al.* [54, 70] did a study on small vegetation fire images taken at about ten meters away from the camera and found that, in comparison with other spectral bands, the near infrared obtains images whose fire zones are closest to those obtained in visible images (Fig. 2.4). It should be noted that no test was performed for camera-fire distances greater than 10 m and that in the presence of a large quantity of soot fire zones may be masked.



(a) (b)

Figure 2.4: Fire images with smoke simultaneously acquired in the visible and near-infrared domains. (*a*) image in the visible spectrum. (*b*) image in the near infrared spectrum [70].

In [86], it is shown images of the same fire taken simultaneously in the FIR and MWIR spectra (Fig. 2.5). FIR and MWIR images present fire areas highly identifiable but they are very different for each spectrum. It can be seen that the MWIR image contains an area of fire integrating the base and the flame, whereas in the FIR image the fire zone is limited to the base due to the difference of temperatures in the fire in correlation with the spectrum used to acquired the image. For the measurement of geometrical characteristics of fire front, the FIR spectrum is not appropriate because parts of the flame front do not appear.




(a) (b)

Figure 2.5: Bi-spectral MWIR-FIR images. (*a*) image in the MWIR spectrum. (*b*) image in the FIR spectrum [86].

Gouverneur *et al.* in [87] present a study in which the best spectral bands for fire detection (long distance fire-camera) are identified: SWIR, MWIR, or LWIR. In this work, the ratio between the radiation emitted by a fire and those emitted by an object at a temperature of about 30°C (representing the average temperature of an object heated by the sun) is calculated. The results show that the image obtained in the MWIR band has the biggest contrast, but the images acquired in the LWIR and in the SWIR bands have also high contrast. It should be noted that the performances of theses bands decrease with the fire-camera distance.

Fig. 2.6 presents fire images acquired simultaneously in the visible and LWIR bands with two cameras having different pixel resolution and focal.

(a) (b)

Figure 2.6: Fire images obtained simultaneously in the visible and LWIR spectra. ($a$) visible image. ($b$) LWIR image.

It can be seen in the LWIR image that the fire area is easily distinguishable. This zone is larger that the corresponding one which appears in the visible image and it has no texture. To conclude, the images of distant fires obtained in the LWIR range show areas of flames with a high contrast of intensity with the environment and with shapes that are an over-envelope of fire zones obtained in visible images. The fusion of the information obtained in the visible and LWIR images should make it possible to detect fire pixels in the visible images efficiently.

## 2.3 Materials composing the vision device

### 2.3.1 Choice of a vision system and its characteristics

Measuring objects outdoors from an image can be a very complex operation because it requires to position in the camera's field of view landmarks whose positions are known. This is very complicated on non-planed, difficult to access or with dense vegetation lands. Stereovision is a process that provides depth information from two simultaneous and scaled images of the same scene. This method was chosen to obtain 3D points of fires and for measuring their geometrical characteristics.

In a stereovision system, it is possible to compute the theoretical depth measurement precision $\Delta z$, given the focal length $f$ of the cameras, the average distance camera-target $z$, the baseline length between the two cameras $B$, and the imprecision disparity $\Delta d$. The relation is

expressed by:

$$\Delta z = \left| \frac{z^2}{Bf} \right| \Delta d \tag{2.1}$$

where $z$, $\Delta z$ and $B$ are expressed in meters and $f$ and $\Delta d$ in pixels.

From this equation it can be seen that the error in $z$ is proportional to the square distance camera-target, so it is important that $z$ be as small as possible. Thus, it was decided to carry the vision system by drone in order to follow the fire throughout its spread maintaining a good view and a minimum (but safe) camera-fire distance. For this reason, part of the choice of vision devices took into account their weight and shape so as not to interfere with the drone's capabilities.

Equation 2.1 shows also that for $z$, $f$ and $\Delta d$ constant, the greater the inter-camera distance and the smaller the error on the estimated distance $z$. The cameras of the stereovision system considered in this work being fixed on an axis, the inter-camera distance is the length of that axis which is carried by the drone. It was established by a professional drone pilot that for a DJI S1000 drone the maximum length of this axis is 1 m. With this value for $B$, an imprecision disparity equal to 1 pixel, and considering the visible cameras used, the theoretical error in $z$ is less than 20 cm for a fire-camera distance equal to 15 m.

The reference spectrum for researchers working on the issue of wildland fires is the visible one. Fire areas appear texured which allows to have features and 3D points by stereovision over the entire zone of fire. However, it is difficult to detect fire pixels in visual images due to the various color and texture of flame and possible presence of smoke. The images of fires obtained in the infrared spectrum show areas of fire with a strong contrast of intensity with the environment. It was decided to use multimodal information from infrared and visible images to segment fire areas in visible images. It has been chosen the LWIR spectrum which is suitable for obtaining information (base and flame) from distant fires and in the presence of smoke, besides the fact that it is possible to find LWIR cameras for drones with an affordable price and small dimensions.

### 2.3.2 Proposed solution

This section presents the stereovision system portable by drone, that has been developed for the measurement by vision of geometric characteristics of wildfires.

The device is composed by two different stereovision modules, one working with visible cameras and the other with infrared cameras.

**Stereovision visible module**

The visible module contains two Leopard Imaging MIPI OV4689 cameras [88] controlled by an image signal processor (ISP) bridge chip called OmniVision's OV580 (Fig. 2.7).

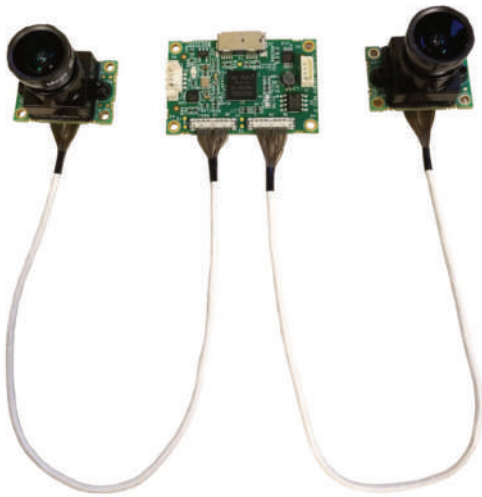


Figure 2.7: Video Modules OV4689 MIPI 4M Camera Module.

An ISP is a type of multimedia processor able to process the images from digital cameras. In this device, the ISP is used to merge the two video streams coming from each camera at 30 Hz and to create a single image containing the two frames one next to the other. The characteristics of the used camera are a 2,8 mm of focal length, a pixel size of 2 $\mu$m, an horizontal field of view of 98°, a CMOS sensor of 1/3", and an IR cut filter at 650 nm. The resolution of the image is set to 2208×1242 pixels. The cameras and the ISP are powered by a USB3 cable with a tension of 5 V and the current required is 296 mA. The weight of the two lenses and the ISP chip is 30 gr.

The cameras are mounted on an axis with a distance inter-camera of 0,9 m. The support is a plywood one composed of fibrous and composite structure that limits its deformations. Holes were made in order to reduce its weight. An aluminum bar was added on the upper side in order to rigidify the system (Fig. 2.8).

Figure 2.8: Visible stereovision system mounted on the drone DJI S1000.

Figure 2.9 presents an image of the vision system carried by the DJI S1000 UAV located at the rear of a fire approximately 12 m away. Figure 2.10 shows two visible stereoscopic images taken by this device in a very near position of the one presented in Fig. 2.9.

Figure 2.9: Vision system carried by the DJI S1000 UAV located at the rear of a fire approximately 12 m away.

(a) (b)

Figure 2.10: Images kept simultaneously by the visible stereovision system. (*a*) Left image. (*b*) Right image.

### Stereovision infrared module

The IR stereovision module is composed by two FLIR Vue Pro R cameras [89] working in the LWIR spectrum.

The characteristics of this camera are a 9 nm of focal length, an horizontal field of view of 69°, a resolution of 640×512 pixels, and an operating frequency of 30 Hz. The weight of each camera is 113 gr.

The image acquisition of this camera can be triggered by an external 50 Hz PWM signal with a duty cycle of 20 ms.

The IR sensors are positioned above the visible ones as it appears in the Fig. 2.11, in such a way to obtain two pair of visible-IR cameras distant of 0,9 m; the distance between their lens centers is 5 cm.

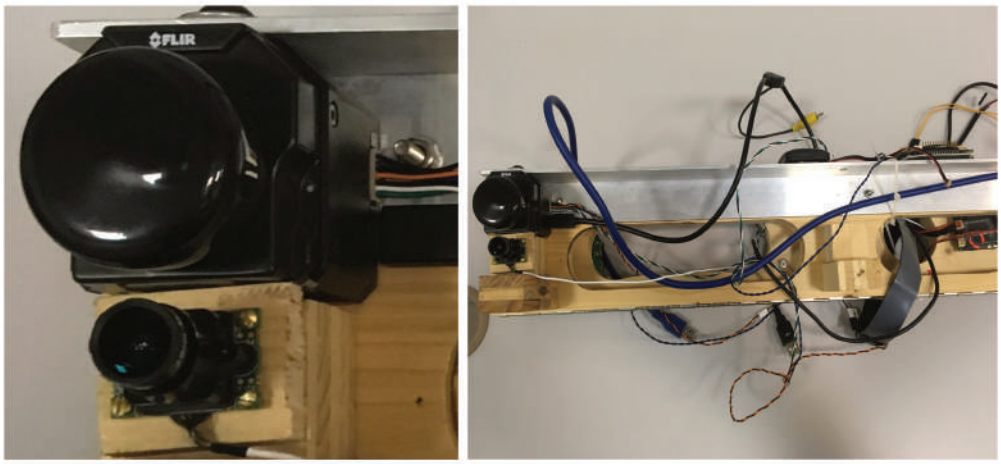



Figure 2.11: IR and visible cameras placed below each other.

Figure 2.12 shows the multimodal stereovision system composed of infrared and visible cameras screwed in the plywood axe and carried by a drone.

Figure 2.12: Visible and infrared stereovision systems mounted on the drone DJI S1000.

The stereovision system is positioned on the undercarriage kept in a fixed position. Figure 2.13 shows the attachment system.

Figure 2.13: Support attachment system.

The IR camera needs a voltage of 5 V and has a maximum consumption of 7,8 W. A power supply is added for the vision device in order to power them. This battery is added on the drone to don't influence the balance of the vision system.

To synchronize the two IR cameras, a Raspberry computer is used to generate the PWM signal required for their external trig. From its GPIO port, a cable is connected and split in two elements, each one connecting a camera. A particular attention was paid in order that the cable lengths are very close (difference less than 1 mm) and that the welds on the cables are as homogeneous as possible and very light.

Figure 2.14 shows two IR stereoscopic images taken by this device at the same time that the visible images shown in fig. 2.15.

(a) (b)

Figure 2.14: IR stereoscopic images. ($a$) Left image. ($b$) Right image.

(a) (b)

Figure 2.15: Visible images. (*a*) Left image. (*b*) Right image.

### Additional sensors and system wiring

Others sensors that are added on the drone are used to obtain the position and the orientation of the vision device. In particular, an IMU board, a barometer sensor, a GPS, and a digital compass are considered. All the data of these sensors are read by a board equipped with an 8-bit microcontroller; an ATmega8 chip mounted on Arduino is used. The values read by the Arduino controller are sent to a Raspberry computer via USB cable.

The position of the vision device is computed by the GPS sensor ans the barometer sensor. The GPS sensor is an "u-blox M8 GNSS" [90] with a position precision of 2 m. The GPS sensor is fixed between the two visible cameras in the center position, and it gives two angular values called "latitude" and "longitude". Latitude is the angular distance, measured in degrees along the meridian arc, between the equator and the parallel passing through the considered point. Longitude is the angular distance, measured in degrees along the parallel arc, between the "Greenwich Meridian" (prime meridian) and the meridian passing through the considered point (Fig. 2.16).

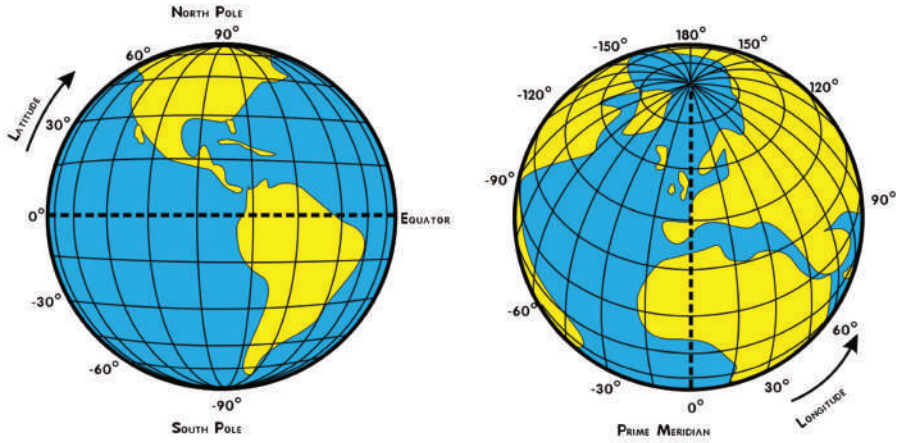



Figure 2.16: Latitude and longitude on the Earth.

Given that the barometer sensor has a precision of 10 cm, and the altitude returned by the GPS sensors has a precision of 2,5 m, to obtain the vertical distance of the device from the sea level (altitude), the barometer sensor is preferred to the GPS sensors.

The orientation of the vision device is given by the IMU board and by the compass sensor. The IMU board is used to obtain the roll and pitch angles that indicate the inclination angles stereovision system, the compass sensor is used to obtain the heading angle of the device (Fig. 2.17).

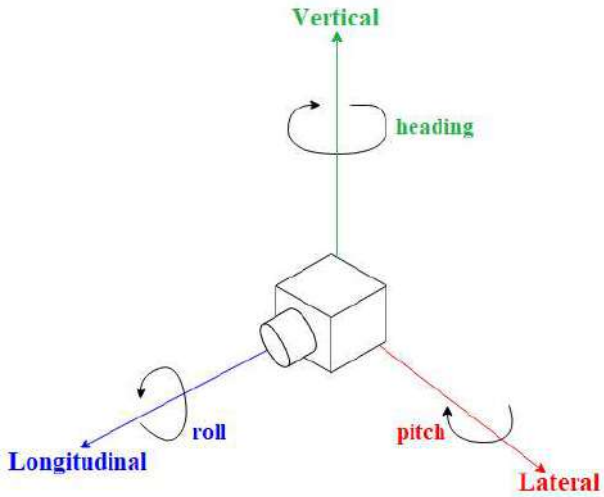



Figure 2.17: Roll, pitch and heading angles of the camera.

Roll, pitch and heading angles are all equal to zero when the stereovision device is horizontal and points to North. Roll angle is positive if the device is rotating to the right, and it is negative if the device is rotating to the left. Pitch angle is negative if the device is pointing down (the condition of the device that is pointing up it is not considered because the cameras are always higher than the fire). Heading angle is positive if the device is pointing to East, negative if the device is pointing to West.

Figure 2.18 presents the physical connections of all the elements of the multimodal stereovision system. Load balancing tests were carried out to find the best position to fix the stereoscopic system on the drone landing gear.

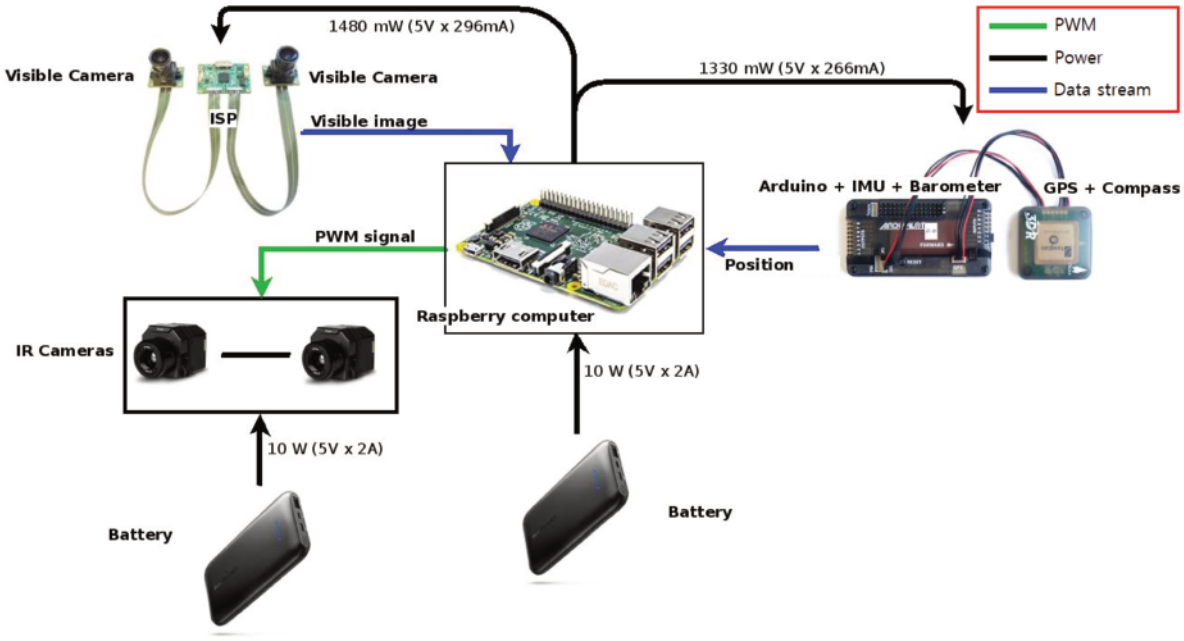



Figure 2.18: Diagram showing elements and connections of the multimodal stereovision system.

As it can be seen, the Raspberry computer is positioned at the center of the logical system and all the connections are directed towards it; it acts as the coordinator of the various components. It is powered by an USB battery with 5V and 2 A. The power entering into the Raspberry computer is 10 W. This value is sufficient for the operation on the Arduino board and the vision stereoscopic device. The power required by the Arduino is 1 W, the one required by the GPS/Compass sensors is 330 mW and the one required by the visible stereovision module is 1480 mW. To function properly the Raspberry computer needs 1,2 W. The total power consumption of the set composed by the Arduino board, the GPS/Compass, the visible camera module and the Raspberry computer is more than 4 W. The remaining power is less than 6 W and it is not sufficient for the IR camera module operating. As mentioned above, infrared cameras need an energy load of 7,8 W and for this reason a separate power supply of 12,5 W is used. It is preferred to use this solution to avoid overloading the Raspberry microchip with a great electrical power.

As it is shown in Fig. 2.18, the ISP is connected to the visible cameras and to the Raspberry computer. This component receives the two frames coming from the visible cameras and joining them into a single image. This image is sent via USB cable to the Raspberry computer in a continuous stream. Given the speed of fire propagation, the Raspberry computer is configured in order to read the image stream each 4 seconds.

The Arduino board reads the continuous data coming from the accelerometer, the GPS/Compass, and the barometer, with a frequency of the order of milliseconds. These three sensors use three separate communication channels of the Arduino board so as not to share the same bus and take three readings while keeping the latency time small between each of

them (estimated in the order of the microsecond). The accelerometer, GPS/Compass and barometer data are sent to the Raspberry computer via a second USB cable using a MAVlink protocol [91].

The Raspberry, the Arduino board and the battery have been added taking into account the balance of the axis. The total payload for the complete multimodal stereovision system is 3,2 kg. Fig. 2.19 shows the complete vision framework before a flight.

Figure 2.19: Complete vision framework before a flight.

**Vision modules synchronization and data storage**

The synchronization between the visible cameras module and the IR cameras module is obtained via software. On the Raspberry computer, an Ubuntu base operating system is mounted, with only the kernel modules useful for the operation of the developed device installed (this choice is made to keep electricity and computational consumption low). A proxy system is implemented in order that the computer becomes a server. The Arduino board and the ISP card act are its clients. With a frequency of 1/4 Hz, the Raspberry computer performs three main operations: store the value of roll, pitch, yaw, heading, GPS, and altitude given by Arduino; send the PWM signal to the IR cameras; and load the current image from the channel connected to the ISP.

A code stores into the computer internal memory the visible images whose names contained the inclination, position and orientation data; the IR images are stored directly into a memory card installed in each IR camera.

Figure 2.20 presents the diagram of the synchronization steps between the visible cameras module and the IR cameras module.

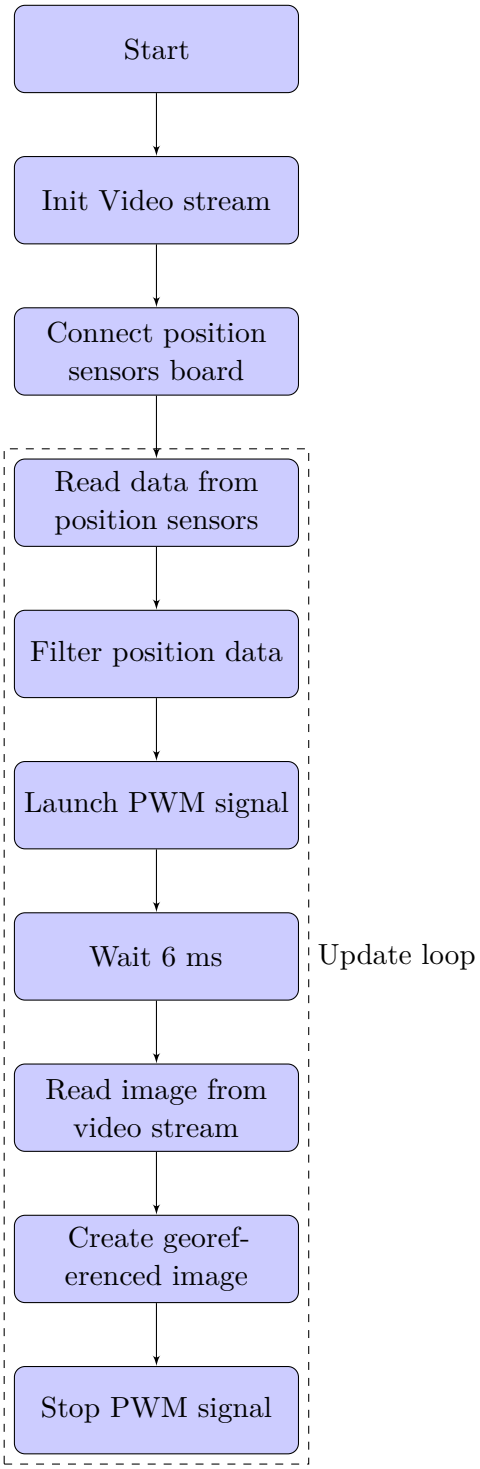



Figure 2.20: Diagram of the synchronization steps between the visible cameras module and the IR cameras module.

The "Start" step represents the boot of the Raspberry computer; in this step the ISP board

and the Arduino board are power on.

The "Init Video Stream" step initializes the video stream channel associated to the port where the ISP board is connected. The "Connect position sensors board" step opens a serial communication channel with a MAVlink protocol on the port where the Arduino is connected, with a bit rate of 115 200 bits/s.

If the previous steps are terminated without errors, the code enters into an infinite loop indicated in the scheme with "Update loop"; in this loop each 4 seconds two pair of stereoscopic images (visible and IR) will be produced. This loop is stopped manually.

In the "Read data from position sensors" step, the data are read from the Arduino. The Arduino board sends a continuous stream via a MAVlink message containing the status of the sensors to which it is connected. The state of the sensors represents the measurements: the accelerations on three axes in m/s (these measurements are converted directly into the Arduino in roll, pitch and yaw angles in radiant), the GPS position in decimal degrees, the orientation in centigrade degrees, and the pressure recorded by the barometer (the pressure is converted into the Arduino in altitude expressed in mm). These data may be influenced by error or noise, and these errors may appear as rapid signal fluctuations with respect to the real value.

In the "Refine position data" step, refining is applied to the sensor signals. This action is carried out during the pause between two pairs of stereoscopic images, and it is used to collect the position data coming from the Arduino. Using these data, a new value between the new received data and all the previous values is computed. The weight of the readings is chosen to be high for older value readings and small for newer value readings. With the assumption that the drone moves very slowly to maintain a good image stabilization, it is assumed a slight variation between two successive positions of drone and images. The last value read is multiplied by a very small factor. The characterization of the refining used is that of considering 95% of the previously calculated value, and 5% of the new value.

$$\overline{m}_i = \overline{m}_{i-1} * 0.95 + m_i * 0.05; \quad \overline{m}_0 = 0. \tag{2.2}$$

where $\overline{m}_i$ is the average position at the instant *i*, $\overline{m}_{i-1}$ is the average position at the instant *i-1*, $m_i$ is the new read data at the instant *i*. The initial value of the average $\overline{m}_0$ is set to 0.

For the roll, pitch, and yaw angles the chosen threshold is 15°, for the heading angle it is 45°, for the altitude measurement it is 5 m, and for the GPS position it is 4 m. The threshold values were chosen for heuristics. As mentioned above, a moving average filter is characterized by good stability but a slow response speed. Avoiding to insert values with a high probability of error in the average calculation increases the speed to get a good estimation of the analyzed value. At the end of the pause state the position values are estimated to their averages.

In the "Launch PWM signal" step, a signal PWM generated by the Raspberry computer

is sent to the IR cameras. Experimentally, it has been found that there is a delay between the instant when the PWM signal instruction is executed by the computer and the instant when the IR images are acquired. An oscilloscope was used to calculate this delay. First, the oscilloscope was synchronized with the Raspberry clock and the signal coming out of the GPIO port to compute the time that elapses between the instant time of the launch of the instruction code and the instant time in which the PWM signal is emitted. In a second step, it was synchronized with the IR camera clock to calculate the instant time between the signal reception and the instant time when the IR camera takes the picture. Even for visible cameras, there is a delay between the instant time when the instruction to read the video channel is executed and the instant time when the picture is received by the Raspberry. This time was measured using the reading of the clock timestamp of the Raspberry and will be subtract to the delay previously computed, and the final delay was found. Taking into account this time shifting, the system has been configured to read visible images from the ISP 6 ms after launching the instruction to generate the PWM signal through the Raspberry computer"

In the Fig. 2.20, the "Wait" step corresponds to the stage during which the code is stationary for 6 ms.

In the "Read image from video stream" step, the instruction to read a new picture coming from the ISP is launched. In the "Create georeferenced image" step, the image coming from the ISP is stored into the memory card of the Raspberry computer. Its name is built using a comma seperated value message structured as follow (the serparator symbol chosen is ";"): "id_number;roll;pitch;yaw;heading;latitude;longitude;altitude.jpg". The sequence number is useful to identify the time between two pictures. Roll, pitch and yaw are expressed in radiant, heading in degree, latitude and longitude in degree multiply by $10^7$, and the altitude is in millimeters. At the end of this step the program provides a pair of geolocalized visible images and a pair of IR images taken at the same instant.

In the "Stop PWM signal" step the PWM signal is reset. This step is useful to recreate the signal transition for the next "Update loop" to be captured by the IR cameras. To reset the PWM signal, its duty cycle is set to 10 ms.

This multimodal stereovision system has as weak points a GPS sensor with a medium level accuracy and a synchronisation of IR and visible images in the order of ms.

#### 2.3.2.1 Drone navigation system

The drone used to carry the developed multimodal stereovision system is the DJI Spring Wing octocopter S1000 [92] with a wing extension of 1 m It was chosen because it can carry a total weight of 11 kg in the air and has a flight time of 15 minutes. Moreover, its frame allows to install additional components. A navigation system integrating an IMU card, a barometer, a GPS/Compass sensor and a radio receiver for manual control has been add to the UAV. Particular attention was paid to position these elements on the drone to optimize its flight. In

Fig. 2.21 is presented the drone adapted with the material required to carry out the operations that this thesis project proposes.

Figure 2.21: Drone mounting the material for power and for video transmission to the ground.

To adjust the speed to be assigned to the UAV engines during flight operations, a Proportional-Integral-Derivative control (PID) regulator is used, as presented in [93] (Fig. 2.22).

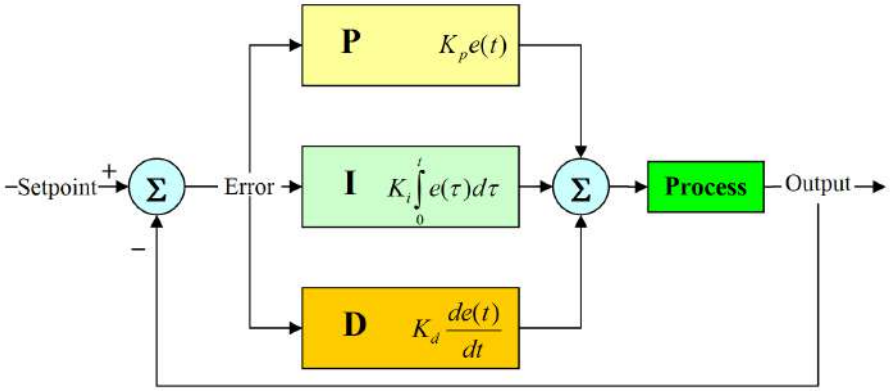



Figure 2.22: Block diagram of a PID.

The controller acquires a value as input from a process and compares it with a reference value. The difference, the so-called error signal, is then used to determine the value of the controller output variable. The PID regulates the output based on the value of the error signal (proportional action), the passed values of the error signal (integral action) and how quickly the error signal changes (derivative action).

The three actions of a PID are calculated separately and added algebraically:

$$u = u_P + u_I + u_D. \tag{2.3}$$

The proportional action is obtained by multiplying the error signal "e" with an appropriate constant:

$$u_P = K_P * e. \tag{2.4}$$

The integral action is proportional to the integral in time of the error signal "e", multiplied by the constant $K_I$:

$$u_I = K_I \int e(t)\,\mathrm{d}t \tag{2.5}$$

The derivative action is used to improve controller performance:

$$u_D = K_D \frac{\mathrm{d}e}{\mathrm{d}t} \tag{2.6}$$

The Ziegler–Nichols method [94] was used to configure the PID regulator values that control the motors. It is an algorithm for finding the called "critical gain", from which the other PID parameters are be derived. First the process is controlled by an exclusively proportional controller ($K_I$ and $K_D$ are set to zero), next the gain of the proportional controller is gradually increased. The critical gain $K_U$ is the value of the gain whereby the controlled variable has sustained oscillations that do not disappear after a transient: this is a measure of the effect of delays and process dynamics. Next the critical period $P_U$ of the sustained fluctuations is recorded. Finally the constants for the PID controller are determined in the following way:

$$K_P = 0.6K_U; \tag{2.7}$$

$$K_I = P_U/2; \tag{2.8}$$

$$K_D = P_U/8. \tag{2.9}$$

The input of the PID regulator are the GPS position, the IMU values, the compass and the signal coming from the radio. The output are the PWM values to be assigned to the eight motors, and the new position values.

The regulator is composed of three rings, where each ring is a PID regulator used to control one navigation variable: the position (first ring), the speed (second ring), and the acceleration (third ring).

In order to place the drone in such a way to acquire stereovision images of the entire fire front, a frontal HD camera was added to the axe (Fig. 2.23).

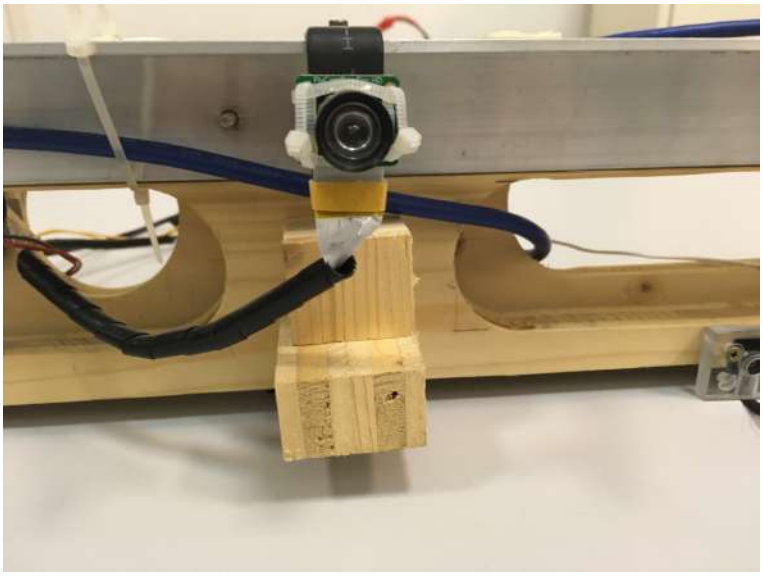


Figure 2.23: Frontal camera to send flying image during navigation.

This camera is connected to a signal transmitter received on the ground. The signal is sent to a field monitor using analog transmission with a different frequency than the radio (5,2 GHz for the video return and 2,4 GHz for the radio).

#### 2.3.2.2 Description of the image processing process

This subsection describes the proposed process for estimating the fire geometric characteristics from stereoscopic images acquired by drone. Figure 2.24 shows the functional diagram of this process.

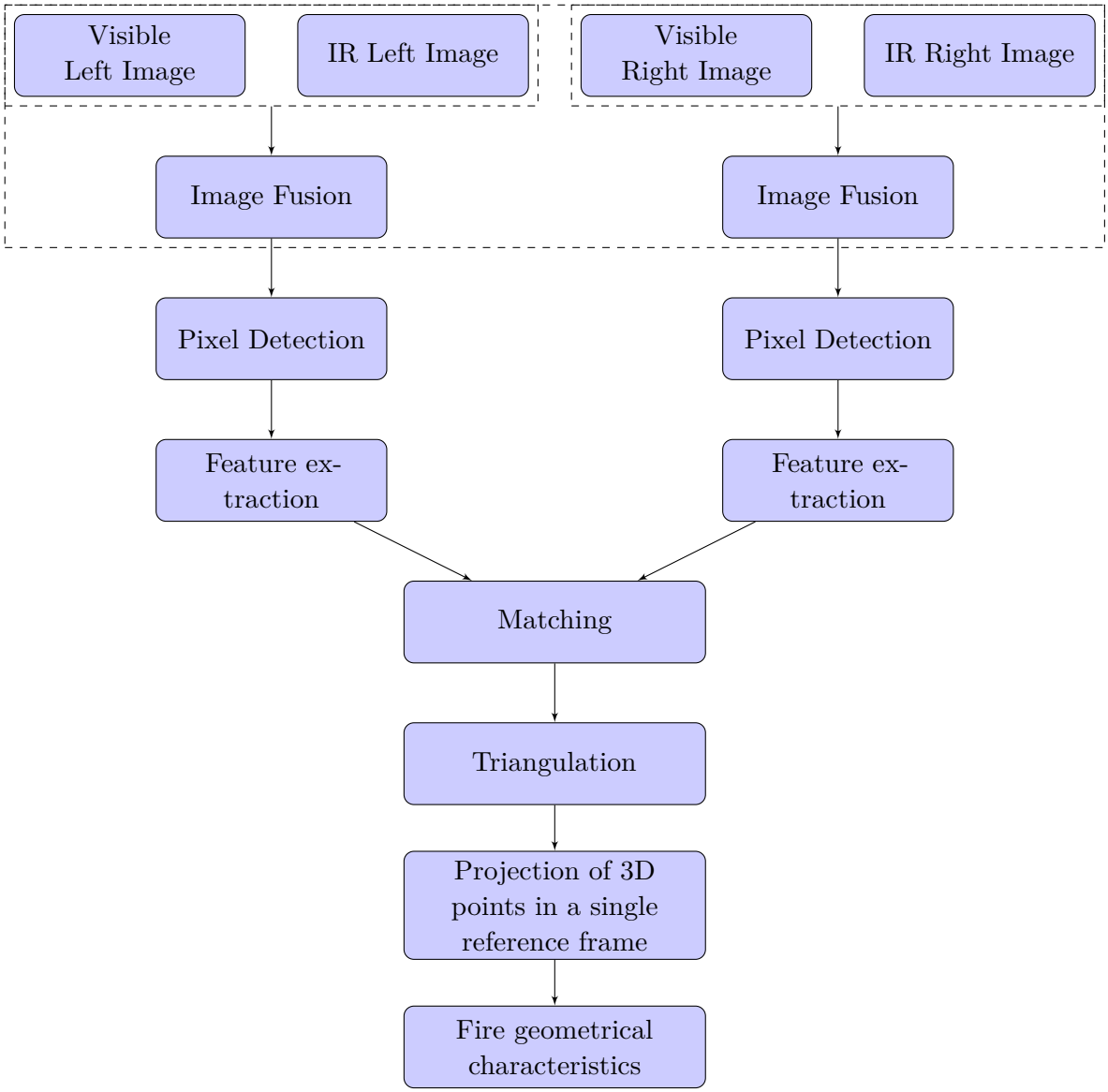



Figure 2.24: Functional diagram of the proposed system.

The first step is a pixel fire detection processing in the images by the multimodal stereovision system carried by drone.

The second step is a procedure for matching points of interest detected in visible stereoscopic images.

The third step allows the calculation of three-dimensional coordinates of points from the paired fire pixels.

The fourth step corresponds to a process of projection of 3D points obtained for different positions of the stereovision system in a common frame.

Finally, the last stage of the process is the estimation of the geometric characteristics of the fire.

This chapter presented the solution proposed in this thesis for the estimation of the geometric characteristics of a propagating fire by drone. The presented device consists of a multimodal stereovision system carried by drone. The cameras were choosen after a bibliographic study on vision fire in the different spectral bands and taking into account their dimensions, weight and possibility of triggering by an external signal. Visible and infrared cameras are synchronised in order to obtain simultaneous images. The overall process of processing stereoscopic images for the estimation of geometric fire characteristics was also presented. In the continuation of this report the developing of each stage will be explained.

Chapter 3

# Theoretical principles of stereovision

Contents



## 3.1 Introduction

This section describes the theoretical principles of the stereovision to reconstruct an object in 3D form from several shifted views of this object. One of the reference works in the field of 3D vision is the book by Hartley and Zisserman [95] and this section is largely inspired by it. In the first part, a geometric model of a camera is described. A second part is devoted to the calculations of the parameters of this model. In a third part, the geometric model of a stereovision system is presented. A fourth part is dedicated to the calculations of the parameters of this system. Finally, one last part is devoted to calculations allowing to obtain three-dimensional points from coordinates of the matched fire pixels.

## 3.2 Camera geometric model

A camera is defined by its optical center $O_c$ (this is also where the camera aperture is located), its optical axis $Z_c$ (perpendicular to the image plane and passing through point $O_c$), the center of the image, called "principal point", ($c_x$, $c_y$) (which is the intersection between the image plane and the optical axis), its focal length $f_c$ (which is the distance between the points $O_c$ and ($c_x$, $c_y$) in mm). These parameters are called "intrinsic parameters" of the camera. ($u$, $v$) are the image coordinates of a pixel $x_p$ corresponding to a point $P$ of coordinates ($x_c$, $y_c$, $z_c$) in the camera frame.

The camera model describes the mathematical relationship between the coordinates of a point in three-dimensional space and its projection onto the image plane. Figure 3.1 shows the geometric model of a generic camera.

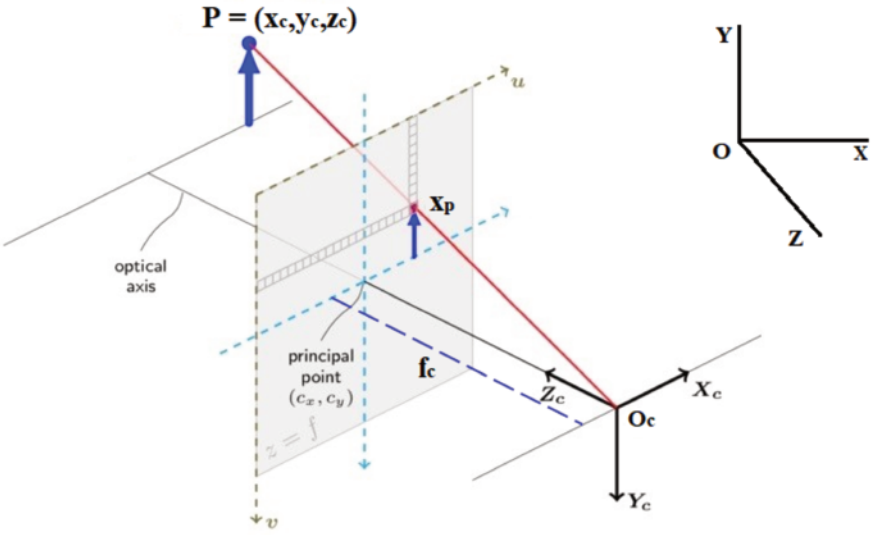



Figure 3.1: Geometric model of a camera.

This model transforms any 3D point of the space into an image point and can be performed by two successive elementary transformations. The first transformation is between the camera frame and the image coordinate system. The second transformation is between the world frame, arbitrarily chosen, and the camera frame, centered to the optical center of the camera.

The first transformation is performed using the projective geometry, that establishes the geometric relationship between a point in the camera frame and its corresponding position in the 2D image. This transformation is described by a matrix $3 \times 4$.

Any point in space $P$ with coordinates ($x_c$, $y_c$, $z_c$) expressed in the coordinate system ($O_c$, $X_c$, $Y_c$, $Z_c$) is projected in the image plane at a point $x_p$ with pixel coordinates ($u$, $v$)

according to the equation:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_c/s_u & \gamma & u_0 & 0 \\ 0 & f_c/s_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}, \tag{3.1}$$

A pixel has an horizontal dimension along the u direction ($s_u$) and a vertical dimension along the v direction ($s_v$). $f_c/s_u$ and $f_c/s_v$ represent the focal length expressed in pixels, on the direction $u$ and $v$, respectively. $\gamma$ represents the skew coefficient between the axis $u$ and $v$ of the image plane.

The projection matrix is defined as:

$$I_m = \begin{bmatrix} \alpha_u & \gamma & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{3.2}$$

with:

$$\alpha_u = f_c/s_u \quad \alpha_v = f_c/s_v. \tag{3.3}$$

This matrix contains the intrinsic parameters of the camera which define its geometric model, and it is given by:

$$K_c = \begin{bmatrix} f_c/s_u & \gamma & u_0 \\ 0 & f_c/s_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.4}$$

The second transformation is a rigid transformation consisting of a rotation and a translation, in homogeneous coordinates; it links the camera frame to the scene frame. The parameters of this transformation are called extrinsic camera parameters. Considering that the point $P$ of coordinate ($x$, $y$, $z$) in the coordinate system of the scene is transformed in a point with coordinate ($x_c$, $y_c$, $z_c$) in the camera frame, the following relation can be written:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} R \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \mathbf{t}. \tag{3.5}$$

where $\mathbf{t}$ is the translation vector between the origin of the scene frame and the origin of the camera frame, and $R$ is the rotation matrix that allows to have the camera frame and the

scene frame oriented in the same manner. In homogeneous coordinates the Eq. 3.5 becomes:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} + \mathbf{t} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} T_{4\times4} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \tag{3.6}$$

The composition of the Eq. 3.1 and 3.6 creates the complete equation of the camera model, defined as:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = I_m T \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \tag{3.7}$$

that relates the coordinates of the point $P$ in the scene frame to that of its representative in the image.

Finally, the camera model is described by 5 intrinsic parameters ($\alpha_u$, $\alpha_v$, $\gamma$, $u_0$, $v_0$), 3 extrinsic parameters for the rotation, expressed in instantaneous rotation vectors (the 3 columns of the matrix $R$), and 3 extrinsic parameters for the translation (the 3 elements of the vector $\mathbf{t}$).

Nonlinear intrinsic parameters such as lens distortion are also important, but they cannot be included in the linear camera model described by the intrinsic parameter matrix. The distortion parameters are computed using a calibration algorithm and the distortion effect is removed using the parametric approach defined in [96]. This approach consists in modeling the distortion by enriching the camera model with additional terms (the model then becomes non-linear). The most important deviation is due to radial distortion phenomena ("curvature" of the lenses). Secondary effects are introduced by tangential distortions ("decentralization" of the components of a lens system and production defects). Thus the model is improved by adding corrective terms corresponding to these two types of distortions, radial and tangential.

Starting from the camera model, the effects of distortions can be modeled by a third transformation, noted $D$, connecting the “ideal” image coordinates $(u, v)$ to the “real” image coordinates $(u_r, v_r)$:

$$(u_r, v_r) = D(u, v) = (u, v) + \Delta(u, v) = (u, v) + \Delta_r(u, v) + \Delta_t(u, v) \tag{3.8}$$

where $\Delta_r$ and $\Delta_t$ are the radial distortion and the tangential distortion, respectively.

The distortion transformation $D$ is defined as:

$$\begin{bmatrix} u_r \\ v_r \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} (1 + r_1(u^2 + v^2) + r_2(u^2 + v^2)^2 + r_3(u^2 + v^2)^3 + \begin{bmatrix} d_1(3u^2 + v^2) + d_2uv \\ 2d_1uv + d_2(u^2 + 3v^2) \end{bmatrix} \tag{3.9}$$

where $r_1$, $r_2$ and $r_3$ are the 3 radial parameters, $d_1$ and $d_2$ are the 2 tangential parameters of the distortion function. These coefficients for the correction of the distortion expand and complete the set of intrinsic parameters of the standard model of a camera. Optical distortion is usually modeled as a transformation that occurs after the 3D coordinates are projected onto the image plane. After that, the intrinsic matrix applies a transformation similar to the image, translating physical coordinates on the image plane into pixel coordinates.

Given $P$, a point of the scene with $(x,\ y,\ z)$ as coordinates, the camera model can be written as a non-linear vector function $F$:

$$(u_r, v_r) = F(I_m, T, D, P). \tag{3.10}$$

Due to the non-linearity of $D$, the function $F$ is not invertible, and it is necessary to use a non parametric method to estimate the function $D(u, v)$.

## 3.3 Estimation of the intrinsic parameters of a camera

The intrinsic parameters of a camera are estimated from established correspondences between the 3D coordinates of points on the plane $P_i$, with coordinates $(x_i,\ y_i,\ z_i)$ and the pixel coordinates of their projections in the camera image $(u_i,\ v_i)$. This is achieved using a test pattern consisting of squares whose geometric characteristics and positions are precisely known. The parameters are obtained by determining the $H_c$ homography as:

$$H_c = \min_{H_c} \left| \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} - H_c \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \right|^2. \tag{3.11}$$

To do this, several steps are necessary. A normalization of the coordinates $P_i$ and $(u_i,\ v_i)$ is made in a first step. This standardization consists in applying a similarity transformation $S_1$ to the points $(u_i,\ v_i)$ and a transformation $S_2$ to the points $P_i$. Each similarity transformation consists of a translation, which moves the points so that their origin corresponds to their barycenter, and a scaling procedure which acts so that the average distance from the points to the origin is equal to $\sqrt{2}$. This scaling, called "isotropic" scaling, has the particularity of generating a mean point of coordinates $(1, 1, 1)^T$. By setting:

$$(\tilde{u}_i, \tilde{v}_i, \tilde{w}_i) = S_1(u_i, v_i) \tag{3.12}$$

and

$$\tilde{P}_i = S_2 P_i, \tag{3.13}$$

from Eq. 3.7, the following equation is obtained [95]:

$$\begin{bmatrix} 0 & -\tilde{w}_i\tilde{P}_i & \tilde{v}_i\tilde{P}_i \\ \tilde{w}_i\tilde{P}_i & 0 & -\tilde{u}_i\tilde{P}_i \end{bmatrix} \tilde{\mathbf{H}} = A\tilde{\mathbf{H}} = 0, \tag{3.14}$$

with:

$$\tilde{\mathbf{H}} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} & h_{21} & h_{22} & h_{23} & h_{24} & h_{31} & h_{32} & h_{33} & h_{34} \end{bmatrix}^T. \tag{3.15}$$

$h_{ab}$ is the element of the matrix $\tilde{H}_c$ where $a$ is the index of row and $b$ is the index of column (standardized projection matrix). $\tilde{\mathbf{H}}$ is therefore a column matrix composed of all the elements of $\tilde{H}_c$. The estimation of $\tilde{\mathbf{H}}$ makes it possible to estimate $\tilde{H}_c$.

From Eq. 3.14, applied to a minimum of twelve matches, it is possible to determine the elements of the matrix $\tilde{H}_c$. To optimize the performance of the estimation, a constraint on $\tilde{H}_c$ can be added such as the following [95]:

$$\|\tilde{H}_c\| = 1. \tag{3.16}$$

By fixing this constraint, a solution of the Eq. 3.14 can be obtained by decomposing the matrix $A$ in singular values; this method is called singular value decomposition (SVD), as presented in [97]. Thus if the matrix $A$ can be written as:

$$A = UDV^T, \tag{3.17}$$

with $D$ a diagonal matrix whose elements are arranged in descending order, then $\tilde{H}$ corresponds to the last column of $V$.

The obtained solution is an estimation of $\tilde{H}_c$ which must be improved by reducing the geometric error, i.e. by reducing the quadratic deviation $s_i$ between the points $(u_i, v_i)$ and their estimation

$$(\tilde{u}, \tilde{v}) = \tilde{H}_c P_i, \tag{3.18}$$

given by:

$$s_i = \sum_i \|(u_i, v_i) - (\tilde{u}_i, \tilde{v}_i)\|^2 \tag{3.19}$$

The iterative technique of Levenberg-Marquardt presented in [98] is used with as initial value the estimation of $\tilde{H}_c$ obtained through $\tilde{H}$. The matrix $\tilde{H}_c$ is that which verifies the following equation:

$$\tilde{H}_c = \min_{\tilde{H}_c} \sum_i \|(\tilde{u}_i, \tilde{v}_i) - \tilde{H}_c P_i\|^2. \tag{3.20}$$

The matrix $\tilde{H}_c$ is then denormalized and generates the matrix $H_c$ according to the formula:

$$H_c = S_1^{-1}\tilde{H}_c S_2. \tag{3.21}$$

The matrix $H_c$ is written in the form:

$$H_c = \begin{bmatrix} \mathbf{h}_c^1 & \mathbf{h}_c^2 & \mathbf{h}_c^3 \end{bmatrix}, \tag{3.22}$$

and it is possible to write [99]:

$$\begin{bmatrix} \mathbf{h}_c^1 & \mathbf{h}_c^2 & \mathbf{h}_c^3 \end{bmatrix} = \lambda K_c \tilde{T}, \tag{3.23}$$

where $\mathbf{h}_c^1$ is the column $i$ of the matrix $\tilde{H}_c$, $\lambda$ is a scale factor, $K_c$ is the intrinsic matrix. $\tilde{T}$ is the matrix defined from extrinsic camera parameters, as:

$$\tilde{T} = \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{bmatrix} \tag{3.24}$$

It is not necessary to estimate the matrix $\tilde{T}$ to calculate the intrinsic parameters contained in the matrix $K_c$. As the first two columns of $\tilde{T}$ are two columns of a rotation matrix, they are orthogonal. This property allows to get the following constraints from Eq. 3.23:

$$\mathbf{h}_c^{1^T} K_c^{-1^T} K_c^{-1} \mathbf{h}_c^{2^T} = 0 \tag{3.25}$$

$$\mathbf{h}_c^{1^T} K_c^{-1^T} K_c^{-1} \mathbf{h}_c^{1^T} = \mathbf{h}_c^{2^T} K_c^{-1^T} K_c^{-1} \mathbf{h}_c^{2^T} \tag{3.26}$$

To estimate the intrinsic parameters that constitute the calibration matrix $K_c$ from the matrix $H_c$, a new matrix $B$ is introduced as follows:

$$B = K_c^{-1^T} K_c^{-1}. \tag{3.27}$$

$B$ is a symmetrical matrix, thus a 6 elements vector $\mathbf{b}$, containing the different elements of $B$, it is used:

$$\mathbf{b} = \begin{bmatrix} B_{11} & B_{12} & B_{22} & B_{13} & B_{23} & B_{33} \end{bmatrix}^T. \tag{3.28}$$

With these notations of $\mathbf{b}$, it is possible to write the following equation:

$$\mathbf{h}_c^{i^T} B \mathbf{h}_c^{j^T} = \mathbf{w}_{ij}^T b \quad \forall i, j \in [1;3], \tag{3.29}$$

where $\mathbf{w}_{ij}$ is equal to:

$$\mathbf{w}_{ij} = \begin{bmatrix} h_{i1}h_{j1} & h_{i1}h_{j2} + h_{i2}h_{j1} & h_{i2}h_{j2} & h_{i3}h_{j1} + h_{i1}h_{j3} & h_{i3}h_{j2} + h_{i2}h_{j3} & h_{i3}h_{j3} \end{bmatrix}^T \tag{3.30}$$

Thanks to this writing, the Eq. 3.25 and 3.26 can be written in the form of a matrix relationship:

$$\begin{bmatrix} w_{12}^T \\ (w_{12} - w_{22})^T \end{bmatrix} \mathbf{b} = W\mathbf{b} = 0 \tag{3.31}$$

The values of the 6 intrinsic parameters ($u_0$, $v_0$, $\alpha_u$, $\alpha_v$, $\gamma$, $\lambda$) are determined by solving the Eq. 3.31. In order to obtain at least 6 different equations, it is necessary to carry out point correspondences $\{(u, v)_i \Leftrightarrow P_i\}$ on at least 3 images of the test pattern taken from the same camera.

A solution to vector $\mathbf{b}$ is then obtained in the same way as for the estimation of $\tilde{\mathbf{H}}$ thanks to

the decomposition of the matrix $W$ into singular values (SVD). The computed vector obtained allows to obtain the matrix $B$ to a $\lambda$ scale factor. It is possible to estimate the intrinsic parameters located in the matrix $K_c$ from the matrix $B$ using the following relationships [99]:

$$v_0 = \frac{B_{12}B_{13} - B_{11}B_{23}}{B_{12}B_{22} - B_{12}^2} \tag{3.32}$$

$$\lambda = B_{33} - \frac{B_{13}^2 + v_o(B_{12}B_{13} - B_{11}B_{13})}{B_{11}} \tag{3.33}$$

$$\alpha_u = \sqrt{\frac{\lambda}{B_{11}}} \tag{3.34}$$

$$\alpha_v = \sqrt{\frac{\lambda B_{11}}{B_{11}B_{22} - B_{12}^2}} \tag{3.35}$$

$$\gamma = -\frac{B_{12}\alpha_u^2\alpha_v}{\lambda} \tag{3.36}$$

$$u_0 = \frac{\gamma v_0}{\alpha_v} - \frac{B_{13}\alpha_u^2}{\lambda} \tag{3.37}$$

In this document the skew coefficient $\gamma$ is equal to 1 (means that the principal plane directions are perpendicular to each other), and the pixels in the image are squared of dimension 1, so:

$$\gamma = 1, \tag{3.38}$$

$$s_u = s_v = 1, \tag{3.39}$$

$$\alpha_u = f_c/s_u = f_c/s_v = \alpha_v = f_c. \tag{3.40}$$

Using the estimation of the homographs, Zhang's method [99] estimates the distortion coefficients $r_1$, $r_2$, $r_3$, $d_1$, and $d_2$. Expanding the Eq. 3.9 it is obtained the following system:

$$\begin{cases} u_r = u + r_1(u^2+v^2)u + r_2(u^2+v^2)^2u + r_3(u^2+v^2)^3u + d_1(3u^2+v^2) + d_2uv \\ v_r = u + r_1(u^2+v^2)v + r_2(u^2+v^2)^2v + r_3(u^2+v^2)^3v + 2d_1uv + d_2(u^2+3v^2) \end{cases} \tag{3.41}$$

where the unknown are the distortion coefficients. Having $m$ points in $n$ images, a linear system of *2mn* equations in 5 unknowns can be set, which can be solved with the least squares method.

## 3.4 Geometric model of a stereovision system

A stereovision system is composed of two cameras each defined by its intrinsic parameters. Figure 3.2 presents a graphic representation of a stereovision system with $O_L$ and $O_R$ the centers optics of the left and right cameras, respectively. $\overrightarrow{O_L Z_L}$ and $\overrightarrow{O_R Z_R}$ are their optical axis and $P$ is a point in space which is projected in $P_L$ in the left camera plane and in $P_R$ in the right camera plane. $P_L$ and $P_R$ are the three-dimensional coordinates of the projections of point $P$ expressed in the left and right camera frames, respectively.

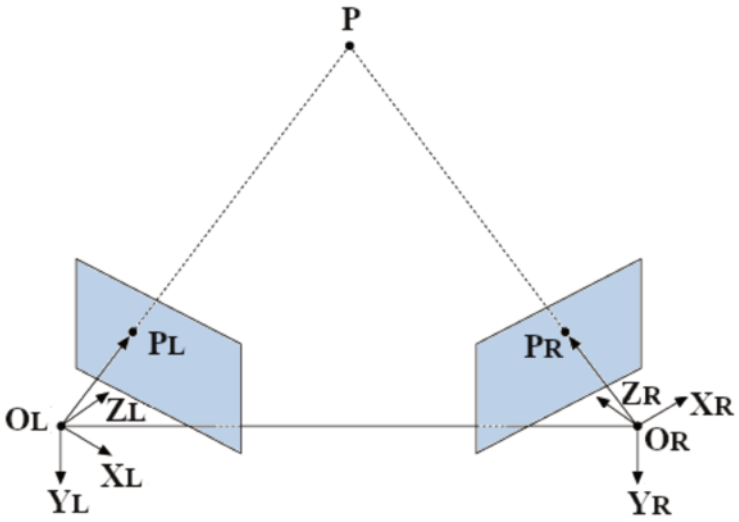



Figure 3.2: Stereovision system geometric model.

The relationship between the frames of two cameras is defined by a homogeneous transformation matrix. In this document, this transformation allows to pass from the "right camera frame" to the "left camera frame". It is defined by an homogeneous matrix as follow:

$$T_s = \begin{bmatrix} R_s & \mathbf{t}_s \\ 0_{1\times 3} & 1 \end{bmatrix}. \tag{3.42}$$

where $R_s$ is the rotation matrix to rotate the $X_r$, $Y_r$ and $Z_r$ axis of the right camera frame to be colinear to the $X_l$, $Y_l$ and $Z_l$ axis of the left camera frame, respectively; and $\mathbf{t}_s$ is the vector that represents the translation between the optical center $O_R$ of the right camera and the optical center $O_L$ of the left camera, expressed as:

$$\mathbf{t}_s = \begin{bmatrix} t_1 & t_2 & t_3 \end{bmatrix}^T \tag{3.43}$$

$\mathbf{t}_s$ and $R_s$ are called extrinsic parameters of the stereovision system.

The antisymmetric matrix of $\mathbf{t}_s$ is defined by:

$$[t_s]_\times = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \tag{3.44}$$

$K_L$ is the intrinsic matrix of the left camera, $T_L$ is the transformation matrix to express the 3D point $P$ in the left camera frame, $K_R$ is the intrinsic matrix of the right camera frame, and $T_R$ is the transformation matrix to express the 3D point $P$ in the right camera frame. In this phase, the cameras are considered without any distortion parameter.

The Eq. 3.7 can be expressed for the left camera and the right camera for any point $P$ in space as:

$$P_L = T_L P, \tag{3.45}$$

$$P_R = T_R P. \tag{3.46}$$

Using the Eq. 3.42 the relationship between $P_R$ and $P_L$ is:

$$P_R = T_s P_L. \tag{3.47}$$

Epipolar geometry is a mathematical model of geometry that describes the geometrical relationships of a stereoscopic pair of images (Fig. 3.3). The fundamental matrix, noted $F$, contains the necessary information of this geometry. It relates the coordinates ($x_L$, $y_L$, $z_L$) of $P_L$, in the left camera frame, to the coordinates ($x_R$, $y_R$, $z_R$) of $P_R$, in the right camera frame, using the relation called "epipolar constraint" which is expressed as:

$$\begin{bmatrix} x_R \\ y_R \\ z_R \end{bmatrix}^T F \begin{bmatrix} x_L \\ y_L \\ z_L \end{bmatrix} = 0. \tag{3.48}$$

By considering the calibration matrices $K_L$ and $K_R$ of the left and right cameras, the fundamental matrix is in the form as:

$$F = K_R^{-1^T} [t_s]_\times R_s K_L^{-1} \tag{3.49}$$

Moreover, at any point on the left image, with the coordinates ($u_L$, $v_L$), corresponds a set of possibilities of projection in the right image ($u_{R_i}$, $v_{R_i}$). This set is represented by the points of a straight line called epipolar line. The epipolar line is the straight line that passing through the point ($u_{L_i}$, $v_{R_i}$) and the epipolar point.

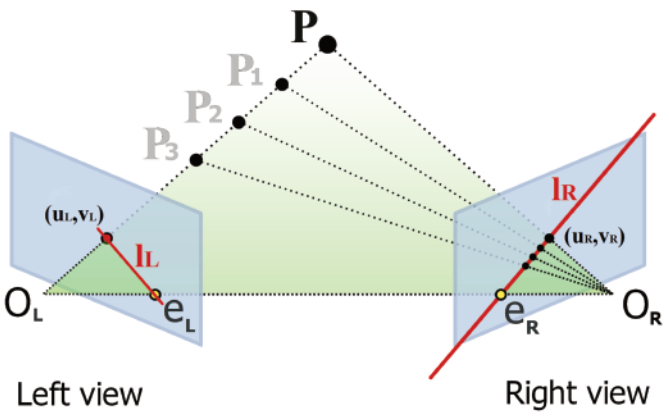



Figure 3.3: Essential elements of an epipolar geometry. $P_1$, $P_2$, and $P_3$ correspond to different 3D positions of the point P that produce the same projection in the letf image and different projections in the right image.

The point $P_L$ can be seen as the intersection of the straight line passing through $O_L$ and P, the optical axis of the left camera, with the image plane of the left camera. This line is projected on the image plane of the right camera. Due that from a single point of the 3D space corresponds a single pixel point in the image plane, the projective transformation is a homography; so if the optical axis connects P to $O_L$ then its projection connects the projections of these two points. The straight line passing through $P_R$ and the right epipolar point $e_R$, the projection of $O_L$ in the right camera, is called "right epipolar line". The same reasoning can be done similarly with the optical axis of the right camera.

The centers of each camera $O_L$ and $O_R$ are projected one onto the other image plane. Applying the Eq. 3.7 to the points $O_L$ and $O_R$, are obtained:

$$e_L = I_{m_L} T_L O_R, \tag{3.50}$$

$$e_R = I_{m_R} T_R O_L. \tag{3.51}$$

$e_L$ and $e_R$ are called the "epipolar points". The epipole of an image plane associated to a camera (noted $e_L$ for the left camera and $r_R$ for the right camera) is the point of intersection of all the epipolar lines of this plane. The epipoles $e_L$ and $e_R$ verify the following property:

$$F e_L = e_R^T F = 0 \tag{3.52}$$

The right epipolar line, noted $l_R$, is written as:

$$l_R = F \begin{bmatrix} u_L \\ v_L \\ 1 \end{bmatrix} \tag{3.53}$$

The epipolar line $l_L$ of a point $(u_L, v_L)$ in the left image is defined as:

$$l_L = F^T \begin{bmatrix} u_R \\ v_R \\ 1 \end{bmatrix} \tag{3.54}$$

Considering the relationships as presented in [95]:

$$\begin{bmatrix} x_R \\ y_R \\ z_R \end{bmatrix}^T K_R^T F K_L \begin{bmatrix} x_L \\ y_L \\ z_L \end{bmatrix} = 0 \tag{3.55}$$

The epipolar constraint is simplified as:

$$\begin{bmatrix} u_R \\ v_R \\ 1 \end{bmatrix}^T E \begin{bmatrix} u_L \\ v_L \\ 1 \end{bmatrix} = 0, \tag{3.56}$$

where:

$$E = [t_s]_\times Rs \tag{3.57}$$

is the essential matrix.

Finally, the calibration of a stereovision system makes it possible to estimate the matrix $R_s$ and the vector $t_s$ from which the essential matrix $E$ is constructed. This matrix relates the image coordinates of the two pixels from the same point in space $P$ and located in the two images of the stereovision system.

## 3.5 Estimation of the essential matrix

The estimation of the essential matrix is done mainly by two methods. The first one, developed by Longuet-Higgins in [100] computes this matrix from 8 matches of points in stereoscopic images. The second method, developed by Nistér in [101], allows to estimate the essential matrix from only 5 matches. This part describes the solution proposed by Nistér. Eq. 3.56 is rewritten as:

$$\tilde{\mathbf{x}}^T \tilde{\mathbf{E}} = 0, \tag{3.58}$$

where $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{E}}$ are equal to:

$$\tilde{\mathbf{x}} = \begin{bmatrix} u_L u_R & v_L u_R & u_R & u_L v_R & v_L v_R & v_R & u_L & v_L & 1 \end{bmatrix}^T, \tag{3.59}$$

$$\tilde{\mathbf{E}} = \begin{bmatrix} E_{11} & E_{12} & E_{13} & E_{21} & E_{22} & E_{23} & E_{31} & E_{32} & E_{33} \end{bmatrix}^T \tag{3.60}$$

From the five pairs of matched points, five vectors $\tilde{x}$ are constructed and concatenated to create a matrix of dimension $5 \times 9$. This matrix is decomposed into the form of the product of an orthogonal matrix and a superior triangular matrix (QR decomposition). The last 4 lines of the orthogonal matrix are reshaped into 4 matrices $X_1$, $X_2$, $X_3$ and $X_4$ of dimension $3 \times 3$.

The essential matrix can be expressed from these 4 matrices as:

$$E = x_1 X_1 + x_2 X_2 + x_3 X_3 + X_4, \tag{3.61}$$

where $x_1$, $x_2$ and $x_3$ are the scalars to be determined.

Considering the following constraints [95, 102]:

$$det(E) = 0, \tag{3.62}$$

$$EE^T E = \frac{1}{2} tr(EE^T)E = 0. \tag{3.63}$$

Using the Eq. 3.61, and performing a Gauss-Jordan elimination, it is obtained a system of 10 equations in $x_1$, $x_2$ and $x_3$. The rearrangement of these equations needs to obtain a matrix $B$ of dimension $3 \times 3$ containing polynomials in $x_3$. As the vector $\begin{bmatrix} x & y & 1 \end{bmatrix}^T$ is a null vector of $B$, the values that produce the determinant of $B$ equal to 0 (which is a polynomial of 10 degrees) are estimated. It is noted:

$$< n > \equiv det(B) \tag{3.64}$$

After a normalization of $<n>$ so that $n_{10} = 1$, the roots are obtained by searching the eigenvalues of the following canonical matrix:

$$\begin{bmatrix} -n_9 & -n_8 & \dots & -n_0 \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix} \tag{3.65}$$

For each $x_3$ root, the variables $x_1$ and $x_3$ can be obtained using the equation system defined by $B$. The essential matrix is then obtained from Eq.3.61.

From the matrix $E$, it is then possible to determine $R_s$ and $t_s$ by considering the theorem presented in [95, 97] and presented below:

**Theorem 3.1**
*Let E be a matrix decomposed into singular values and*

$$E = U diag(1, 1, 0) V^T \tag{3.66}$$

*with U and V matrices chosen so that their determinant is strictly positive, then*

$$\boldsymbol{t}_s = \pm \begin{bmatrix} u_{13} & u_{23} & u_{33} \end{bmatrix} \tag{3.67}$$

*and*

$$R_s = UDV^T \doteq UD^TV^T \tag{3.68}$$

*where:*

$$D = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.69}$$

The matrix $T_s$ defined in Eq. 3.42 has 4 possible solutions which satisfy the epipolar constraint of the Eq. 3.56, and one of these solutions corresponds to the real matrix $T_s$. The scene points must be in front of the cameras to find the correct estimation of $T_s$. A single point match is sufficient to determine the solution. It only needs to calculate the three-dimensional triangulation of the point $P$ using each of the 4 possible solutions and checking that $P$ is in the field of view of the two cameras.

This method computes the essential matrix from 5 matching points. This same computation can be done with a larger number of points so to improve the precision of the estimation of the matrix $E$.

## 3.6 Correspondence between points

3D points are obtained by stereovision using information of points matched in the left and right stereoscopic images.

The simplest method for the matching of points is called 2D search (or sparse search). This method searches for each point of interest into an image (for example the left image), its correspondence in any position of the other image (the right image). This search can take a long process time and generate false matches. One solution to this problem is to do a 1D type search. In this type of research, it is assumed that two corresponding points are located on the same line in the two stereoscopic images.

If the right camera is only offset horizontally compared to the left camera, and not rotated, then each pair of pixels that match are at the same vertical position in the two images. In general, even with high-precision equipment, it may be impractical to maintain perfect coplanarity between cameras.

Image rectification is a transformation process used to project images onto a common image plane. All the epipolar lines are then parallel to the horizontal axis and corresponding points have identical vertical coordinates. Any point must lie on the same horizontal line in each image plane, that allows to search matching points only in 1D.

### 3.6.1 Rectification process

In rectified images two corresponding points have identical vertical coordinates. In order to transform the original image pair into a rectified image pair, it is necessary to find a projective transformation $H$ that constrains the epipolar lines to be parallel with the horizontal axis. After this transformation the epipole points $e_L$ and $e_R$ are mapped to the infinite points $e_{L_\infty}$ and $e_{R_\infty}$, respectively (Fig. 3.4). The transformation $H$ is computed as composition of two elementary transformations, $H_1$ and $H_2$. $H_1$ is the transformation that rotates both images to look perpendicular to the line joining their collective optical centers. $H_2$ is the transformation that twists the optical axes so the horizontal axis of each image plane directs to the other image's optical center.

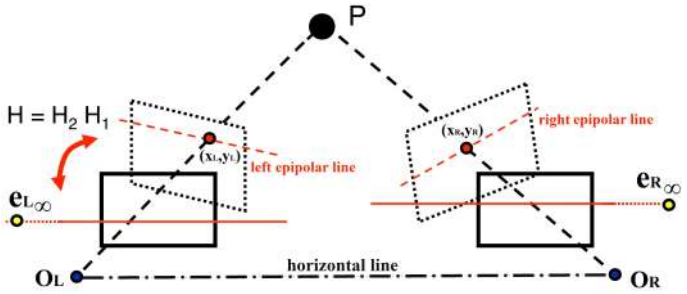



Figure 3.4: Stereovision model with rectified image planes (the dotted rectangles are the original planes and the black line ones are the rectified image plane).

It is assumed that the left camera frame is the reference world frame, so the transformation matrix $T_L$ of the left camera frame is equal to:

$$T_L = \begin{bmatrix} I_{3\times3} & 0_{3\times1} \\ 0_{1\times3} & 1 \end{bmatrix}, \tag{3.70}$$

The epipole points $e_{L_\infty}$ is computed as:

$$e_{L_\infty} = I_{m_L} T_L \begin{bmatrix} O_R \\ 1 \end{bmatrix}, \tag{3.71}$$

where the origin of the right camera frame $O_R$ is expressed in the left camera frame by the

combination of the rotation matrix $R_S$ and the translation vector $t_S$ of the Eq. 3.42:

$$O_R = R_S t_S. \tag{3.72}$$

Using the Eq. 3.70 and 3.72, the Eq. 3.71 becomes:

$$e_{L_\infty} = I_{m_L} T_L \begin{bmatrix} R_S t_S \\ 1 \end{bmatrix} = I_{m_L} R_S t_S, \tag{3.73}$$

The epipole point $e_{R_\infty}$ is computed by:

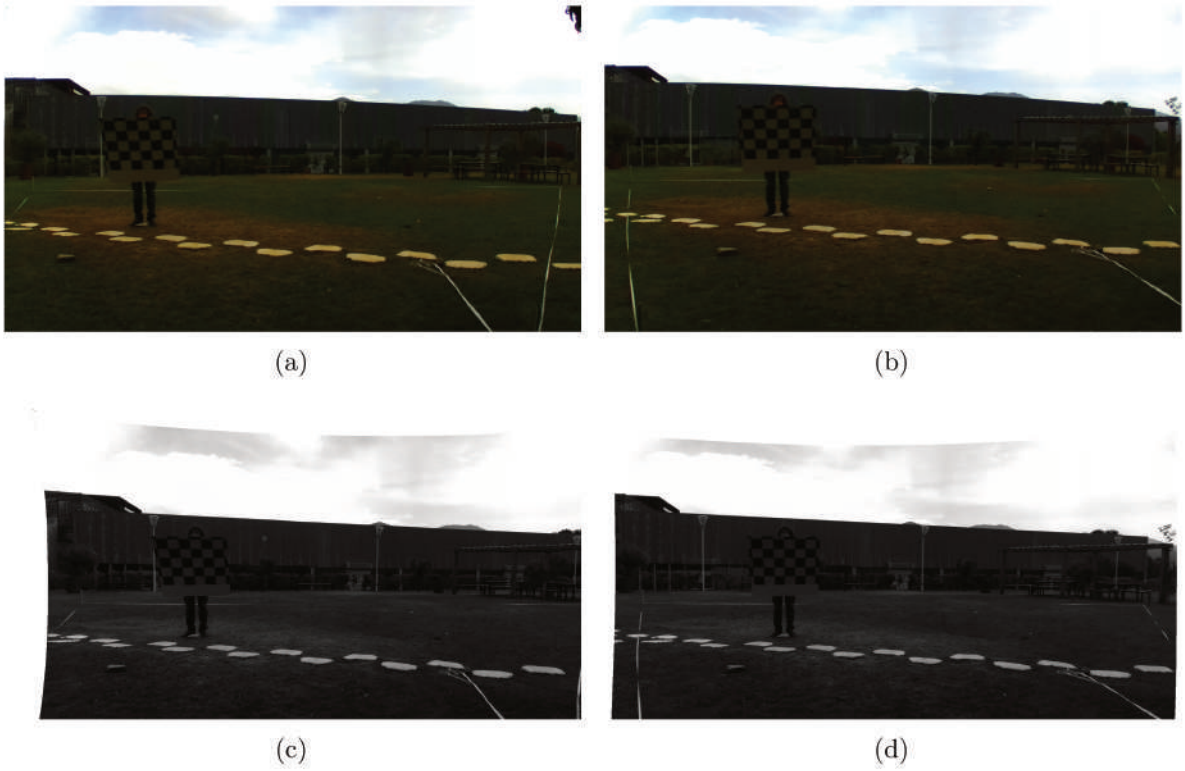
$$e_{R_\infty} = I_{m_R} T_R \begin{bmatrix} O_L \\ 1 \end{bmatrix}, \tag{3.74}$$

Substituting $T_L$ as the Eq. 3.95, the Eq. 3.75 becomes:

$$e_{R_\infty} = I_{m_R} T_R \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = I_{m_R} \begin{bmatrix} t_{1_R} \\ t_{2_R} \\ t_{3_R} \\ 1 \end{bmatrix}, \tag{3.75}$$

Szeliski in [103] presents a method to compute $H_1$ and $H_2$ using the epipolar line left and right as computed in the Eq. 3.50 and 3.51, the epipolar points $e_{L_\infty}$ and $e_{R_\infty}$, and the origin points $O_L$ and $O_R$ of the camera frames.

The left rectified epipolar line is the line parallel to the horizontal line and passing through the points $e_{L_\infty}$. The horizontal line is the line passing thought the points $O_L$ and $O_R$. Knowing all the points, it is possible to compute the director parameters of these lines. Using these director parameters, it is possible to find the rotation matrix to rotate the left image plane to be superposed to the left rectified image plane. Through an equivalent operation, it is possible to find the matrix *H'* that rectifies the right image. Figure 3.5 shows two original images (Fig. **??** and **??**), and the corresponding rectified images (Fig. **??** and **??**); it is possible to note that the vertical position of the checkboard in the original images is different, while in the rectified images is in the same vertical coordinates.



(a) (b)

(c) (d)

Figure 3.5: Original images and rectified images. ($a$) Left original image. ($b$) Right original image.($c$) Left rectified image. ($d$) Right rectified image.

### 3.6.2 Feature detection and matching points algorithms

The matching point algorithm is based on comparing and analyzing point correspondences between the two images, left and right. The evaluation of each 3D point requires the identifying, in the two stereoscopic images, of the pair of corresponding left and right points originating from the projection of the same 3D point, with pixel coordinates $(u_{L_i}, v_{L_i})$ in the left image, and $(u_{R_i}, v_{R_i})$ in the right image. When pairing, the most likely match for each point in the left image is searched in the right image. This process is performed into 3 steps: the detection of the points of interest, the descriptor extraction of the points of interest, and the matching of the points of interest by comparison of their descriptors.

#### 3.6.2.1 Detection of features

A point of interest, also called feature, in an image corresponds to a double discontinuity of the intensity function. They can be, for example, line junctions corresponding to textures, object corners, isolated points, or connected regions. There are many methods of detecting

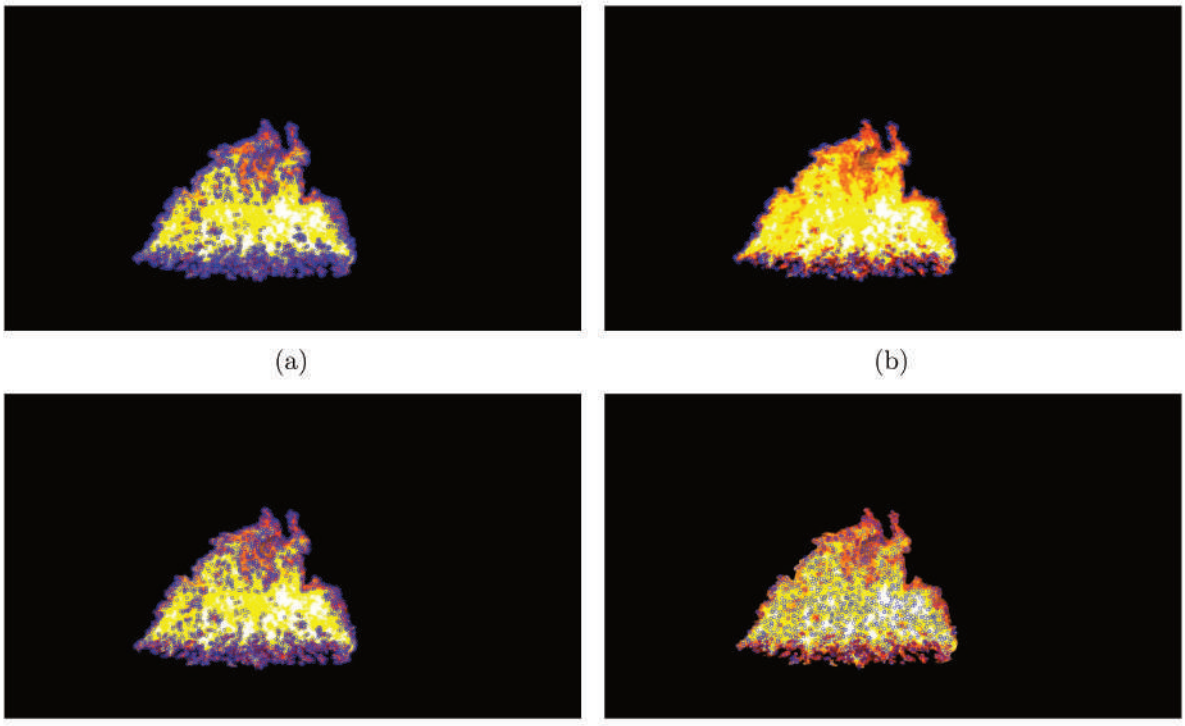
points of interest in an image; the best known are [104] the Harris method [105], the FAST algorithm (Feature from Accelerated Segment Test) [106], the SURF algorithm (Speeded Up Robust Features) [107], the MSER algorithm (Maximally Stable Extremal Regions) [108], and the detection by the minimum eigenvalue method [109].

The Harris method calculates the second order matrix obtained from the neighboring pixels of the pixel considered. The FAST algorithm is based on the comparison between the gray level of the current point and that of certain points close to the treated point The SURF algorithm uses an integer approximation of the determinant of Hessian, and it is based on the sum of the Haar wavelet response around the point of interest. The MSER algorithm extracts from an image a number of co-variant regions, defined by a stable connected component of some gray-level sets of the image. The detection by the minimum eigenvalue method calculates the minimum of the eigenvalues of the Harris matrix.

In [110] all these methods are compared, the result is that the SURF method is the fastest method, but the number of detected points is small. The methods that detect a significant number of points are the Harris method and the Eigen method, with the Harris algorithm faster than the Eigen algorithm.

Harris, FAST, SURF, MSER, and Eigen methods were tested alone and by combination on the wildfire images used in this thesis. The best result was achieved by performing a detection of Harris, followed by the SURF method. This procedure is simple and fast.

Figure 3.6 shows an example of features detected in an wildfire image by the FAST, Harris, Eigen and SURF algorithms; as the MSER method returns feature regions, is not considered in this work. The input image is smoothed by a Gaussian kernel in a scale-space representation.



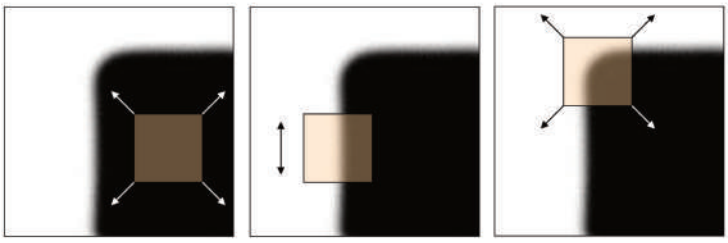
(a) (b)

(c) (d)

Figure 3.6: Comparison Harris, FAST, Eigen and SURF algorithms. (*a*) Harris (2 829 feature points). (*b*) FAST (664 points). (*c*) Eigen (2 087 points). (*d*) SURF (1 152 points).

### Harris method

The Harris detector detects corners sliding a small window over the image. This can causes or not gradient changes in different directions (Fig. **??**).



(a) (b) (c)

Figure 3.7: Harris Corner Point Detector. (*a*) "Flat region", no change in all directions. (*b*) "Edge region", no change along the edge directions. (*c*) "Corner region", significant change in all directions.

This method consists in calculating the second order moment matrix, called the Harris

matrix, and allows the corners to appear by autocorrelation independently of the direction. This matrix, denoted $H_x$, is defined for each pixel $x$ of the image $I$ as follows:

$$H_x = \sigma_d^2 \begin{bmatrix} I_a(x) & I_c(x) \\ I_c(x) & I_b(x) \end{bmatrix} \tag{3.76}$$

with the images $I_a$, $I_b$ and $I_c$ calculated by:

$$I_a = g_{\sigma_i} \left( \frac{\partial g_{\sigma_i} I^2}{\partial l} \right) \tag{3.77}$$

$$I_b = g_{\sigma_i} \left( \frac{\partial g_{\sigma_i} I^2}{\partial c} \right) \tag{3.78}$$

$$I_c = g_{\sigma_i} \left( I_a I_b \right) \tag{3.79}$$

$g_{\sigma_d}$ and $g_{\sigma_i}$ are the Gaussian standard deviations of $\sigma_d$ and $\sigma_i$, $\partial/\partial l$ the derivative of an image on the rows and $\partial/\partial c$ the derivative of an image on the columns. As in [74], $\sigma_d$ is fixed equal to 1 and $\sigma_i$ is fixed equal to 1,5.

The force function of the points of interest $f_H$ is then defined as the ratio of the determinant of the Harris matrix on its trace [111]:

$$f_{H_l}(x) = \frac{\det H_l(x)}{tr H_l(x)} = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} \tag{3.80}$$

with $\lambda_1$ and $\lambda_2$ the eigenvalues of the matrix $H_x$.

The points of interest $x$ of an image have a value $f_{H_l}(x)$ greater than that of the other pixels of its environment. For each pixel of the image, a neighborhood block of size $n \times n$ pixels is considered and if the pixel corresponds to the local maximum of this neighborhood then it is considered as a point of interest.

**SURF method**

This approach uses a very basic Hessian matrix approximation, applied on the "integral image". The integral image is used as a quick and effective way of calculating the sum of values (pixel values) in a given image, or a rectangular subset of a grid of the given image. The entry of an integral image $I_\Sigma(x)$ at a location $\mathbf{x} = (u, v)^T$ represents the sum of all pixels in the input image $I$ within a rectangular region formed by the origin and $x$.

$$I_\Sigma(\mathbf{x}) = \sum_{i=0}^{i \leq u} \sum_{j=0}^{j \leq v} I(i, j) \tag{3.81}$$

With $I_\Sigma$ calculated, the algorithm only takes four additions to compute the sum of the intensities over any upright, rectangular area, independent of its size.

SURF uses the Hessian matrix because of its good performance in computation time and accuracy. Rather than using a different measure for selecting the location and the scale (Hessian-Laplace detector), SURF relies on the determinant of the Hessian matrix for both. Given a pixel, the Hessian of this pixel is:

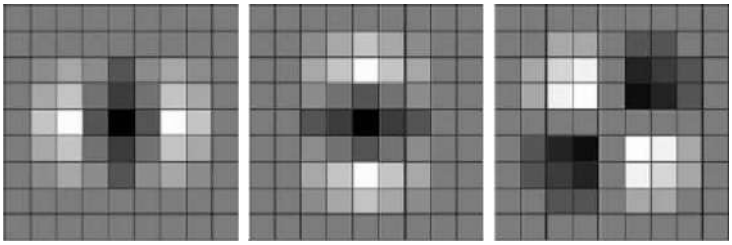
$$H(f(u,v)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial v} \\ \\ \frac{\partial^2 f}{\partial x \partial v} & \frac{\partial^2 f}{\partial v^2} \end{bmatrix}. \tag{3.82}$$

To adapt to any scale, SURF filters the image by a Gaussian kernel; so given a point $\mathbf{x} = (u, v)$, the Hessian matrix $H(\mathbf{x}, \sigma)$ in $\mathbf{x}$ at scale $\sigma$ is defined as:

$$H(\mathbf{x}, \sigma) = \begin{bmatrix} L_{uu}(\mathbf{x}, \sigma) & L_{uv}(\mathbf{x}, \sigma) \\ L_{uv}(\mathbf{x}, \sigma) & L_{vv}(\mathbf{x}, \sigma) \end{bmatrix}, \tag{3.83}$$

where $L_{uu}(\mathbf{x}, \sigma)$ is the convolution of the Gaussian second order derivative with the image $I$ in $\mathbf{x}$ along the horizontal direction, $L_{vv}(\mathbf{x}, \sigma)$ is the convolution of the Gaussian second order derivative along the vertical direction and $L_{uv}(\mathbf{x}, \sigma)$ is the convolution of the Gaussian second order derivative along the diagonal direction.

Figure 3.8 shows a box filter $9 \times 9$ of the Gaussian second order derivative, with $\sigma = 1.2$ [107].



(a) (b) (c)

Figure 3.8: Box filter $9 \times 9$ constructed with the Gaussian second order partial derivatives, with $\sigma = 1.2$. $(a)$ Horizontal direction of the derivative. $(b)$ Vertical direction of the derivative. $(c)$ Diagonal direction of the derivative.

Gaussian second order derivative are optimal for scale-space analysis but in practice, they have to be discretized and cropped. This leads to a loss in repeatability under image rotations around odd multiples of $\pi/4$. This weakness holds for Hessian-based detectors in general. Nevertheless, the detectors still perform well, and the slight decrease in performance does not outweigh the advantage of fast convolutions brought by the discretization and cropping.

In order to calculate the determinant of the Hessian matrix, it is applied a convolution with an approximation of the Gaussian kernel of the second order derivative, that can be evaluated at a very low computational cost using integral images and independently of size

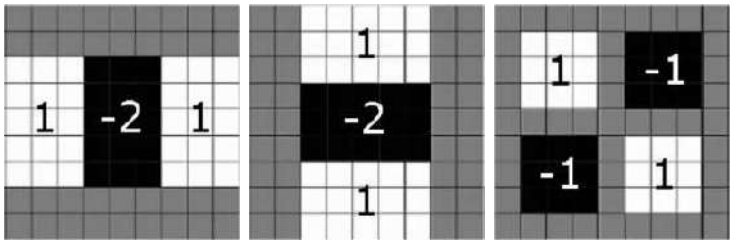
(this makes SURF a fast algorithm). $D_{uu}$ is the approximation of the Gaussian second order derivative along the horizontal direction, and similarly for $D_{vv}$ and $D_{uv}$, as:

$$D_{uu}(u,v) = G(u-1,v,\sigma) + G(u+1,v,\sigma) - 2G(u,v,\sigma) \tag{3.84}$$

$$D_{vv}(u,v) = G(u,v-1,\sigma) + G(u,v+1,\sigma) - 2G(u,v,\sigma) \tag{3.85}$$

$$\begin{aligned} D_{uv}(u,v) = {} & G(u-1,v-1,\sigma) + G(u+1,v+1,\sigma) - 2G(u,v,\sigma) \\ & -[G(u-1,v+1,\sigma) + G(u+1,v-1,\sigma) - 2G(u,v,\sigma)] \end{aligned} \tag{3.86}$$

Figure 3.9 shows these approximations for a box filter $9 \times 9$.




(a) (b) (c)

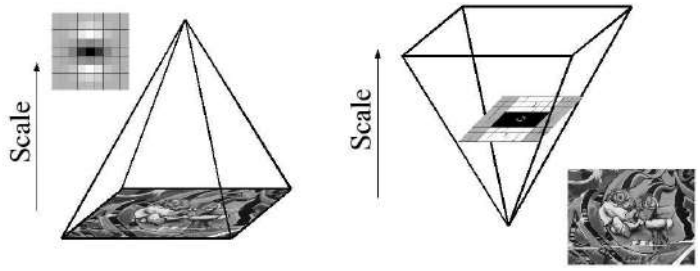
Figure 3.9: Box filter $9 \times 9$ constructed with the approximations of the Gaussian second order partial derivatives, with $\sigma = 1.2$. ($a$) Horizontal direction of the derivative. ($b$) Vertical direction of the derivative. ($c$) Diagonal direction of the derivative.

The approximated determinant of the Hessian is:

$$\det H_{approx} = D_{uu}D_{vv} - (wD_{uv})^2 \tag{3.87}$$

with $w = 0.9$ [107].

The method works on the images in scale spaces, implementing the image as an image pyramids. The images are repeatedly smoothed with a Gaussian and subsequently sub-sampled in order to achieve a higher level of the pyramid. Due to the use of box filters and integral images, SURF does not have to iteratively apply the same filter to the output of a previously filtered layer but instead can apply such filters of any size at exactly the same speed directly on the original image, and even in parallel. Therefore, the scale space is analyzed by up-scaling the filter size ($9 \times 9$, $15 \times 15$, $21 \times 21$, $27 \times 27$, etc) rather than iteratively reducing the image size. The filter size increases is doubled simultaneously the sampling intervals for the extraction of the interest points x can be doubled as well which allow the up-scaling of the filter at constant cost (Fig. 3.10).




(a) (b)

Figure 3.10: Scale-space representation. (*a*) The image size is iteratively reduced. (*b*) Integral images that allow the up-scaling of the filter at constant cost.

A point is compared with 26 points around it. These points are found into a $3 \times 3 \times 3$ neighborhood matrix, corresponding to the $3 \times 3$ box of the neighbors, for 3 scales. A detected point corresponds to an extreme point and has its determinant of Hessian matrix higher than a threshold.

#### 3.6.2.2 Points of interest descriptors

This step consists in describing the pixels surrounding the points of interest. Generally the descriptors are invariant to geometric and light transformations. Many descriptor extraction techniques were proposed in the literature: Scale-Invariant Feature Transform method (SIFT) [112], Gradient Location Orientation Histogram method (GLOH) [113], descriptor method based on filters [114], Linear Discriminant Analysis Hashing method (LDAHash) [115] or descriptors based on color [116]. In this thesis, a simple form of descriptors proposed by Brown *et al.* [74] is used.

These descriptors are in the form of $8 \times 8$ matrices obtained from the intensity of the pixels located into a window where the point of interest is the top left corner. Each matrix is obtained by first considering an area of $41 \times 41$ pixels, and then by sampling it in order to work on a grouping of $8 \times 8$ pixels. This sampling is equivalent to leaving four pixels between each selected pixel. To avoid aliasing effects, the pixels are sampled by interpolation, that is to say by making a weighted average of the neighboring pixels. In order to be robust to changes in brightness between the images containing the points to be matched, the block of $8 \times 8$ pixels named $B$ is normalized as follows:

$$\overline{\mathbf{x}^j} = \frac{\mathbf{x}^j - \mu_B}{\sigma_B} \tag{3.88}$$

With $\mathbf{x}^j$ a pixel belonging to the block of pixels $B$, $\mu_B$ and $\sigma_B$ the mean and the standard deviation of the pixels intensity of the block $B$. Figure 3.11 shows an example of feature description.

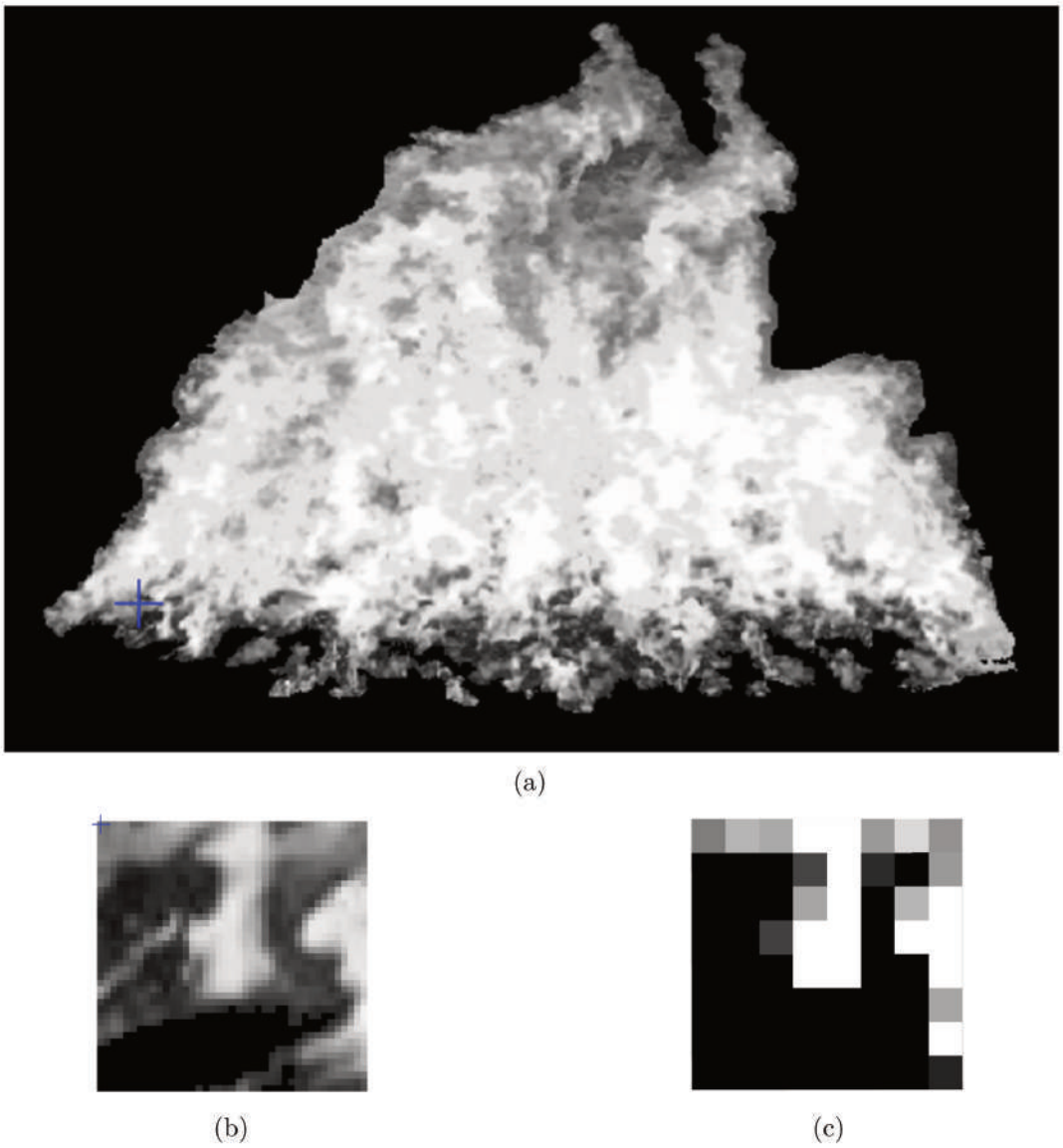


Figure 3.11: Example of a feature descriptor box for one point of interest. ($a$) Example of a point of interest (blue cross). ($b$) $41 \times 41$ descriptor box around the point of interest. ($c$) $8 \times 8$ descriptor box around the point of interest.

#### 3.6.2.3 Matching points of interest by comparison of their descriptor

The previous steps make it possible to obtain points of interest in each stereoscopic image with their associated descriptor.

The matching method consist to find into the two images (left and right) if there are points in common. The method verify for each feature in the left image if there is a corresponding feature into the right image. Thus, this method requires a measure of similarity of their

descriptor. A large number of similarity measurement methods was listed by Chambon and Crouzil [117]. Among the methods with a good ratio of execution time to number of good matches, the method called "Zero mean Normalized Sum of Squared Differences" (ZNSSD) was chosen for this thesis. The correlation coefficient which measures the similarity between two descriptors is defined by:

$$ZNSSD(D_1, D_2) = \frac{\|(D_1 - \overline{D_1}) - (D_2 - \overline{D_2}\|^2}{\sqrt{\|D_1\|^2 \|D_2\|^2}}, \tag{3.89}$$

with $D_1$ and $D_2$ the descriptors to compare, $\overline{D_1}$ and $\overline{D_2}$ the average of the values of the descriptors $D_1$ and $D_2$, respectively.

The smaller is the ZNSSD, the greater is the similarity between $D_1$ and $D_2$.

In the rectified image, if a feature position is on the line $i$ of the left image, and if exists a corresponding point in the right image, the interest point in the right image has to be on the same line $i$. This constraint is motivated because to avoid error of the rectification process, due to the not perfect estimation of the camera parameters. Figure 3.12 show an example of a pair of original images and a pair of rectified images, where it is possible to note that the $Y$ coordinates of the selected pixel, different in the original images, is the same in the rectified images.

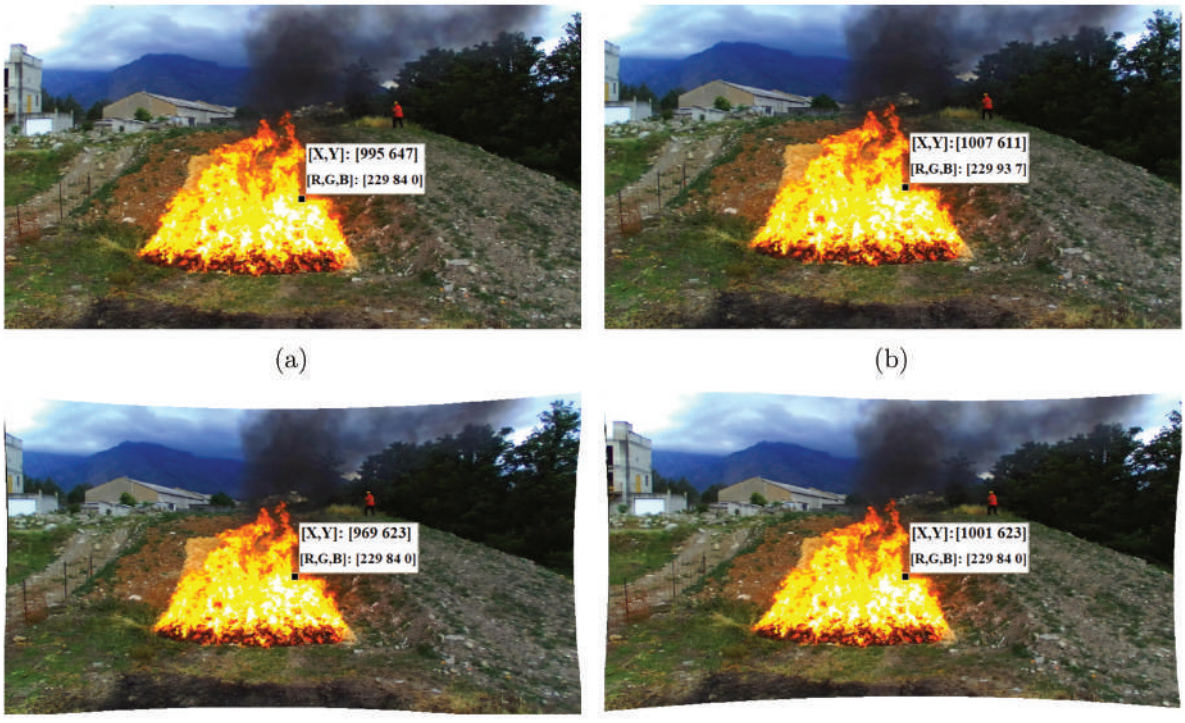
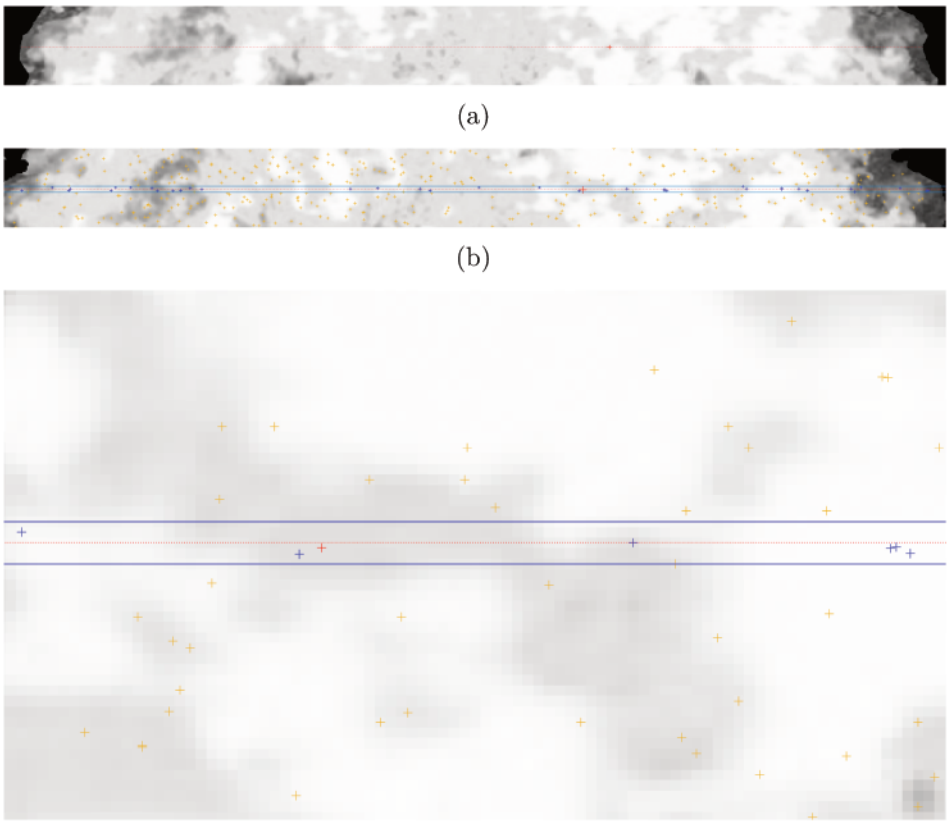



Figure 3.12: Original and rectified fire images. ($a$) Original left image. ($b$) Original right image. ($c$) Rectified left image. ($d$) Rectified right image.

The matching procedure is carried out in the following manner: for each feature in an image, a procedure is done in the other stereoscopic image by searching candidates along the epipolar line which has the same y-coordinate as the feature. A vertical tolerance of $\mp 2$ pixels is applied. The pixel that obtains the smallest ZNSSD coefficient is paired.

Figure 3.13 shows an epipolar line, associated to a point of interest in the left image, and the candidate points for pairing in the right image.



(a)

(b)

(c)

Figure 3.13: Example of points matched in a fire image of the visible domain. ($a$) Left image: the red cross represents the point to match, the dotted line is the left epipolar line. ($b$) Right image: the dotted red line is the right epipolar line, the blue lines delimit the region to search candidate points to match, the yellow crosses are the points too far from the epipolar line, the blue points are the points at a $\mp$ 2px of distance from the epipolar line (37 candidate points to match), the red cross is the matched point. ($c$) Zoom of the right image.

## 3.7 3D reconstruction of matching points

The three-dimensional reconstruction of matching points is carried out by a so-called triangulation method as presented in [95].

**Definition 3.1**
*Let $M_L$ be the projection matrix of the left camera and $M_R$ the projection matrix of the right camera, defined as the Eq. 3.7:*

$$M_L = I_{m_L} T_L, \tag{3.90}$$

$$M_R = I_{m_R} T_R, \tag{3.91}$$

*then $\forall P \in R^3$ of coordinates ($x_p$, $y_p$, $z_p$):*

$$\begin{bmatrix} u_L \\ v_L \\ 1 \end{bmatrix} = M_L \begin{bmatrix} x_p \\ y_p \\ z_p \\ 1 \end{bmatrix} \tag{3.92}$$

$$\begin{bmatrix} u_R \\ v_R \\ 1 \end{bmatrix} = M_R \begin{bmatrix} x_p \\ y_p \\ z_p \\ 1 \end{bmatrix} \tag{3.93}$$

Expressing the matrix $T_L$ and $T_R$ as:

$$T_L = \begin{bmatrix} r_{11_L} & r_{12_L} & r_{13_L} & t_{1_L} \\ r_{21_L} & r_{22_L} & r_{23_L} & t_{2_L} \\ r_{31_L} & r_{32_L} & r_{33_L} & t_{3_L} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3.94}$$

$$T_R = \begin{bmatrix} r_{11_R} & r_{12_R} & r_{13_R} & t_{1_R} \\ r_{21_R} & r_{22_R} & r_{23_R} & t_{2_R} \\ r_{31_R} & r_{32_R} & r_{33_R} & t_{3_R} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3.95}$$

the Eq. 3.92 and 3.93 can be rewrited as a system of four equation in three unknown:

$$\begin{cases} u_L = f_L \frac{r_{11_L} X + r_{12_L} Y + r_{13_L} Z + t_{1_L}}{r_{31_L} X + r_{32_L} Y + r_{33_L} Z + t_{3_L}} + c_{x_L} \\ v_L = f_L \frac{r_{21_L} X + r_{22_L} Y + r_{23_L} Z + t_{2_L}}{r_{31_L} X + r_{32_L} Y + r_{33_L} Z + t_{3_L}} + c_{y_L} \\ u_R = f_R \frac{r_{11_R} X + r_{12_R} Y + r_{13_R} Z + t_{1_R}}{r_{31_R} X + r_{32_R} Y + r_{33_R} Z + t_{3_R}} + c_{x_R} \\ v_R = f_R \frac{r_{21_R} X + r_{22_R} Y + r_{23_R} Z + t_{2_R}}{r_{31_R} X + r_{32_R} Y + r_{33_R} Z + t_{3_R}} + c_{y_R} \end{cases} \tag{3.96}$$

where $f_L$ is the focal length of the left camera, ($c_{x_L}$, $c_{y_L}$) is the principal point of the left camera, $f_R$ is the focal length of the right camera, and ($c_{x_R}$, $c_{y_R}$) is the principal point of the right camera.

The equations presented in 3.96 give a system of 4 equations. If ($u_L$, $v_L$), ($u_R$, $v_R$), $M_L$ and $M_R$ are known, the three-dimensional coordinates ($x_p$, $y_p$, $z_p$) of the point $P$ are obtained by the resolution of the overdetermined equation system (3 unknowns in 4 equations).

In this thesis, the intrinsic and extrinsic parameters of the stereovision system are estimated by calibration. The three-dimensional coordinates of $P$ in the left camera frame, ($x_L$, $y_L$, $z_L$), can be computed using the Eq. 3.92-3.93 where $M_L$ and $M_R$ are equal to:

$$M_L = I_{m_L} \tag{3.97}$$

$$M_R = I_{m_R} T_s \tag{3.98}$$

where $I_{m_L}$ and $I_{m_R}$ are the calibration matrices of the left and right cameras and $T_s$ is the transformation matrix from the right camera frame to the left camera frame.

When the cameras are not calibrated, it is possible to compute the projection matrices from the fundamental matrix $F$ as it is described in [95]:

$$M_L = \begin{bmatrix} Id_{3\times3} & 0_{3\times1} \end{bmatrix} \tag{3.99}$$

$$M_R = \begin{bmatrix} [e_R]_\times F & e_R \end{bmatrix} \tag{3.100}$$

where $I_d$ is the identity matrix, $e_R$ is the epipole of the right image, and $[e_R]_\times$ is the antisymmetric matrix of $e_R$.

Chapter 4

# Measurement of wildfire geometrical characteristics by multimodal stereoscopic images acquired by drone

## Contents

## 4.1 Introduction

This chapter presents the algorithms for the measurement of fire geometrical characteristics from 3D points obtained from the images of a multimodal (visible - LWIR) stereoscopic system carried by a drone. The following data are thus calculated: the position on the ground of the fire, its rate of spread, local direction, width, height, length, inclination angle, 3D shape and surface. In addition, from the fire points that are on the ground, the area of the fuel that burns is estimated.

Section 4.2 describes the calibration procedure used for the estimation of intrinsic and extrinsic parameters of the vision system considered in this thesis.

Section 4.3 explains the multimodal fire detection procedure used in this work.

Section 4.4 explains the obtaining of the 3D fire points.

Section 4.5 describes the method to project the 3D fire points into several frames to simplify the fire geometrical estimation.

Section 4.6 presents the methodology to measure the geometrical characteristics from the fire transformed 3D points.

## 4.2 Calibration of the vision system

The intrinsic and extrinsic parameters of the stereovision system are estimated using the methods presented in the previous chapter. Bouguet [118] developed a MATLAB toolbox containing the programs for all of these methods. Since its 2014b version, this toolbox is included into the the standard Matlab toolbox called Computer Vision System, to calibrate a vision systems from checkboard images (Fig. 4.1).

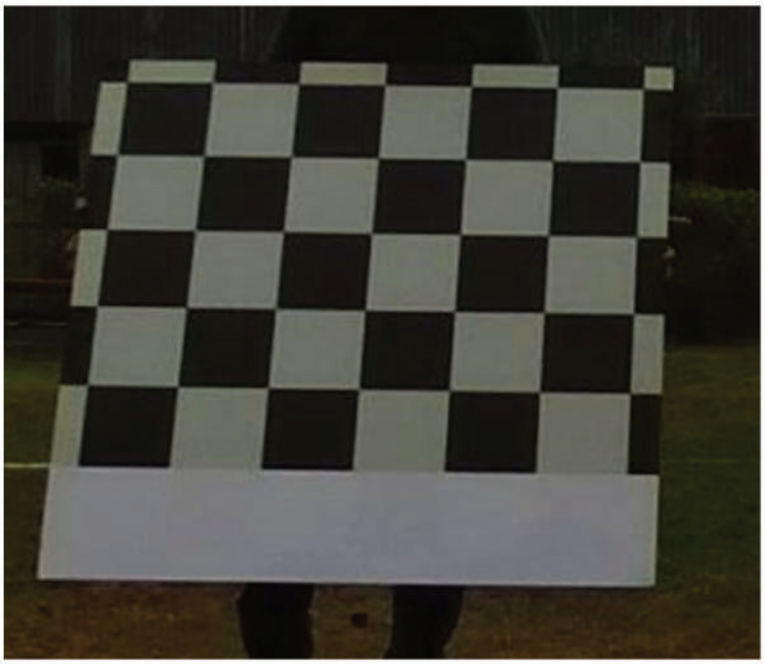


Figure 4.1: Checkboard used for the camera calibration.

This checkboard is a square of 1 m of side, composed by squares black and white of 15 cm of side. The number of squares of the checkboard is 6 along the width side and 5 along the height side. The number of squares must be different in each direction of the pattern in order to be rotation-invariant and avoid a 180° rotation ambiguity, which is important in the case of a stereo calibration.

During the calibration procedure, a series of 20 images of the checkboard with different positions, orientations and angulations is acquired. The number of 20 was chosen taking into account the advice of the Bouguet toolbox, as the right compromise between accuracy of the result and speed of execution. In each image are detected the corner location of the internal squares; in this case 42 points. The pattern is positioned at several places of the field of view of the cameras and between 10 m and 15 m in order to correctly calibrate the distortion function and to work at the same distance that the fire will be from the stereovision system (Fig. 4.2).

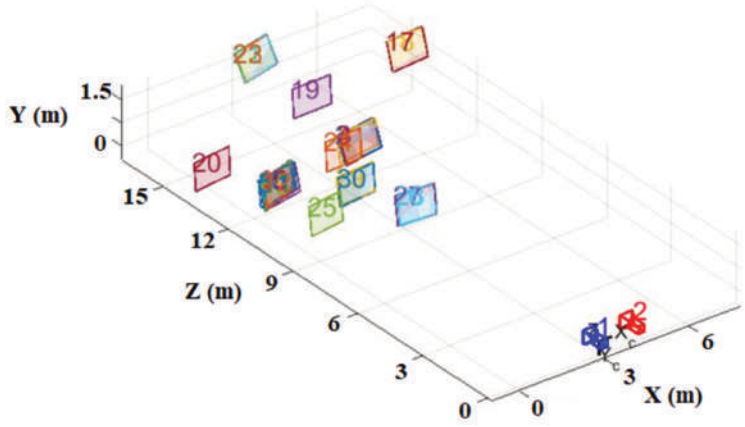


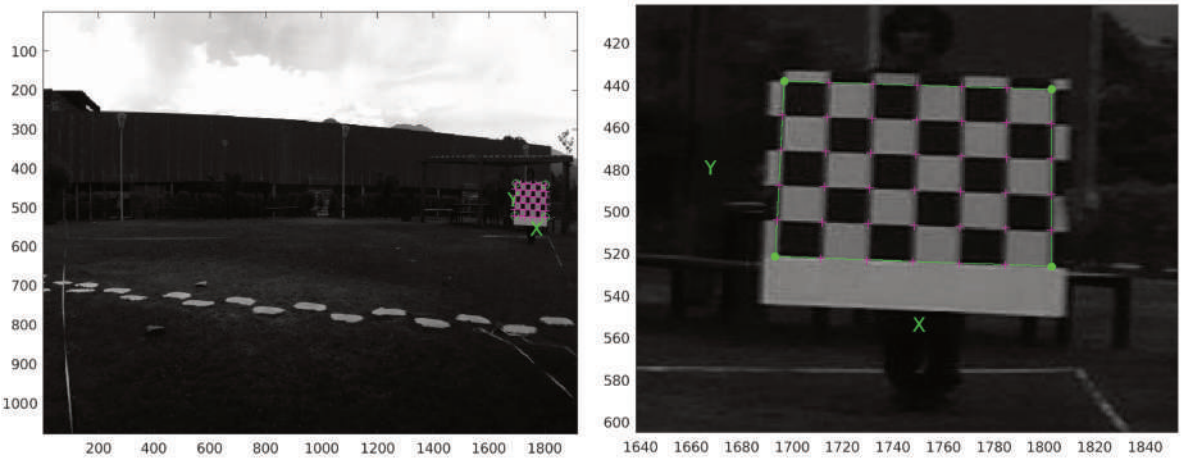

Figure 4.2: Example of 20 checkerboard positions in the World space.

The maximal distance camera-target of 15 m was used because beyond, the detection of corner points was poor Figure 4.3 shows an example of the checkboard at 12 m from the camera.

Figure 4.3: Image of the checkboard with a distance camera-target equal to 12 m.

### 4.2.1 Calibration of the visible stereovision system

At a given instant, two images of the checkboard, in the visible domain, are acquired simultaneously. For each image $i$ are detected the corner location of the 42 points $M_j^i$ (Fig. 4.4).




(a) (b)

Figure 4.4: Checkboard pattern recognized. (*a*) Checkboard pattern in the original image. (*b*) Zoom of the checkboard (the purple crosses are the identified corners, the green circles are the corners selected by hand).

Using the points $M_j^i$, the intrinsic and the extrinsic parameters are estimated using the procedure described in Ch. 3.

To validate the calibration method, for each image $i$, are used the parameters estimated to compute the projection points $\tilde{m_j^i}$ of the checkboard corner points in the space. These projections are compared with the corresponding starting corner location points $M_j^i$. The pixel error of the image $i$ is defined as:

$$pixel\ error_i = \frac{\sum_{j=1}^{42} \tilde{m}_j^i - M_j^i}{42} \tag{4.1}$$

Figure 4.5 shows an example of two pixels (the starting pixel and the reprojection pixel) with a sub-pixel resolution.

Figure 4.5: Starting pixel and reprojection pixel in sub-pixel resolution.

To compute the pixel error associated at one camera, Eq. 4.1 is computed for all the 20 image as:

$$error\ camera = \frac{\sum_{i=1}^{20} pixel\ error_i}{20} \tag{4.2}$$

Figure 4.6 shows the results of the pixel error for the 20 images captured with the two visible cameras, with an error pixel of 0,08.

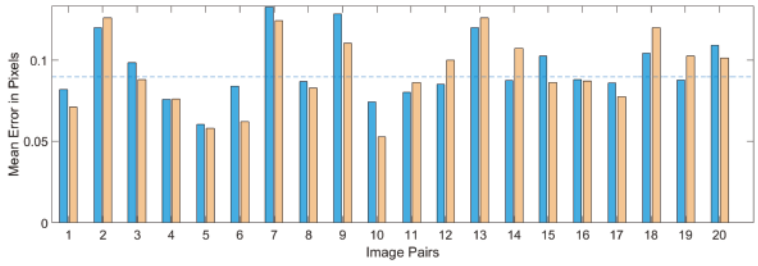



Figure 4.6: Mean reprojection error of the two visible cameras.

### 4.2.2 Calibration of the LWIR stereovision system

The procedure described previously and used to calibrate the visible cameras is also used to calibrate the LWIR cameras. The checkboard was modified in order to be visible in the LWIR spectral band. This modification consists in covering the white squares with aluminium (Fig. 4.7). The black squares have the ability to absorb part of the solar radiation, and the aluminum squares reflect part of solar radiation. This allows to identify the corner points in

an LWIR image. This modified checkboard was made by hand, and leads to a less accurate estimate of camera parameters than the one of the visible cameras. As the 3D fire points are obtained from features identified in the visible images, it was chosen to use the modified checkboard only to calibrate the LWIR cameras.

Figure 4.7: Image of the checkboard modified to be visible by the LWIR cameras.

During the calibration procedure of the LWIR cameras, their sensibility, exposition time and contrast were adjusted in order to obtain images with the best contrast (Fig. 4.8).

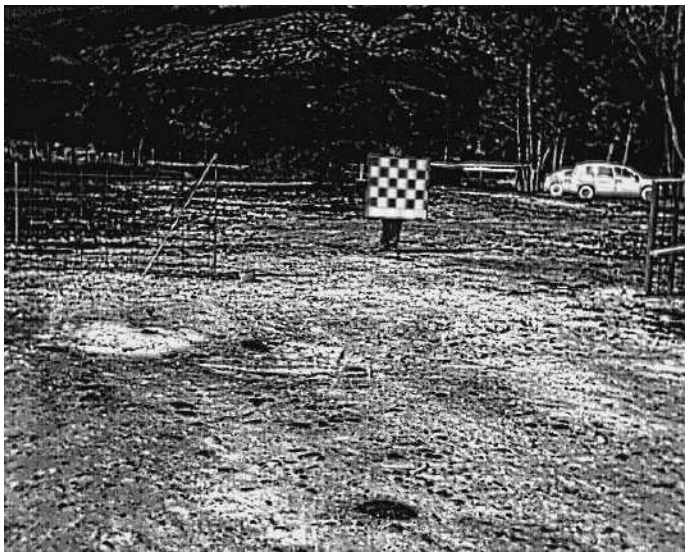
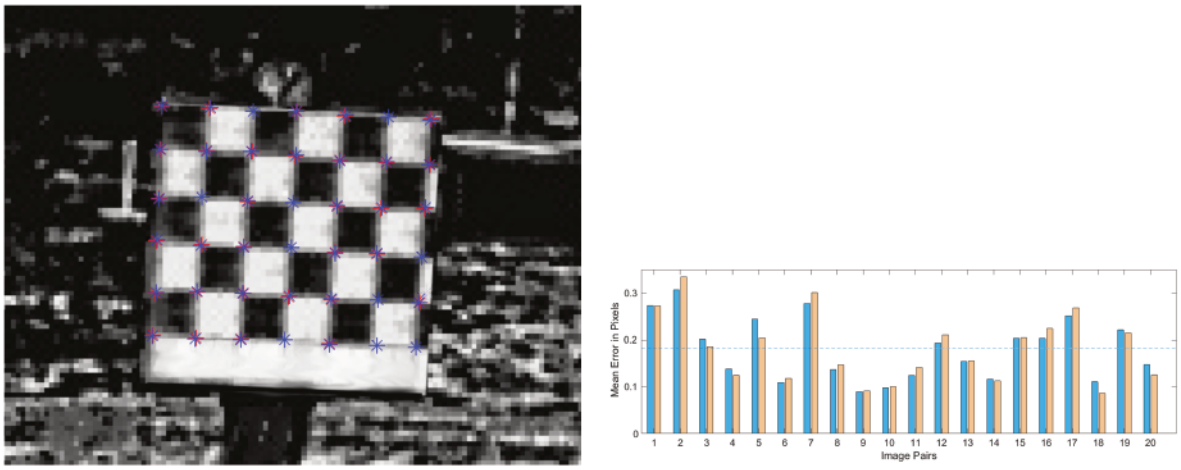


Figure 4.8: Image of the checkboard acquired using a LWIR camera.

Figure 4.9 shows the results obtained with the two LWIR cameras, with an error pixel of 0,19.




(a) (b)

Figure 4.9: Results of the LWIR cameras calibration. (*a*) Example of an image in LWIR domain with the starting points (red stars) and the reprojection points (blue stars). (*b*) Mean reprojection error of the two LWIR cameras.

## 4.3 Obtaining fire pixel points from multimodal stereoscopic images

### 4.3.1 Multimodal fire pixel detection

The step of detection of fire pixels in an image is essential in the process of measuring fire by vision because it determines the accuracy with which the characteristics of the phenomenon can be estimated. As discussed in the previous chapters, all reference measurements of wildfire obtained by vision and used by the scientific community are estimated from images acquired in the visible domain. However, this detection is complex because of the multiple and in-homogeneous colours of the flames, the presence or absence of smoke and the luminous intensity of the environment.

A reference work on the performance of the state-of-art fire pixel detection algorithms on visible band images is that of Toulouse *et al.* [119]. 11 wildland fire colour segmentation algorithms [61, 120, 121, 122, 123, 124, 125, 126, 127, 53, 128] were benchmarked using a dataset containing more than 500 wildfire images. The obtained results show that the tested

colour segmentation techniques performance is dependant on the image category (lighting, predominant colour, smoke).

The problems encountered to detect fire pixels into visible images do not occur in fire images acquired in the infrared domain. However, as mentioned in the chapter 2, the fire areas that appear in infrared images are not identical to those that can be identified in the visible images because the elements emitting in the infrared domain are not the same as those radiating in the visible spectrum. Thus, the use of infrared images alone cannot be considered for measuring the geometric characteristics of fire.

It was chosen in this thesis to use a multimodal pixel detection algorithm.

At each instant, visible stereoscopic images and LWIR stereoscopic images are acquired simultaneously. The images kept by the adjacent visible and LWIR cameras are processed by pair in a multimodal fire pixel procedure. Fig 4.10 shows an example of the visible image and the LWIR image of a fire acquired simultaneously with cameras with different focal lengths and resolutions.

(a) (b)

Figure 4.10: Example of visible and LWIR images acquired by the vision system used in this work. ($a$) Image acquired in the visible spectrum. ($b$) Image acquired in the LWIR spectrum.

The detection method is carried out in two steps, using the images of the cameras situated at the same side of the stereovision system: first the LWIR image is processed in order to localize a pre-selected fire area. Then, in the visible image, only the pixels that are at the same position that the one pre-selected are taken into account to detect fire pixels.

A modified Matlab toolbox's code is used to produce superimposed images [129]. The proposed method is a modification of the calibration method used in Matlab toolbox to allow a calibration of cameras with different resolution and focal. The intrinsic parameters of the single cameras are that obtained in the Par. 3.3. Theoretically, without knowing the depth

information, it is not possible to superimpose each pixel in the right place, to obtain an image in another image plane (superposition). The method therefore assumes that all objects are on the image plane, the plane $Z = 0$, where the upper left corner of the checkboard in the RGB image is the reference point. The method is able to find the transformation to be applied to the LWIR image to be superposed to the RGB image. Figure 4.11 shows the superposition of the visible and LWIR images: only the pixels situated into the fire area detected in the LWIR image are in colour.

The fire area detected in the LWIR image determines the location of the pixels in the visible image to process in order to detect the fire pixels.

Figure 4.11: Fire zone detected in the LWIR image used as a pre-selected area in the visible image.

### Detection of fire pixels in the LWIR image

In infrared images, fire pixels (characterized by high temperature) have higher pixel intensity than background pixels (Fig. 4.12). All pixels are normalized to increase the contrast between fire and background zones. This normalization is done using the Matlab function "imadjust", that maps the intensity values in grayscale image to new values. In particular, imadjust saturates the bottom 1% and the top 1% of all pixel values. Using JPEG image coded with 8 bits, the intensity values is into a range from 0 to 255.

Figure 4.12: Fire image acquired in the LWIR domain.

The Otsu method [130] is one of the most used threshold search methods in image processing. This technique makes it possible to choose a threshold that minimizes the variance of the pixel intensities between two classes of pixels and which maximizes the difference of the means intensity of these two classes. In this case, the two classes are fire and background. The images are considered to have 256 gray levels. Let $h$ the histogram of an image, composed by 256 classes. $p_k$ is the probability of distribution of level $k$, defined by:

$$p_k = \frac{h(k)}{N}, \tag{4.3}$$

where $N$ is the total number of pixels in the image.

The weighted inter-class variance is defined by:

$$\sigma_\omega^2(s) = \omega_1(s)\sigma_1^2(s) + \omega_2(s)\sigma_2^2(s), \tag{4.4}$$

with $s$ the separation threshold of the two classes; $\omega_1$ and $\omega_2$ are defined as:

$$\omega_1(s) = \sum_{k=1}^{s} p_k \tag{4.5}$$

$$\omega_2(s) = \sum_{k=s+1}^{256} p_k = 1 - \omega_1(s) \tag{4.6}$$

$\sigma_1$ and $\sigma_2$ are defined by the following equations:

$$\sigma_1^2(s) = \sum_{k=1}^{s} (k - \mu_1(s))^2 \frac{p_k}{\omega_1(s)} \tag{4.7}$$

$$\sigma_2^2(s) = \sum_{k=s+1}^{256} (k - \mu_2(s))^2 \frac{p_k}{\omega_2(s)}. \tag{4.8}$$

The means of the classes, $\mu_1$ and $\mu_2$, are defined as:

$$\mu_1(s) = \sum_{k=1}^{s} \frac{kp_k}{\omega_1(s)} \tag{4.9}$$

$$\mu_2(s) = \sum_{k=s+1}^{256} \frac{kp_k}{\omega_2(s)}. \tag{4.10}$$

Otsu's algorithm consists of finding the threshold $s$ which minimizes $\sigma_\omega^2(s)$. This algorithm was used in the "fire" pixel detection methods developed by Gouverneur *et al.* [87], Verstockt *et al.* [131] and Martínez-de-Dios *et al.* [132].

Let's $I_{i,j}$ the intensity value of the pixel of position $(i,j)$, $p_{i,j}$. In this detection, all the pixels of the infrared image whose the intensity is greater or equal than the threshold $s$ are selected.

$$fire\ pixels = \{p_{i,j} \mid I_{i,j} \geq s\} \tag{4.11}$$

These selected zones contain all the pixels relative to the fire.

**Detection of fire pixels in the visible image**

The detection of fire pixels in the visible image is the second step of the detection method of this thesis. Only the pixels situated in the visible image at the same position as those selected in the LWIR image are considered to detect fire pixels.

As discussed above, the performance of the eleven state-of-art fire colour segmentation algorithms are dependant on fire characteristics such as color, texture, presence of smoke and luminosity of the background. To overcome this problem, it has been decided to use, for each experiment, the algorithm among the eleven ones considered which visually gets the best pixel detection on the first image of the image sequence. The method developed by Rossi *et al* [53] was adapted in order to use only the second step of this method (the first step corresponding to a pixel pre-selection procedure). A graphical interface associated to the work described in [119] and developed by Tom Toulouse (Fig. 4.13).

Figure 4.13: Graphical interface developed by Tom Toulouse.

For all the experimental fire image sequences acquired for this PhD thesis, the adapted Rossi's method gives mostly the best detection results. As the experiments were carried using the same fuel, on the same field, and practically at the same period of the day and the year, this could be explained by the repetitive production of the same type of fire whose characteristics are favourable to the Rossi's method.

In this algorithm, a 3D Gaussian model is used to represent the pixels present within the fire zone. A pixel is identified as fire pixel if the values of the three RGB channels are close to the triple reference RGB of the distribution.

Defined:

$$\overline{m} = (\, m_r, m_g, m_b) \tag{4.12}$$

color means of extracted area's channels,

$$\sigma = max(\, \sigma_r, \sigma_g, \sigma_b) \tag{4.13}$$

the highest standard deviation of extracted area's channels, and

$$p = (\, p_r, p_g, p_b) \tag{4.14}$$

the RGB color value of one pixel of the RGB image; the pixel $p$ is considered a fire pixel fire or not using the following condition:

$$\begin{cases} |p - \overline{m}| \leq k * \sigma & p \in Fire \\ Otherwise & p \notin Fire \end{cases} \tag{4.15}$$

It can happen that the fire front splits into several parts as it can be seen in Fig. 4.14. As researchers working on fire modelling want information on the front of the fire, it was decided

to process the largest part of the fire located at the front of the fire. To do so, a k-clustering technique was applied to discriminate these zones, with $k = 1$ cluster.

(a) (b)

Figure 4.14: Example of the fires front. ($a$) Example of a fire front homogeneous.($b$) Example of a fire front separated into several parts.

The output of the fire pixel detection procedure is the visible image area containing all the selected color pixels. In this area will then be the detection of features.

First, a binary mask of the image it is obtained. In this mask the value 1 means that in that position there is a pixel of the fire zone, and 0 means that in that position there is a pixel of the background. An example of the binary image associated to an image containing fire is shown in Fig. 4.15.

Figure 4.15: Example of the binary mask of the fire image (in white are the pixels of the fire zone, in black are the pixels of the background).

Second, from the original image are selected all the pixels where their value in the binary mask is 1 (Fig. 4.16).

Figure 4.16: Example of output of the fire pixels detection.

### 4.3.2 Feature detection and matching points procedures

As mentioned in Section 3.6.2, the features are detected in the image fire areas by processing these zones successively with the Harris detection algorithm and the SURF procedure. The first method is applied with a box filter of $5 \times 5$ pixels. The second one uses 6 scale levels and the following filters size dimensions: $27 \times 27$, $51 \times 51$, $75 \times 75$, $99 \times 99$, $123 \times 123$ and $147 \times 147$. The points obtained by each procedure are added, and the duplicate points are deleted. Figure 4.17 shows the features selected by the complete procedure, where 3 238 points are detected.

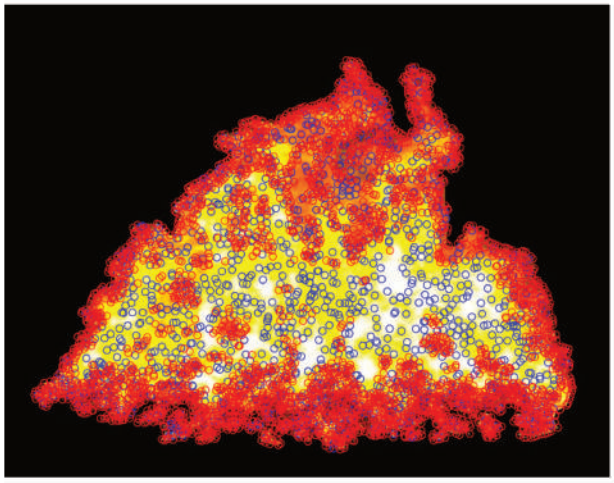
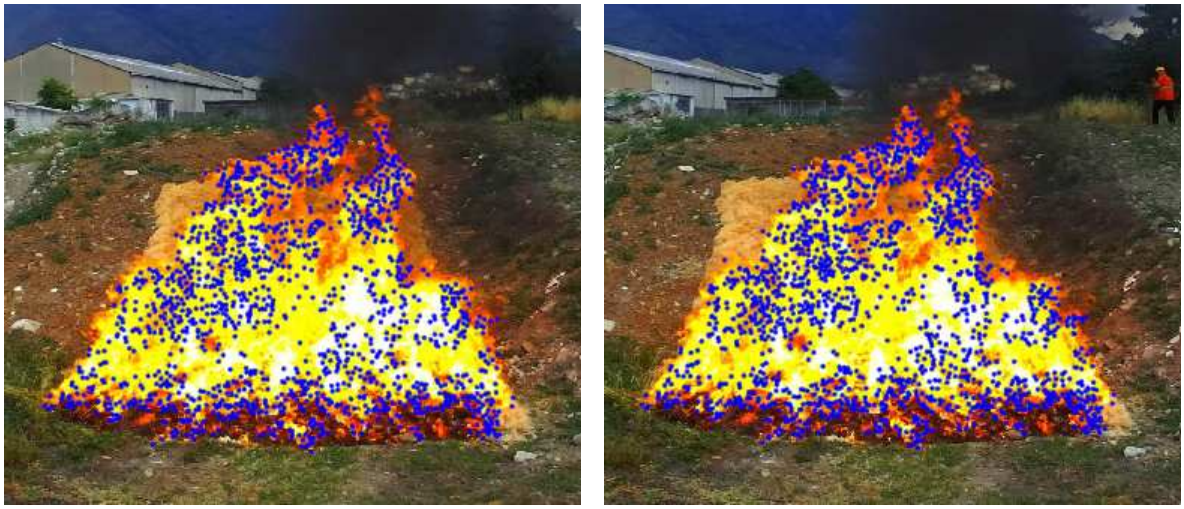


Figure 4.17: Features identified using the Harris algorithm (red circles) and the SURF algorithm (blue circles).

The feature matching procedure uses a zero mean normalized sum of square difference (ZNSSD) to measure descriptor similarities [74]. Figure 4.18 shows the features that are matched (1 803 points).



(a) (b)

Figure 4.18: Example of matched points. ($a$) Left image. ($b$) Right image.

## 4.4 3D reconstruction of fire points

The 3D fire points are obtained from the matched points using the Eq. 3.96.

Outliers points are eliminated by a a two-stage refining process. In the first step, are eliminated the isolated points. A 3D point is an isolated point if the mean distance between it and its 4 neighbors is too great. In the second step, are eliminated the 3D points too distant to the camera frame origin.

All the points that are located far from the principal 3D point cloud are considered as outliers. If one of the two conditions presented in the Eq. 4.16 is true, a point is identified as outlier and eliminate.

$$|d_1 - \mu_{d1}| \geq 3\sigma_{d1} \quad , |d_2 - \mu_{d2}| \geq 3\sigma_{d2}, \tag{4.16}$$

where $\mu_{d1}$, $\sigma_{d1}$ are respectively the mean and standard deviation of the distance $d_1$ between this point and its four neighbors; $\mu_{d2}$, $\sigma_{d2}$ are respectively the mean and standard deviation of the distance $d_2$ between this point and the camera frame origin.

Figure 4.19 shows the 3D fire points obtained from the fire image presented in Fig. 4.10.

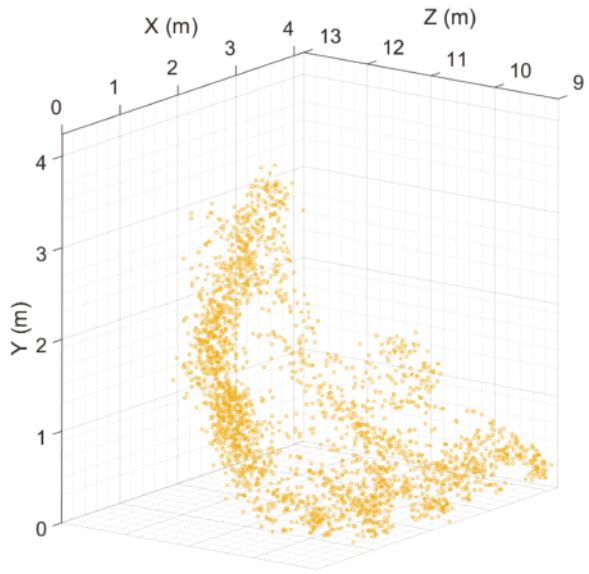



Figure 4.19: 3D fire points.

## 4.5 3D fire points transformation for the estimation of geometrical characteristics

At each instant of image acquisition, the vision system carried by drone has a position and orientation different from those of the previous moment. As the 3D points obtained by stereovision are expressed in a frame positioned in the vision device, and as this frame is moving over time, it is necessary to project all the 3D points in a common frame to produce results that show the temporal evolution of the fire geometrical characteristics; this reference frame will be noted as "Global reference frame". Moreover, these data have to be estimated independently of the field local plane on which the fire is located. This requires the application of a transformation function on the 3D fire points in order to work as if the slope of the propagation plane was zero; the x-axis of this frame being at each instant parallel to the field local plane and the y-axis will be normal to it. This frame will be called "Slope reference frame". In addition, as the fire may change its direction over time, to compute geometrical characteristics such as width, flame inclination and length, it is necessary to rotate the 3D points in such a way that their depth axis corresponds to the instantaneous main direction; the z-axis will be at each instant parallel to the fire instantaneous main direction; this frame will be called "Local direction reference frame".

### 4.5.1 Inputs/outputs of the transformation method

The inputs of the method are the position, the altitude and the orientation of the stereovision device at each instant.

The position is expressed by the latitude and longitude angles in degree; latitude is an angle between -90° and 90°, longitude is an angle between 0° and 180°. These angles are expressed with seven decimals. The measure of the altitude is expressed in mm as an entire scalar between 1 and 1 000 000 (1 km of maximal altitude). The orientation is composed by three values, two returned by the IMU board and one returned by the compass sensor. The values returned by the IMU board are the roll and pitch of the camera, expressed in radiant with fourteen decimals, between $-\pi$ and $\pi$. The value returned by the compass is an integer whose value is between 0° and 360°. It is in degree and corresponds to the heading angle of the camera.

The outputs of the method are the 3D points expressed into several reference frames. The first reference frame is positioned in the back of the burning area and it is used to obtain the evolution of the fire characteristics. The second one has its orientation that changes over time depending on the local slope of the spreading area. This frame is used to compute geometrical characteristics of the fire such as height. The third one is obtained from the second one applying a rotation in such a way that the longitudinal axis follows the instantaneous main direction of the fire. This frame is used to estimate the fire width and depth, for example.

### 4.5.2 Determination of the transformation matrix from the "Camera frame" to the "Global frame"

The 3D fire points obtained from the stereoscopic images are expressed in a frame attached to the stereovision system. The origin $O_C$ of this frame is the optical center of the left visible camera, the $X_C$ x-axis corresponds to the axis going from the left camera towards the right camera, the $Z_C$ z-axis is perpendicular to $X_C$ and directed forward the cameras, the $Y_C$ y-axis is the axis perpendicular to the others two such that the resulting triad is right-handed, as it can been seen in Fig. 4.20.

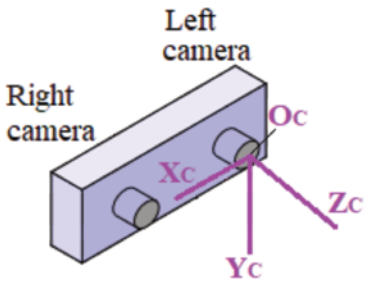



Figure 4.20: Stereovision system and Camera frame ($O_C$, $X_C$, $Y_C$, $Z_C$).

To express the 3D points in a world reference frame, the position given by the GPS sensor fixed on the vision system is used. Due that the GPS sensor and the left visible camera are not in the same place, the information given by the GPS sensor is not exactly the one of the origin of the Camera frame. The position of the Camera frame, which is the frame in which the 3D points obtained by stereovision are expressed, is in a first step, considered to be at the same position as the GPS sensor. A transformation is applied in order to take into account this position offset. The 3D points are expressed in a frame which is the same than the Camera frame but positioned on a point $O_D$ situated on the line joining the optic center of the two cameras and at the vertical position of the GPS sensor (Fig. 4.21). This frame is called "Drone frame" (D frame), its axis $X_D$, $Y_D$ and $Z_D$ are respectively colinear to the axis $X_C$, $Y_C$ and $Z_C$.

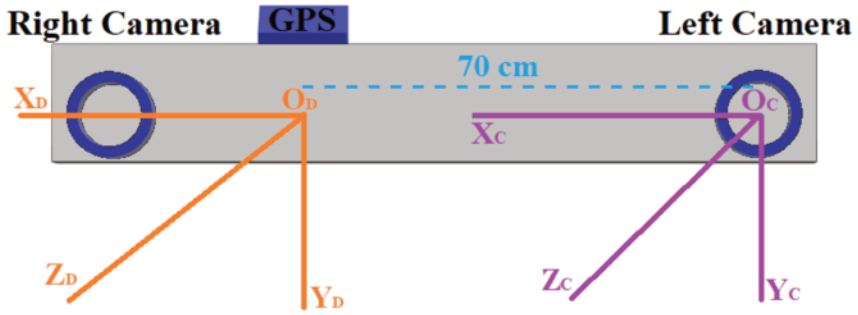



Figure 4.21: Camera frame (purple color) and Drone frame (orange color).

The Drone frame is an intermediate frame used to compute the position of $O_C$ in the World. Due that the height of the 3D points is represented by the y-coordinate, and due that the $Y_C$ axis of the Camera frame is direct downwards, the measure on the y-axis of the 3D points are not coherent with reality. For this reason, it is added another intermediate frame, called "Local frame". This frame has the origin $O_L = O_C$, the $X_L$ axis directs to the right direction of the camera, the $Z_L$ axis directs forward the camera, and the $Y_L$ axis directs upwards.

#### 4.5.2.1 Determination of the transformation matrix from the Drone frame to the Local frame

The coordinates of $O_C$ in the Drone frame are:

$$\begin{bmatrix} Ox_C^D \\ Oy_C^D \\ Oz_C^D \end{bmatrix} = \begin{bmatrix} -70 \\ 0 \\ 0 \end{bmatrix} \tag{4.17}$$

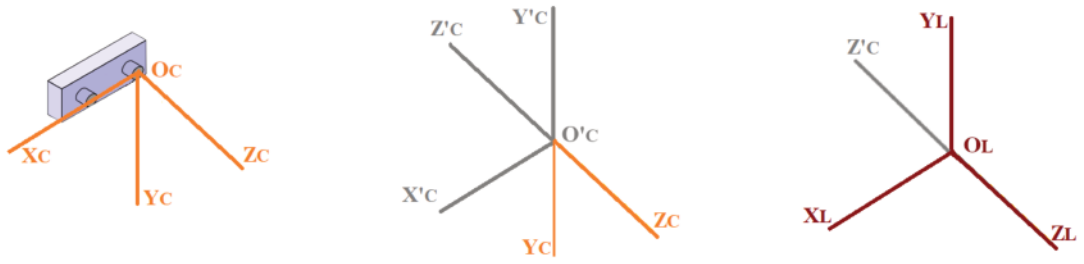
Due that the $X_D$, $Y_D$ and $Z_D$ axis of the Drone frame are respectively colinear to the $X_C$, $Y_C$ and $Z_C$ axis of the Camera frame, the matrix $M_0$ that transforms the Drone frame to the Camera frame is:

$$M_0 = \begin{bmatrix} & Ox_C^D \\ I_{3\times3} & Oy_C^D \\ & Oz_C^D \\ \mathbf{0}_{(1\times3)} & 1 \end{bmatrix}. \tag{4.18}$$

Taken into account the Eq. 4.18, the coordinates of a 3D point expressed in the Drone frame ($x_i^D$, $y_i^D$, $z_i^D$) can be considered to obtain the coordinates of this point in the Camera frame ($x_i^C$, $y_i^C$, $z_i^C$ ) by using the following formula:

$$\begin{bmatrix} x_i^C \\ y_i^C \\ z_i^C \\ 1 \end{bmatrix} = M_0 \begin{bmatrix} x_i^D \\ y_i^D \\ z_i^D \\ 1 \end{bmatrix}. \tag{4.19}$$

The y-axis of the Camera frame is perpendicular to the z-axis and directs downwards the cameras (Fig. 4.22a). In order that all the fire geometrical characteristics are expressed with positive values, a transformation is made on the Camera frame so that the y-axis is increasing upwards. The Camera frame is therefore rotated of 180° along the x-axis to obtain a new frame (Fig. 4.22b), called "Intermediate frame". In this frame, the $Z'_C$ z-axis is directed to the rear of the cameras device. In order to express the depth with positive values an z-axis inversion is applied to the intermediate frame to obtain a new frame called "L frame" (Fig. 4.22c).




(a) (b) (c)

Figure 4.22: Frames used in the transformation procedure. ($a$) Camera frame (orange color). ($b$) Intermediate frame: 180° x-axis rotation of the Camera frame (gray color). ($c$) L frame: $Z_L = -Z'_C$ (red color).

The transformation to pass from the Camera frame to the L frame is expressed by:

$$\begin{bmatrix} X_L \\ Y_L \\ Z_L \\ 1 \end{bmatrix} = M_1 \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} \tag{4.20}$$

where $M_1$ is given by:

$$M_1 = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & cos(180) & -sin(180) & 0 \\ 0 & sin(180) & cos(180) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{180° \text{ x-rotation}} \cdot \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\text{z-axis inversion}} \tag{4.21}$$

In order to produce fire geometrical characteristics data in the same format that the one generated by the fire behavior model developed by the researchers of the "Fire"project of the UMR SPE, it is necessary to project the 3D points in a fixed frame situated at the rear of the fire ignition line. It was chosen to take as origin of this frame (named "Global reference frame - G frame") the GPS data ($O_G$) associated to the first image obtained by the stereovision system before the take-off of the drone. The $X_G$ x-axis of the Global reference frame is parallel to the width side of the burn zone, its $Z_G$ z-axis corresponds to the depth of the burn zone, and the $Y_G$ y-axis is such that the triad is left-handed, corresponding to the altitude (Fig. 4.23).

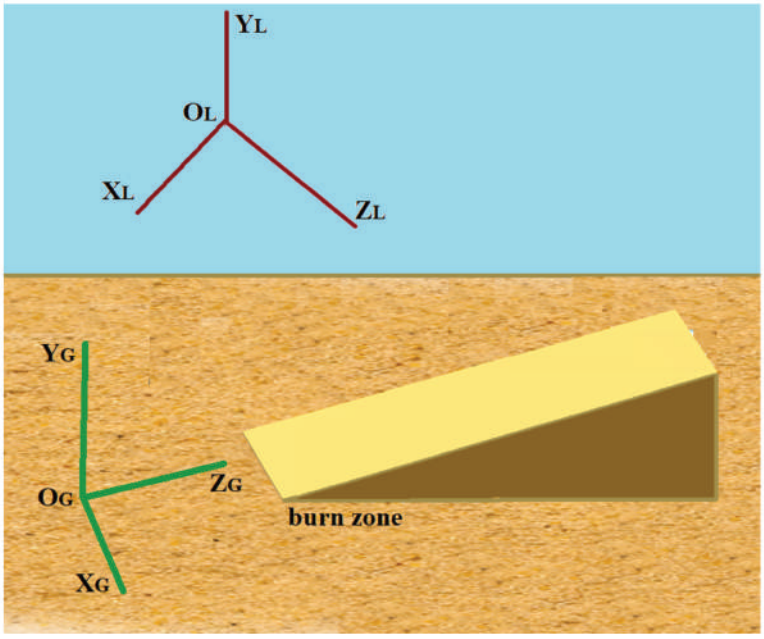



Figure 4.23: L frame and G frame (Global reference frame).

#### 4.5.2.2 Determination of the transformation matrix from the Local frame to the Global frame

The transformation from the Local frame associated with the stereovision system to the fixed Global frame situated on the ground is carried out in several stages.

The positions returned by the GPS sensor are expressed in angular form into the geodetic reference system LLA (Latitude, Longitude, Altitude) defined by $\varphi$ for the latitude, $\lambda$ for the longitude, and $h$ for the altitude. As the 3D points are expressed into the Cartesian system, it is necessary to convert polar GPS coordinates to Cartesian coordinates. The global Cartesian frame WGS84, also called ECEF frame, in which the GPS coordinates are converted, has its origin in the center of the earth $O$ and its axes are $X_{ecef}$, $Y_{ecef}$ and $Z_{ecef}$; where $X_{ecef}$ joins the center $O$ to the intersection between the "Prime meridian" and the latitude of 0°, $Z_{ecef}$ joins the center $O$ to the center of the terrestrial pole, and $Y_{ecef}$ is the axis perpendicular to the others two such that the resulting triad is right-handed. Figure 4.24 presents the ECEF frame, the G Global frame and the L frame.

**Conversion from LLA coordinates to Cartesian coordinates**

Due that the GPS polar angles depend on the eccentricity of the terrestrial form, it is not possible to convert directly polar coordinates to Cartesian coordinates using only the latitude and longitude angles and the altitude value; but it is necessary to also consider information about the Earth form.

The World Geodetic System 1984 defines a mathematical model of the Earth. The measurement of the semi-major axis $a$ is approximately 6 378 km and the measurement of the semi-minor axis $b$ is approximately 6 356 km. Using these values, the ellipsoid flatness factor of the Earth $f$ is:

$$f = 1 - \frac{b}{a} \approx 0.003 \tag{4.22}$$

and the terrestrial eccentricity factor $e_s$ is:

$$e_s = 2f - f^2 \approx 0.007. \tag{4.23}$$

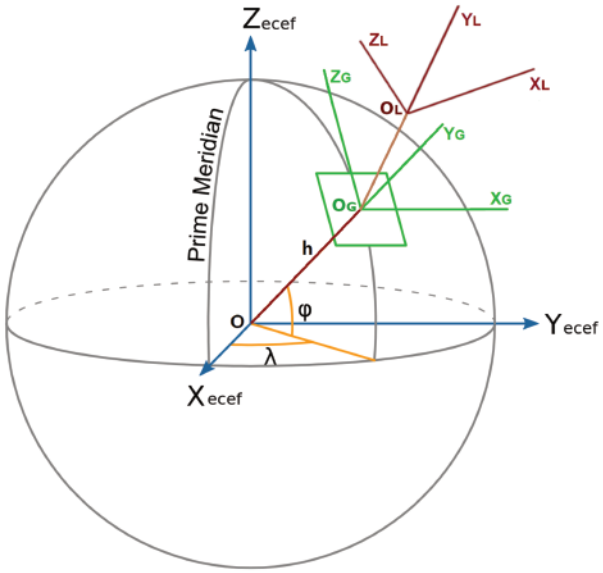



Figure 4.24: G frame (blue color), L frame (red color) and ECEF frame (green color).

The Cartesian coordinates of a point ($x$, $y$, $z$) can be obtained from the ellipsoidal coordinates of a point ($\varphi$, $\lambda$,h) as follows:

$$x = (h + N)\cos\varphi\cos\lambda \tag{4.24}$$

$$y = (h + N)\cos\varphi\sin\lambda \tag{4.25}$$

$$z = [h + (1 - e_s)N]\sin\varphi. \tag{4.26}$$

where $N$, the correction factor that considers the terrestrial form, is

$$N = \frac{a}{\sqrt{1 - e_s \sin\varphi^2}}. \tag{4.27}$$

Using the Eq. 4.24-4.26, the Cartesian coordinates of the point $O_G$ ($x_0$, $y_0$, $z_0$) and of the point $O_L$ ($x_i$, $y_i$, $z_i$), with $i$ corresponding to the instant $i$, are obtained.

The vector $\overline{O_G O_L} = \begin{bmatrix} u, v, w \end{bmatrix}^T$ represents the euclidean distance from $O_L$ to $O_G$ in the ECEF system, and is defined by:

$$u = x_i - x_0, \tag{4.28}$$

$$v = y_i - y_0, \tag{4.29}$$

$$w = z_i - z_0. \tag{4.30}$$

This vector is used to find the coordinates of $O_L$ in the Global frame.

The standard reference system used to compute the distance between two points of the Earth is the ENU frame (East,North,Up). In this frame, the origin can be any point on the Earth. In this thesis, it is positioned at the reference point $O_G$. The $X_{enu}$ x-axis is directed towards East, the $Y_{enu}$ y-axis is directed towards North, and the $Z_{enu}$ z-axis is directed upwards (Fig. **??**). In this thesis, the ENU frame is used as an intermediate frame to obtain the transformation matrix from the Local frame to the Global frame (Fig. 4.25).

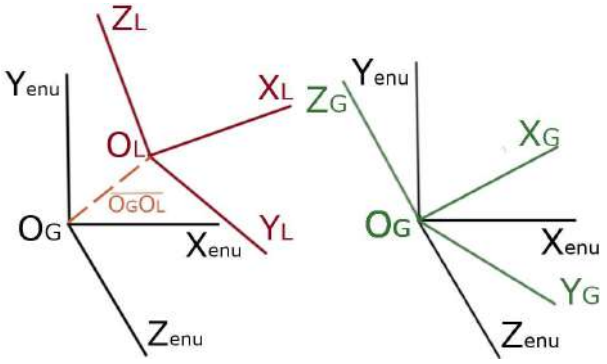



Figure 4.25: Local, Global and ENU frames. *(a)* Local frame (red color) and ENU frame centered in $O_G$ (black color). *(b)* ENU frame (black color) and Global frame (green color).

**Determination of the transformation matrix from the Local frame to the ENU frame**

Using the LLA coordinates of $O_L$ ($\varphi_i$, $\lambda_i$, $h_i$) and the coordinates of the vector $\overline{O_G O_L}$ expressed in the ECEF frame $\begin{bmatrix} u, v, w \end{bmatrix}^T$, the Cartesian coordinates of $O_L$ in the ENU frame centered in $O_G$ $(O_{L_x}, O_{L_y}, O_{L_z})$ are obtained with the relation:

$$\begin{bmatrix} O_{L_x} \\ O_{L_y} \\ O_{L_z} \end{bmatrix} = \underbrace{\begin{bmatrix} -sin(\lambda_i) & cos(\lambda_i) & 0 \\ -sin(\varphi_i)cos(\lambda_i) & -sin(\varphi_i)sin(\lambda_i) & cos(\varphi_i) \\ cos(\varphi_i)cos(\lambda_i) & cos(\varphi_i)sin(\lambda_i) & sin(\varphi_i) \end{bmatrix}}_{\text{Transfomation matrix from ECEF to ENU}} \begin{bmatrix} u \\ v \\ w \end{bmatrix}. \tag{4.31}$$

$\left[O_{L_x}, O_{L_y}, O_{L_z}\right]^T$ is the translation vector that allows to translate the origin of the L frame to the origin of the ENU frame.

The heading angle $\delta$ measured by the compass mounted on the stereocamera device gives the angle between the $Z_L$ axis and the magnetic north, defined by the $Y_{enu}$ axis. The measure of the roll data ($\alpha$) represents the angle between $X_L$ and $X_{enu}$, and the measure of the pitch data ($\beta$) represents the angle between $Y_L$ and $Z_{enu}$. Figure 4.26 shows the three angles $\delta$, $\alpha$ and $\beta$.

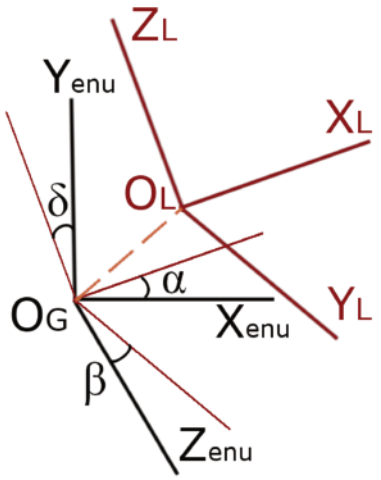



Figure 4.26: Angles between the Local frame and the ENU frame.

The rotation around the $Y_L$ axis of an angle equal to $\delta$ allows to align $Z_L$ and $Y_{enu}$, and its corresponding matrix is noted $R_\delta(\delta)$. $R_\alpha(\alpha)$ is the matrix associated to the rotation around the $Z_L$ axis of an angle equal to $\alpha$, and $R_\beta(\beta)$ is the matrix corresponding to the rotation around the $X_L$ axis of an angle equals to $\beta$. $R_\delta(\delta)$, $R_\alpha(\alpha)$ and $R_\beta(\beta)$ are given by:

$$R_\delta(\delta) = \begin{bmatrix} cos(\delta) & 0 & -sin(\delta) \\ 0 & 1 & 0 \\ sin(\delta) & 0 & cos(\delta) \end{bmatrix}, \tag{4.32}$$

$$R_\alpha(\alpha) = \begin{bmatrix} cos(\alpha) & -sin(\alpha) & 0 \\ sin(\alpha) & cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.33}$$

$$R_\beta(\beta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\beta) & -sin(\beta) \\ 0 & sin(\beta) & cos(\beta) \end{bmatrix}. \tag{4.34}$$

The Local frame is oriented such that its $X_L$ axis is colinear with the $X_{enu}$ axis, its axis $Z_L$ is colinear with the $Y_{enu}$ axis, and its axis $Y_L$ is colinear with the $Z_{enu}$ axis (Fig. 4.27).

The rotation matrix $R_L$ is obtained by the successive application of the rotation $R_\alpha$, $R_\beta$ and $R_\delta$, applied with the angle $-\alpha$, $\beta$, and $\delta$, respectively, as:

$$R_L = R_\alpha(-\alpha) * R_\beta(\beta) * R_\delta(\delta). \tag{4.35}$$

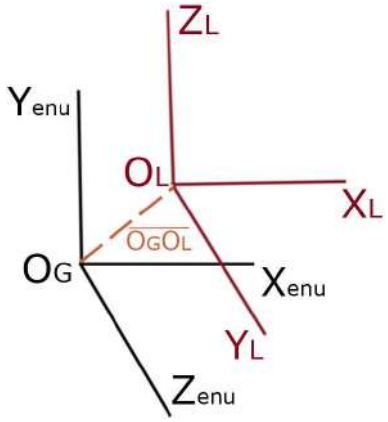



Figure 4.27: Local frame colinear to the ENU frame.

Due that the Local frame is a left-handed system and the ENU frame is a right-handed system, it is necessary to swap the $Y_L$ and $Z_L$ axis, applying a transformation whose matrix is:

$$M_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.36}$$

The roto-translation matrix $M_3$ that allows the transformation from the Local frame to the ENU frame centered in $O_G$ is:

$$M_3 = \begin{bmatrix} & & O_{L_x} \\ R_{L3\times3} & & O_{L_y} \\ & & O_{L_z} \\ \mathbf{0}_{(1\times3)} & & 1 \end{bmatrix} M_2. \tag{4.37}$$

The ENU frame is obtained from the Local frame using the relation:

$$\begin{bmatrix} X_{enu} \\ Y_{enu} \\ Z_{enu} \\ 1 \end{bmatrix} = M_3 \begin{bmatrix} X_L \\ Y_L \\ Z_L \\ 1 \end{bmatrix} \tag{4.38}$$

**Determination of the transformation matrix from the ENU frame to the Global frame**

The same procedure as the one described in 4.5.2.2 is used to obtain the transformation matrix from the ENU frame to the Global frame.

The $\delta_0$ value returned by the compass in position $O_G$ is the angle between the $Z_G$ and the $Y_{enu}$ axis. The angle between the $X_G$ and the $X_{enu}$ axis is equal to $\alpha_0$ (the value of the roll angle in position $O_G$), and the angle between the $Y_G$ and $Z_{enu}$ axis is equal to $\beta_0$ (the value of the pitch angle in position $O_G$) (Fig. 4.28).

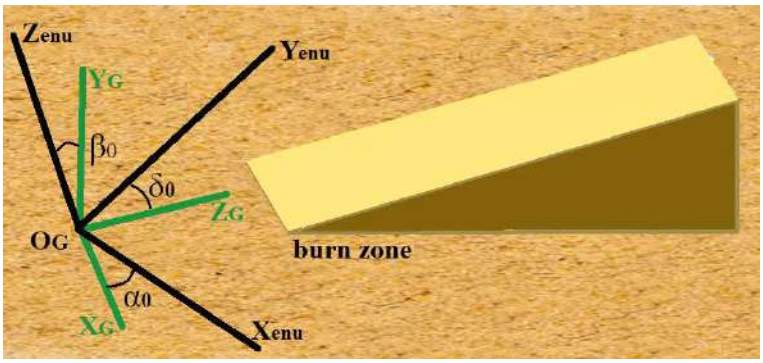



Figure 4.28: Angles between Global frame (green color) and ENU frame (black color).

Due that $\alpha_0$, $\beta_0$ and $\delta_0$ are the angles between the axis of the Global frame and the axis of the ENU frame, the transformation matrix to pass from the ENU frame to the Global frame has to use the opposite angles -$\alpha_0$, -$\beta_0$ and -$\delta_0$.

The rotation matrix is obtained by the successive application of the rotation $R_\delta$, $R_\beta$ and $R_\alpha$ applied with the -$\delta_0$, -$\beta_0$ and $\alpha_0$ angles is noted $M_4$, and is equal to:

$$M_4 = \begin{bmatrix} (R_\delta(-\delta_0) * R_\beta(-\beta_0) * R_\alpha(\alpha_0))_{3\times3} & \mathbf{0}_{(3\times1)} \\ \mathbf{0}_{(1\times3)} & 1 \end{bmatrix}. \tag{4.39}$$

The Global frame is obtained from the ENU frame using the relation:

$$\begin{bmatrix} X_G \\ Y_G \\ Z_G \\ 1 \end{bmatrix} = M_4 * M_2 \begin{bmatrix} X_{enu} \\ Y_{enu} \\ Z_{enu} \\ 1 \end{bmatrix}. \tag{4.40}$$

Finally, 3D points obtained in the Camera reference frame with the coordinates $(x_i^C, y_i^C, z_i^C)$

are projected in the Global frame by using the relation:

$$\begin{bmatrix} x_i^G \\ y_i^G \\ z_i^G \\ 1 \end{bmatrix} = M_4 * M_2 * M_3 * M_1 \begin{bmatrix} x_i^C \\ y_i^C \\ z_i^C \\ 1 \end{bmatrix}. \tag{4.41}$$

In this way, the point of view of the fire is changed, from the point of view of the Camera frame attached to the vision device to the point of view of the fixed frame attached to the Global reference frame (Fig. 4.29).

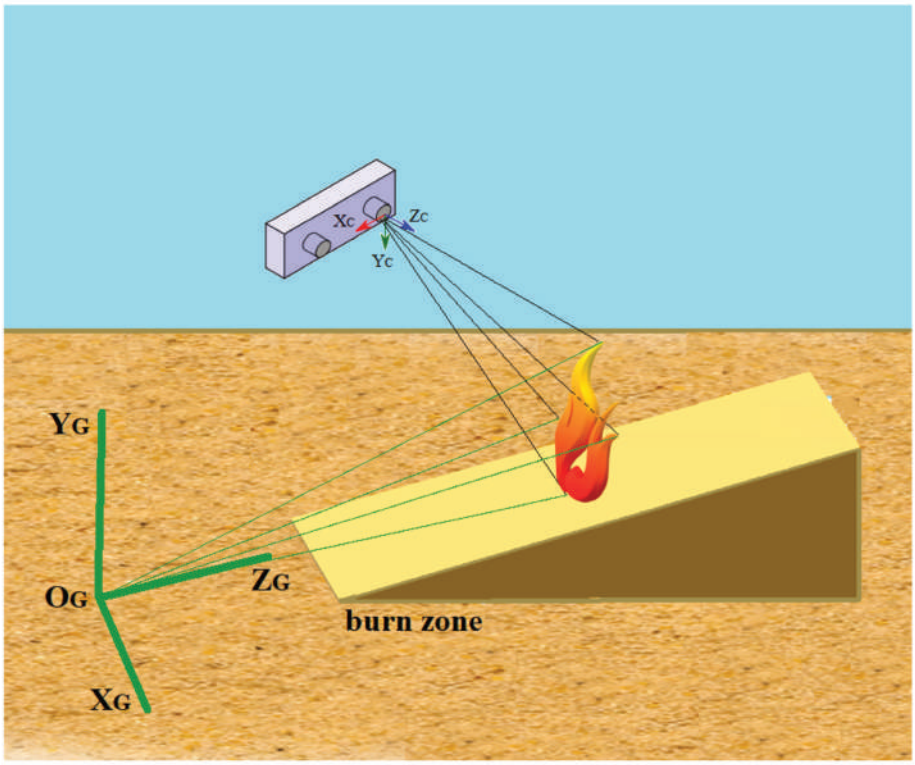



Figure 4.29: Camera frame and Global frame points of view.

### 4.5.3 Determination of the transformation matrix from the Global frame to the Slope frame

It is considered in this thesis that fires can propagate on a terrain whose lateral and longitudinal slopes are not constant (Fig. 4.30).

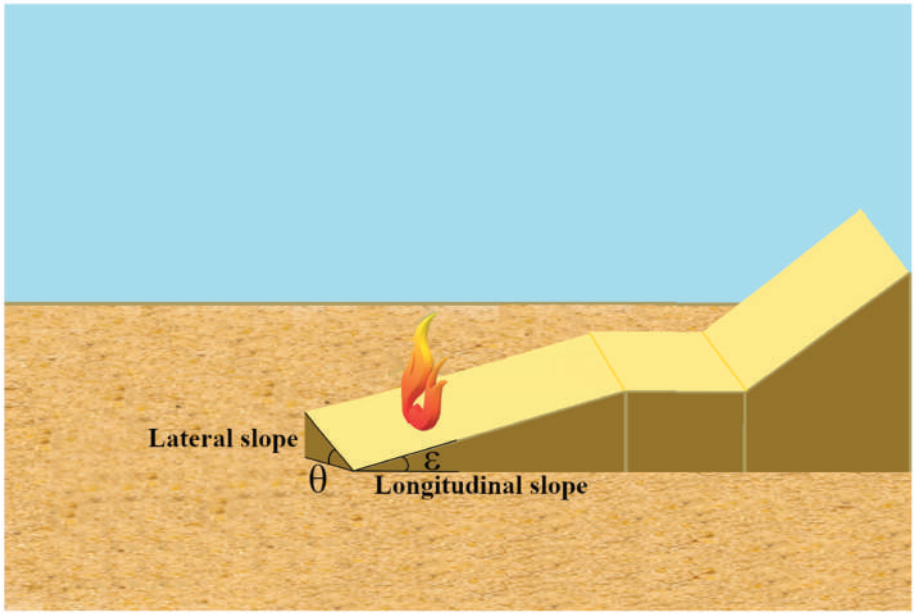



Figure 4.30: Lateral and longitudinal slope of the ground.

Fire geometrical characteristics are intrinsic information and must not incorporate information related to the ground. The height $h$ of a 3D fire point, named $P_H$ for example, is the distance between this point and the orthogonal projection of this point $P'_H$ on the ground plane (also called propagation plane) (Fig. 4.31). To do so, it is necessary to compute at each acquisition time the equation of the local propagation plane.

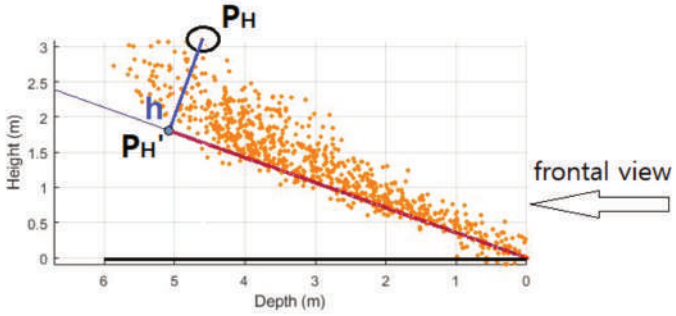



Figure 4.31: Height of a 3D fire point.

A differentiation between the points of the base and the points of the upper part of the fire front is made in order to estimate the position of the fire on the ground and its height. A procedure containing two steps is conducted using the 3D points belonging to the fire extracted at successive times and corresponding to a fire propagation path. Firstly, at each acquisition time, the equation of the local propagation plane is computed. Secondly, the 3D points belonging to planes obtained at successive instants with slope variations of less than 3° are gathered and used to compute the equation of the local plane passing thought them.

The information given by the average planes are used to compute the orientation of the Slope frame for each acquisition.

#### 4.5.3.1 Estimation of the fire base plane equation

When a fire is propagating, the vegetation is destroyed as the fire passes through; the rear of the fire appears to be burning areas without fuel (Fig. 4.32). Images taken from the rear of the light reveal flames going to the ground.

Figure 4.32: Example of burning area appearing in the back of the fire front.

A method has been developed to estimate the local propagation plane of a fire using the 3D flame points situated on the ground. As it appears in the Fig. 4.33, fire ground points have their height coordinate that changes along the slope. Therefore, it is not possible to select them by considering only this information.

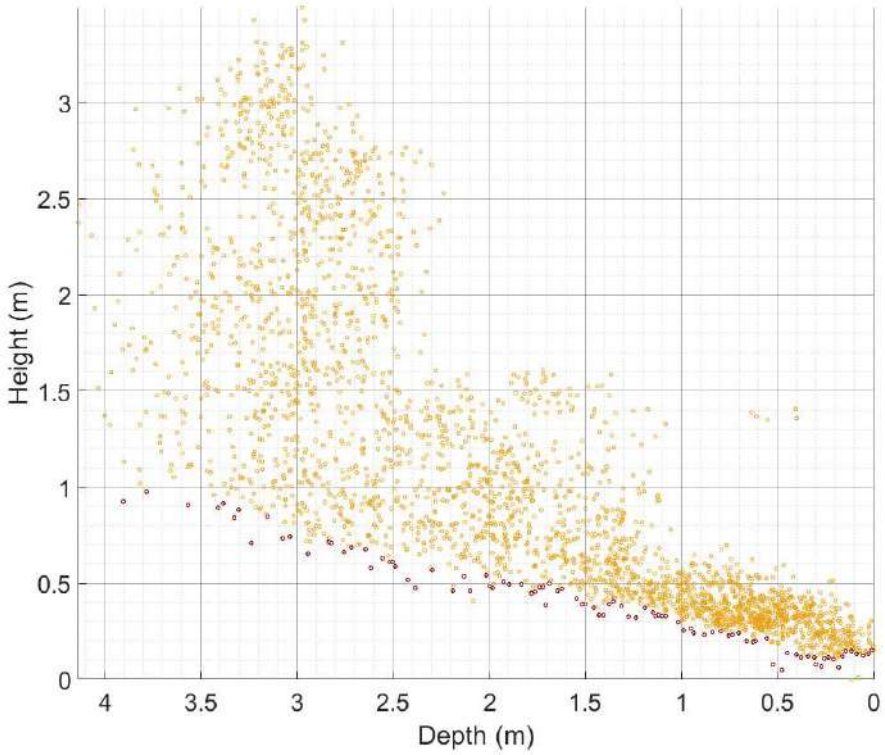



Figure 4.33: Lateral view of 3D fire points obtained at a given acquisition instant (the red points are the points on the ground).

The procedure is carried out at each acquisition instant, and it is done in two steps. Firstly, the equation of the fire base plane is obtained from the selected 3D fire point that are on the ground. From this equation, the lateral and longitudinal angles of the plane are obtained. Secondly, taken into account this information, a refining procedure is done using a rotation of all the 3D fire points obtained at the corresponding instant and a second estimation of the base plane equation.

To start the procedure, the left image obtained at the corresponding acquisition instant is processed in order to identify the fire pixels that are on the ground. These pixels correspond to the ones that are on the back front line of the fire. To identify them, a sort is carried out on the detected fire pixels in such a way to process them by increasing column value. For each column, the pixel that has the highest value is selected. Spurious points corresponding to pixels having a $v$ value very different from the others are eliminate using a Gauss function on the average of the $v$, with a standard deviation of $3\sigma$. Figure 4.34 shows an example of pixels (red color) selected by this procedure among the pixels of fire detected at a given moment (a red outline rectangle has been added to the figure to focus the reader's attention on the selected points).

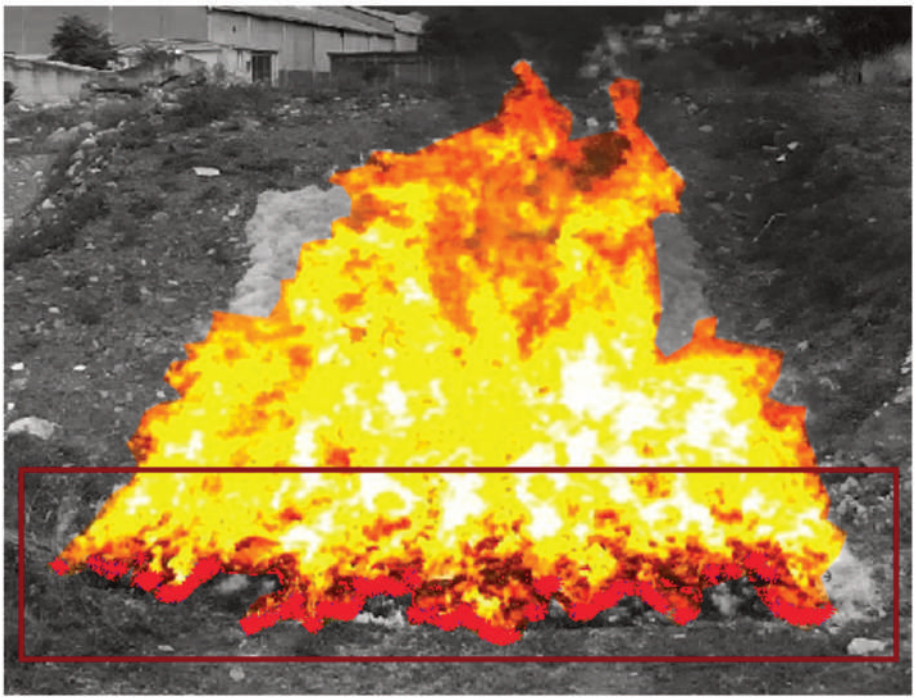
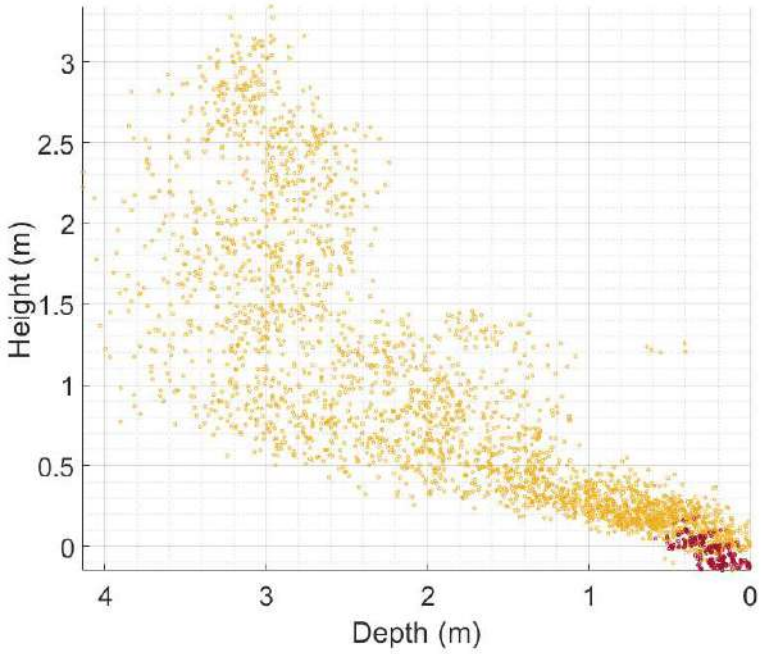
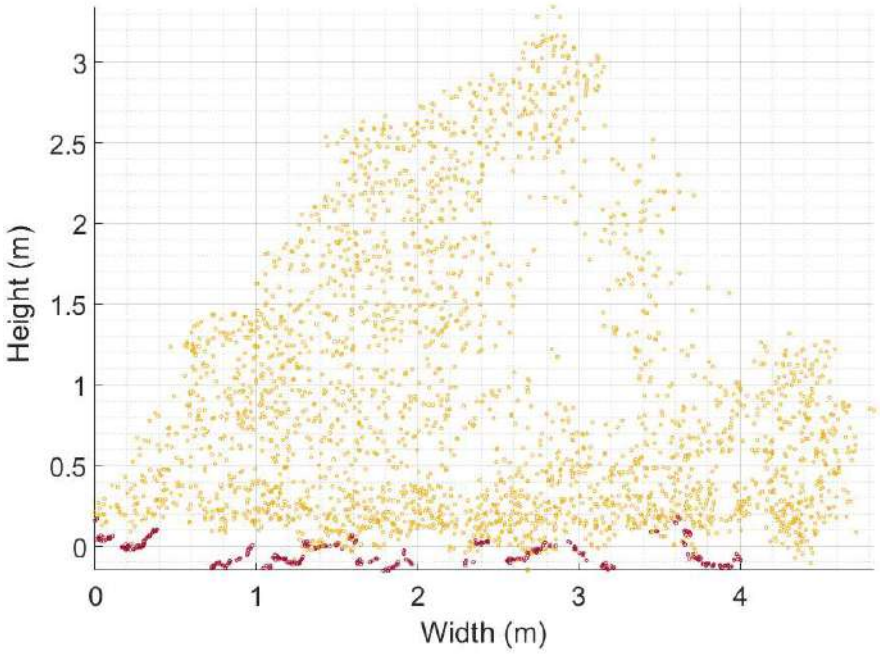


Figure 4.34: Example of ground fire pixels selected by the proposed method (red crosses).

From the selected 3D points an average point is computed. The 3D points that are not more than 30 cm above the mean point are used to obtain the equation of the fire base plane (Fig. 4.35).




(a)




(b)

Figure 4.35: Selected 3D fire points used in the base plane estimation procedure. (a) Lateral view. (b) Frontal view.

The RANSAC (RANdom SAmple Consensus) method [133] is used to obtain the Cartesian

equation of the base plane also called "local propagation plane". Figure 4.36 shows an example of the 3D fire points obtained at a given instant, the 3D fire points identified as being on the ground, and the base plane drawn from its equation obtained with the 3D ground fire.

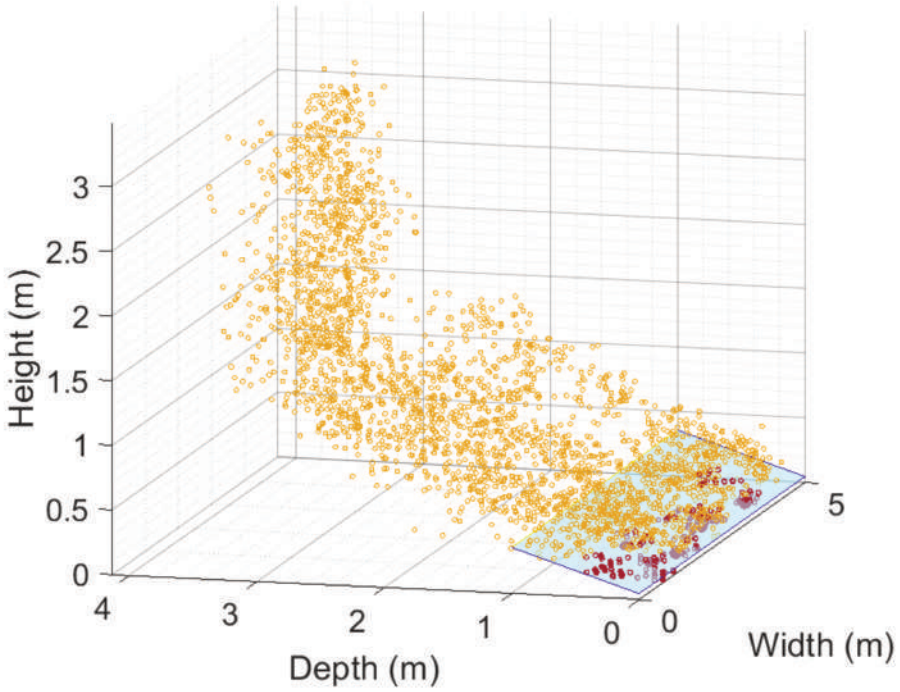
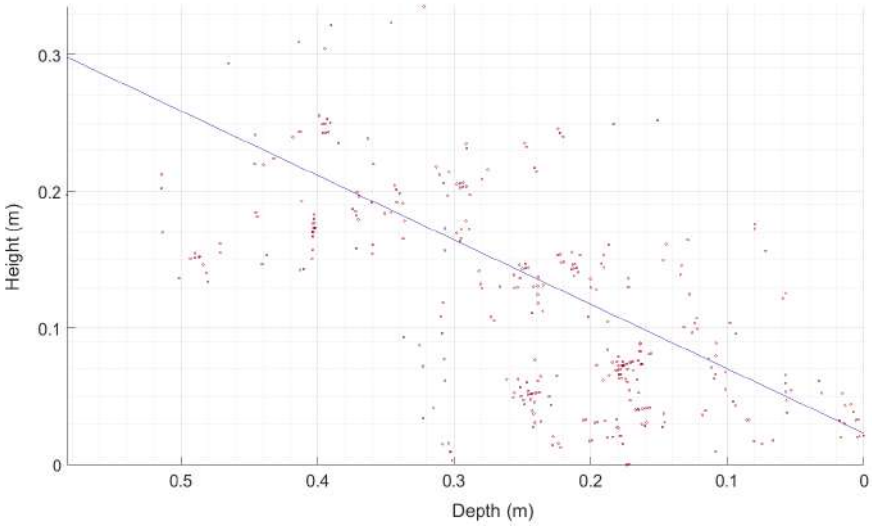
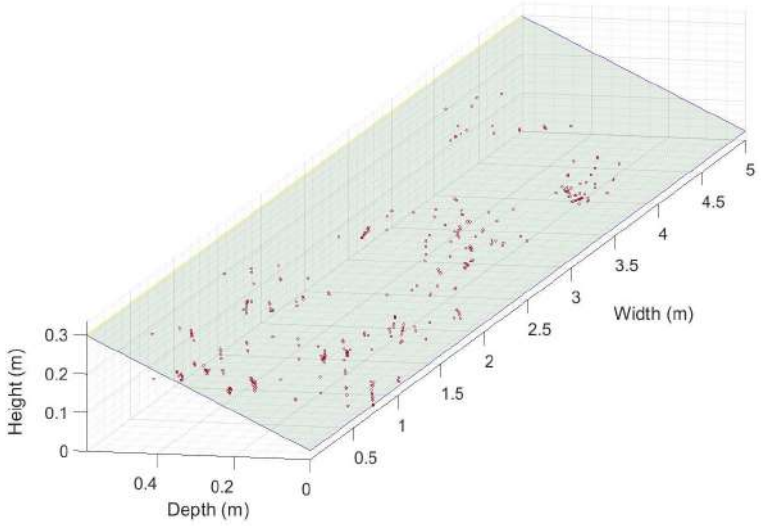



Figure 4.36: Example of 3D fire points, identified 3D ground fire points and base plane drawn from its equation.

Figure 4.37 shows the 3D points which are used for the estimation of the plane, and the plane which passes through these points.




(a)




(b)

Figure 4.37: Estimation of the base plan from the 3D lowest points of a fire. (a) Lateral view. (b) Global view.

The intersection of this base plane with the Y-Z plane defines a line, and the angle noted $\epsilon$ between this intersection line and the depth axis gives the longitudinal angle of the base (Fig. 4.38).

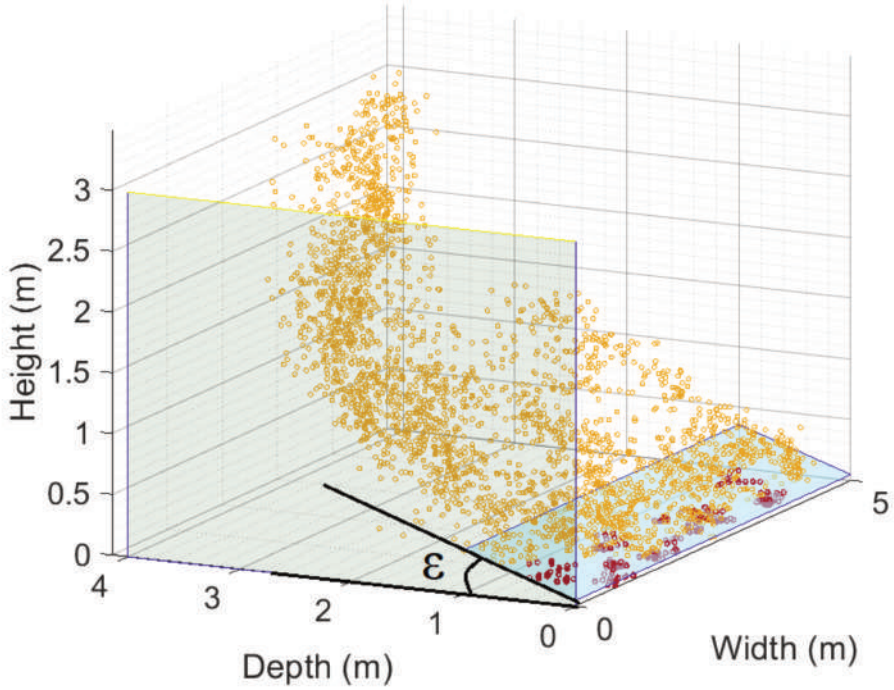



Figure 4.38: Longitudinal angle of the fire base.

The intersection of the base plane with the Y-Z plane is defined by the equation system:

$$\begin{cases} ax + by + cz + d = 0 \\ x = 0 \end{cases} \tag{4.42}$$

and the equation of the intersection line is:

$$z = \frac{b}{c}y + \frac{d}{c}. \tag{4.43}$$

The angle between the intersection line and the depth axis, noted as $\epsilon$, is obtained by the equation:

$$\epsilon = \arctan\frac{b}{c}. \tag{4.44}$$

The same method is used to find the angle of the lateral slope, noted $\theta$. In this case, the intersection line between the base plane and the X-Z plane is given. The X-Z plane is the plane between the width and the depth axis that has the equation $y = 0$. The angle $\theta$ is equal to:

$$\theta = \arctan \frac{c}{a}. \tag{4.45}$$

The inclination of the slope affects the detection of the low points of the fire. The points that are on the ground may be in an insufficient number and they can be concentrated. Therefore, these points are used to have a first estimation of the inclination of the fire plane. A refining procedure is carried out to identify new points of the base, and on this new set of 3D points a new base plane is estimated. The refining procedure begins by rotating all the 3D fire points (obtained at the time in question) around the width axis of the angle $\epsilon$ and around the depth axis of the angle $\theta$. Figure 4.39 shows the 3D points presented in the Fig. 4.36 after the first rotation.

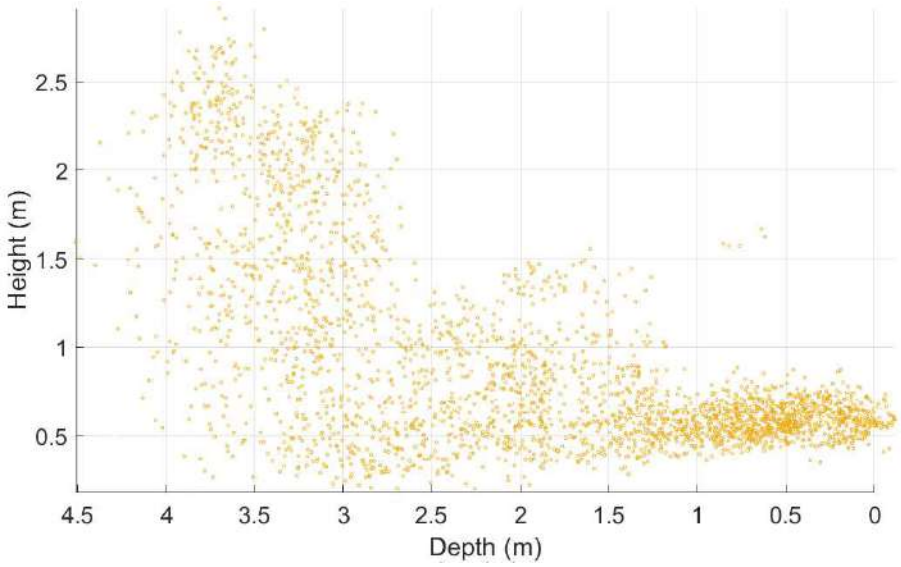
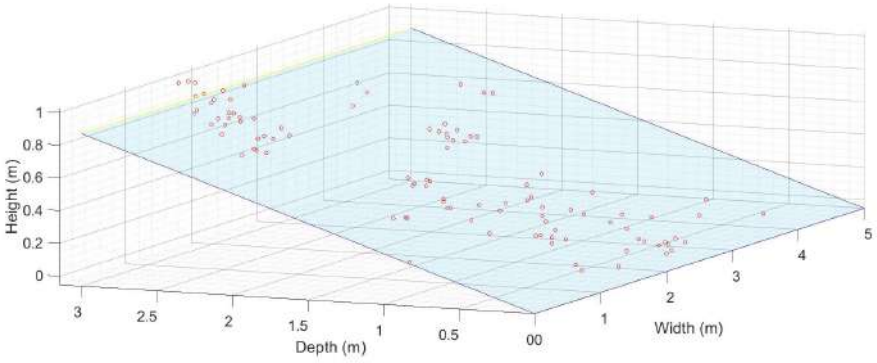
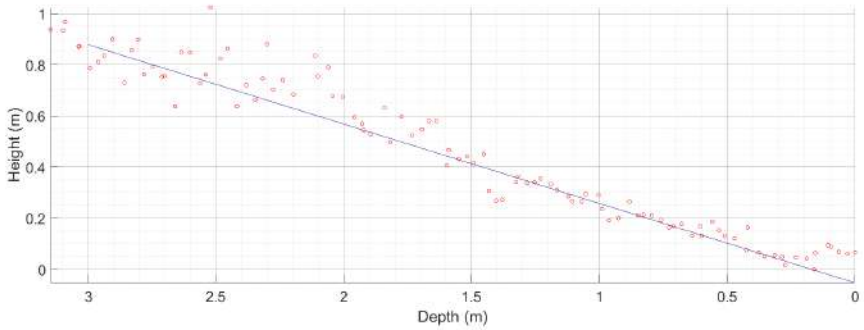



Figure 4.39: 3D points after the rotation using the angles of the base plane.

A procedure for selecting the 3D lowest fire points is performed a second time. Points that are not more than 30 cm above the new average 3D point are used to obtain the refined fire base plane equation by using a RANSAC algorithm (Fig. 4.40).



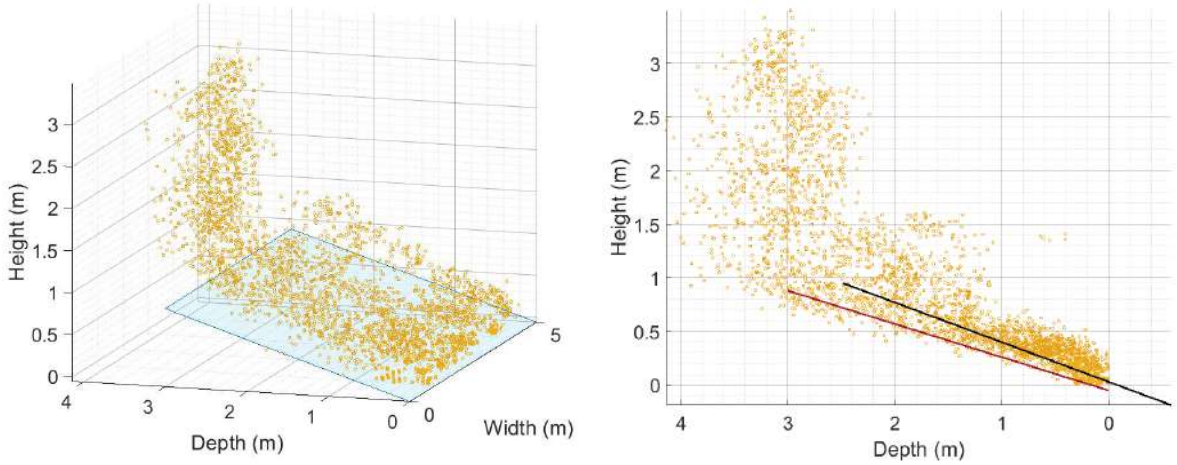

(a)




(b)

Figure 4.40: Refined estimation of the base plan from the 3D lowest points of a fire. (a) Lateral view. (b) Global view.

Figure **??** shows the base plane obtained after the refinement procedure. The longitudinal inclination line of the fire base plane is refined and improved compared to the inclination line obtained firstly (Fig. **??**).




(a) (b)

Figure 4.41: Final fire base plane. *(a)* Refined fire base plane and 3D fire points. *(b)* 3D fire points and inclination lines of estimated base planes: first plane (black color) and refined plane (red color).

A procedure is then carried out taking into account the base plane information obtained over time. At each image acquisition instant, it is estimated 3D fire points and part of them are used to compute the equation of the base plane (which angles are $\epsilon$ and $\theta$) corresponding to the ground plane. A procedure is carried out to improve the estimation of this information by considering 3D points obtained at successive times when it is possible. To initiate the algorithm, at the instant 0, a base plane of reference with its angles equal to 0° and with an empty set of 3D fire base points is considered. These reference angles are called $\epsilon_r$ and $\theta_r$. At each image acquisition instant $i$, the difference, in absolute terms, between the angle $\epsilon_i$ of the estimated base plane and the angle $\epsilon_r$ of the reference plane, and the difference in absolute terms between the angle $\theta_i$ of the base plane and the angle $\theta_r$ of the reference plane are computed.

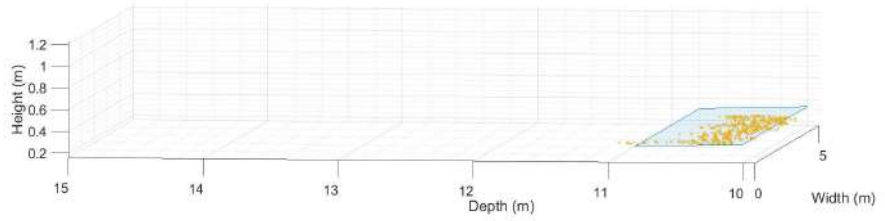
If the two calculated values are less than 3°, the 3D fire base points of the acquisition $i$ are added to the set of 3D fire base points of the reference plane.

$$|\epsilon_r - \epsilon_i| \leq 3^\circ \quad and \quad |\theta_r - \theta_i| \leq 3^\circ \tag{4.46}$$

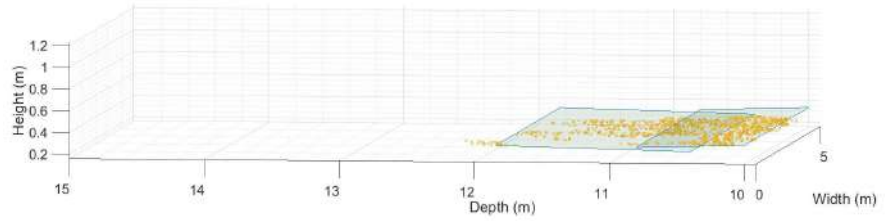
For each successive image acquisition $i$, until it is found a base plane whose angles satisfy the equation 4.46, the reference angles $\epsilon_r$ and $\theta_r$ doesn't change and the 3D fire base points are accumulated.

When it is found a base plane whose the angle differences are greater than 3°, two angles named $\overline{\epsilon}$ and $\overline{\theta}$ are computed considering all the set of base points previously accumulated starting from the instant of the reference plane (initially the instant 0). The angles of the current base plane become the new reference angles for the subsequent base planes.

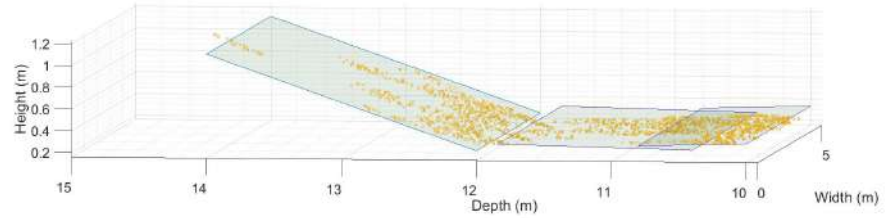
The number of instants considered to accumulate the 3D points, used to compute a base plane, is variable and depends on the instantaneous base plane angles. Considering $n$ point clouds, obtained at $n$ successive acquisition instants, the process gives as result a set of $m \leq n$ averages planes, where each one is identified by the mean angles values $\overline{\epsilon}$ and $\overline{\theta}$ and by the list of indices of the local planes that are part of it. By construction, these average planes are successive to each other and with inclination angles that differ from the ones of the previous average plane more than the threshold chosen, equal to 3°. Figure 4.42 shows with an example the successive estimation of base planes. Figure 4.42b presents two successive base planes with similar inclination. Figure 4.42c shows a third local plane, with an inclination variation greater than 3°. From the first two local planes an average plane is computed, and the third plane becomes the new reference plane for the subsequent planes (Fig. 4.42d). Figure 4.43 shows the result of the ground inclination method.



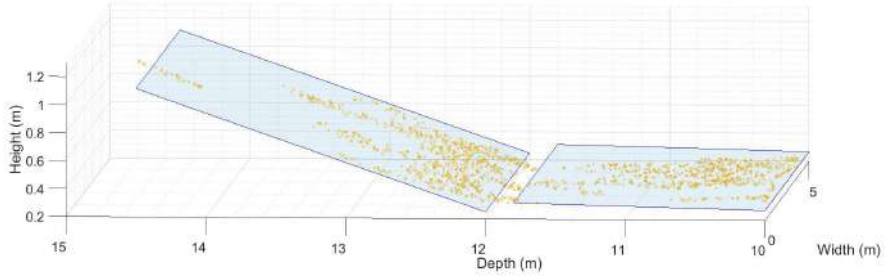

(a)




(b)




(c)




(d)

Figure 4.42: Example of successive estimation of base planes, in yellow are drawn the 3D points used to compute the local planes, in blue are presented the computed planes. *(a)* First fire base plane. *(b)* Two successive base planes with close inclination. *(c)* Three local planes obtained at successive instants, one has an inclination variation higher than more than 3°. *(d)* Average plane of the first two local planes.

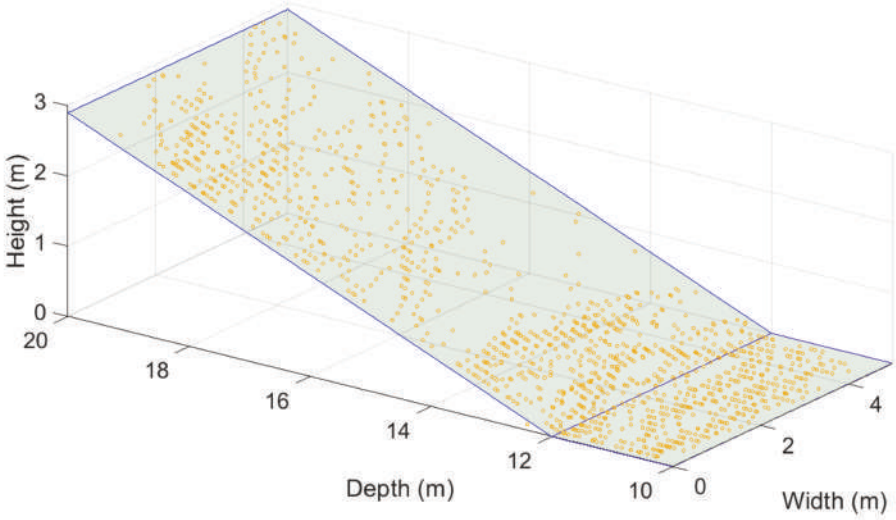



Figure 4.43: Result of the ground inclination method.

#### 4.5.3.2 Determination of the Slope frame

The Slope frame (S frame) is a frame with its origin $O_S$ equal to $O_G$, the $X_S$ x-axis parallel to the ground plane lateral slope, the $Z_S$ z-axis parallel to the ground plane longitudinal slope, and the $Y_S$ y-axis parallel to the normal of the ground plane (Fig. 4.44).

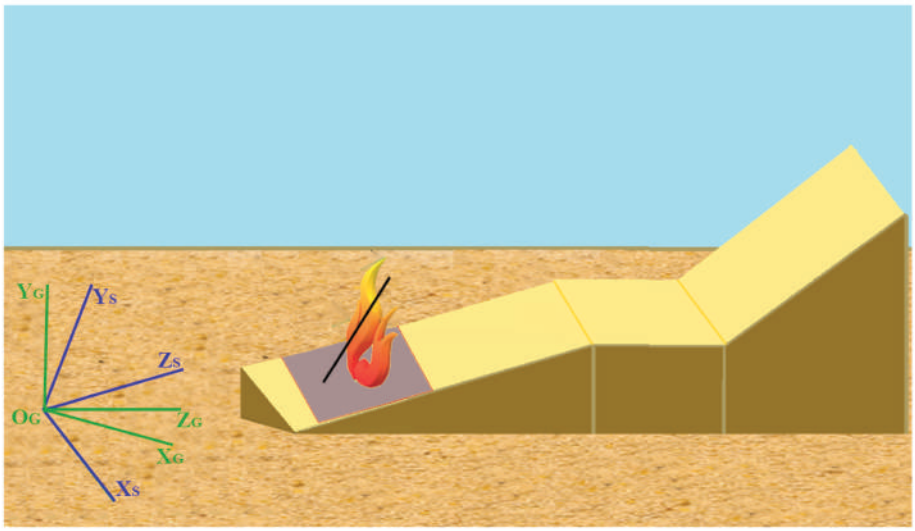



Figure 4.44: Slope frame and Global frame. The black line represents the normal of the ground.

This frame is considered at each moment of image acquisition (and 3D points calculation), and each average plane (characterized by its average angles $\overline{\epsilon}$ and $\overline{\theta}$) imposes an orientation

of the S frame. The Slope frame is obtained from the Global frame by performing a frame transformation whose equation is:

$$\begin{bmatrix} x_i^S \\ y_i^S \\ z_i^S \\ 1 \end{bmatrix} = M_5 \begin{bmatrix} x_i^G \\ y_i^G \\ z_i^G \\ 1 \end{bmatrix}. \tag{4.47}$$

where the matrix $M_5$ is:

$$M_5 = \begin{bmatrix} (R_{\overline{\epsilon}} * R_{\overline{\theta}})_{3\times3} & \mathbf{0}_{(3\times1)} \\ \mathbf{0}_{(1\times3)} & 1 \end{bmatrix} \tag{4.48}$$

with the matrix $R_{\overline{\epsilon}}$ and $R_{\overline{\theta}}$ defined by:

$$R_{\overline{\epsilon}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\overline{\epsilon}) & -sin(\overline{\epsilon}) \\ 0 & sin(\overline{\epsilon}) & cos(\overline{\epsilon}) \end{bmatrix} \tag{4.49}$$

$$R_{\overline{\theta}} = \begin{bmatrix} cos(\overline{\theta}) & -sin(\overline{\theta}) & 0 \\ sin(\overline{\theta}) & cos(\overline{\theta}) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.50}$$

This transformation is applied to the 3D points coordinates obtained at a given time $i$ expressed in the G frame ($x_i^G$, $y_i^G$, $z_i^G$) to obtain the 3D points of coordinates ($x_i$,$y_i$,$z_i$) in the S frame.

The Eq. 4.47 produces as results the fire points rotated as if the ground was flat (Fig. 4.45).

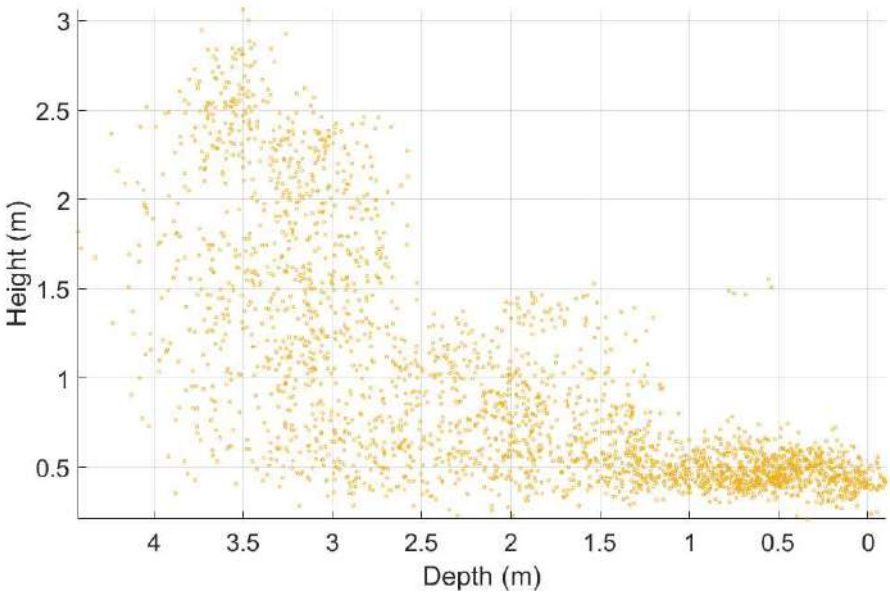



Figure 4.45: 3D points positioned as if the ground was flat.

### 4.5.4 Determination of the transformation matrix from the Slope frame to the Local Direction frame

During its propagation the fire can change its direction, and to correctly calculate the dimension of the geometrical characteristics such as the width, depth, length, surface and inclination of the fire front, it is important to rotate the reference system so that the depth axis is parallel to the instantaneous direction of the fire. An example of a local direction is shown in the Fig 4.46.

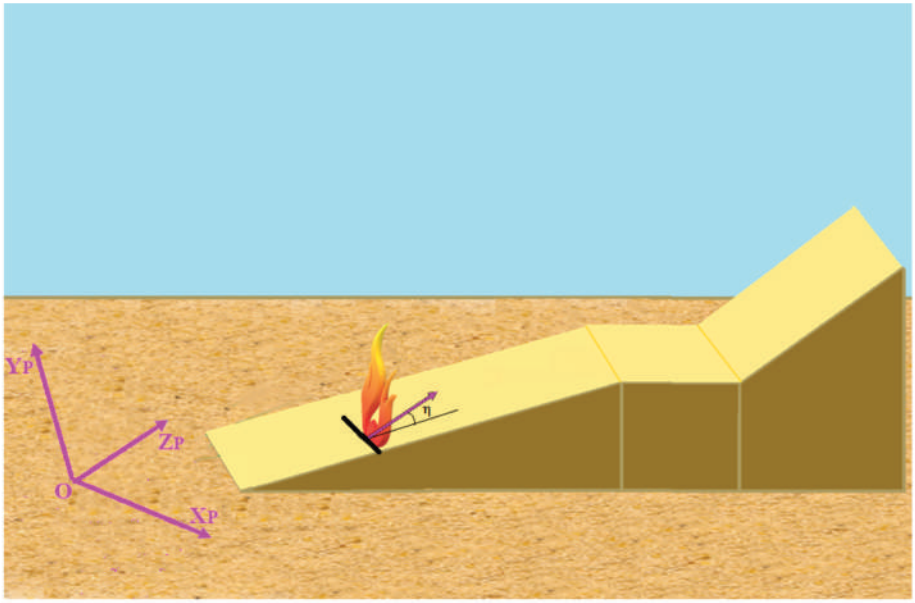



Figure 4.46: Local direction of the fire front. The black line represents the instantaneous fire front line, the purple arrow represents the instantaneous fire direction, $(O, X_P, Y_P, Z_P)$ represents the Local Direction frame.

To calculate this direction, they are considered the sets of base points of two successive acquisition. For each set of points it is computed its centroid. Figure 4.51 shows an example of centroid of one fire ground points.

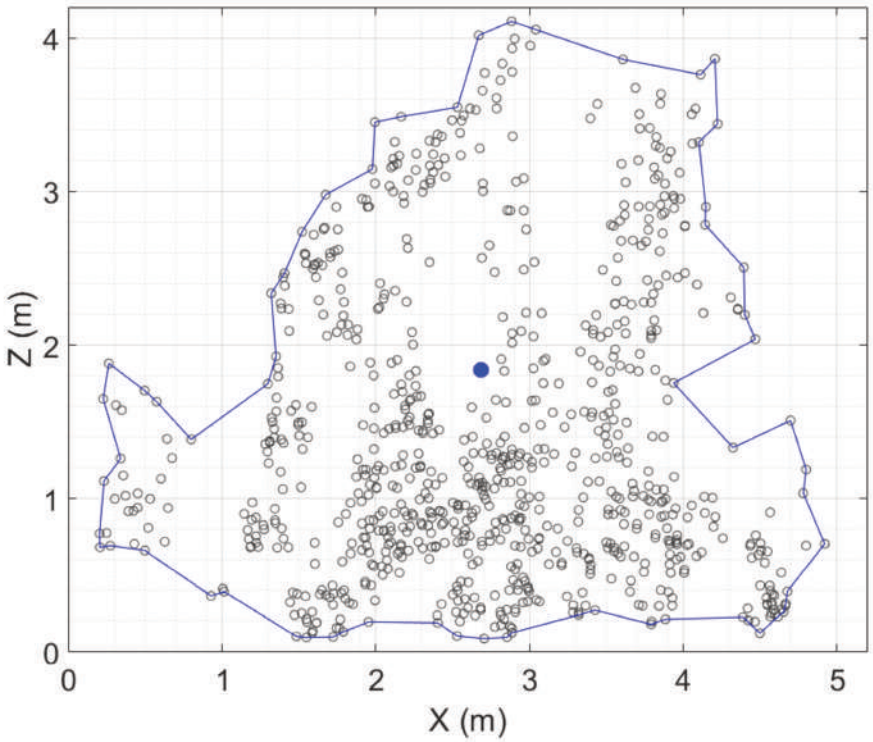



Figure 4.47: Position of the fire ground points and its centroid.

Figure 4.48 shows the centroids of two successive fire ground points. Let $\eta$ the angle between the straight line that connects the two centroids and the $z$ axis, it represents the local direction of the fire.

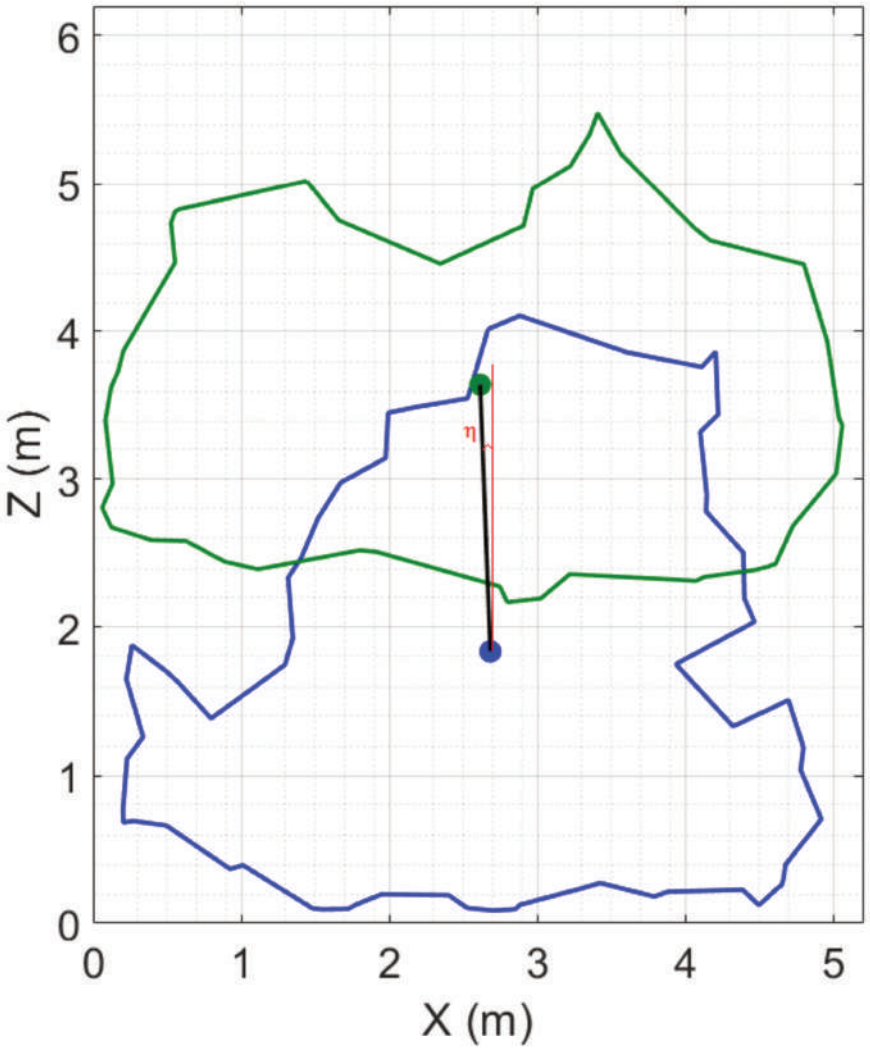



Figure 4.48: Direction of fire propagation. The angle $\eta$ represents the local direction angle of the fire.

The 3D fire points are rotated around the y-axis by the angle $\eta$ using the rotation matrix:

$$M_6 = \begin{bmatrix} cos(\eta) & 0 & -sin(\eta) \\ 0 & 1 & 0 \\ sin(\eta) & 0 & cos(\eta) \end{bmatrix} \tag{4.51}$$

The coordinates of the 3D points expressed in the Slope frame $(x_i^S, y_i^S, z_i^S)$ are expressed in

the Local Direction frame $(x_i^P, y_i^P, z_i^P)$ by the equation:

$$\begin{bmatrix} x_i^P \\ y_i^P \\ z_i^P \\ 1 \end{bmatrix} = M_6 \begin{bmatrix} x_i^S \\ y_i^S \\ z_i^S \\ 1 \end{bmatrix}. \tag{4.52}$$

## 4.6 Fire geometrical characteristics estimation

This section presents the estimation of fire geometrical characteristics like the fire front position, fire local direction, rate of spread, depth, width, fire base perimeter, combustion surface, length, height, fire front area, and fire front volume; from the 3D fire points projected so that the corresponding base plane has a zero slope.

### 4.6.1 Fire front position

The 3D fire points considered for the estimation of the front position, local direction, base perimeter and base area, width and depth are the lowest points positioned on the ground (Fig. 4.49).

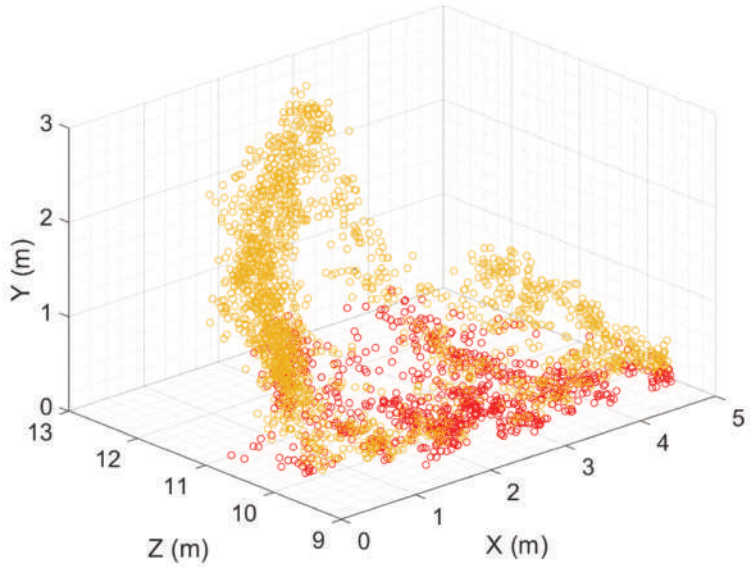



Figure 4.49: 3D fire points (yellow circles), and 3D fire ground points (red circles).

Fig. 4.50 shows the fire points that are on the ground and for which $y = 0$.

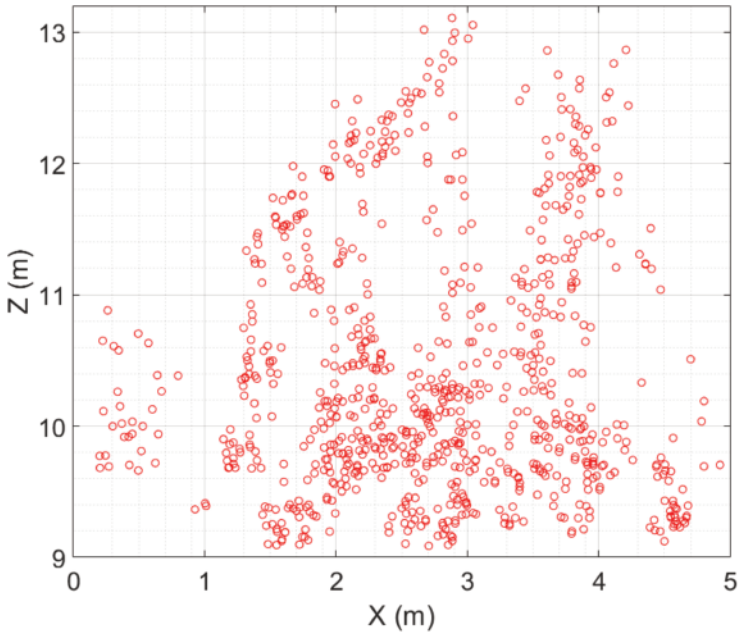



Figure 4.50: Ground fire 3D points used to estimate the fire front position, on the $X$-$Z$ plane.

To find the front part and the back part of the fire base, the 3D points are sorted according to their $x$ coordinate; regions of 15 cm along the $x$ axis are considered separately. It has been choosen to work considering areas 15 cm wide rather to prevent the influence of a single point in the front line form, as it is shown in Fig. 4.51.

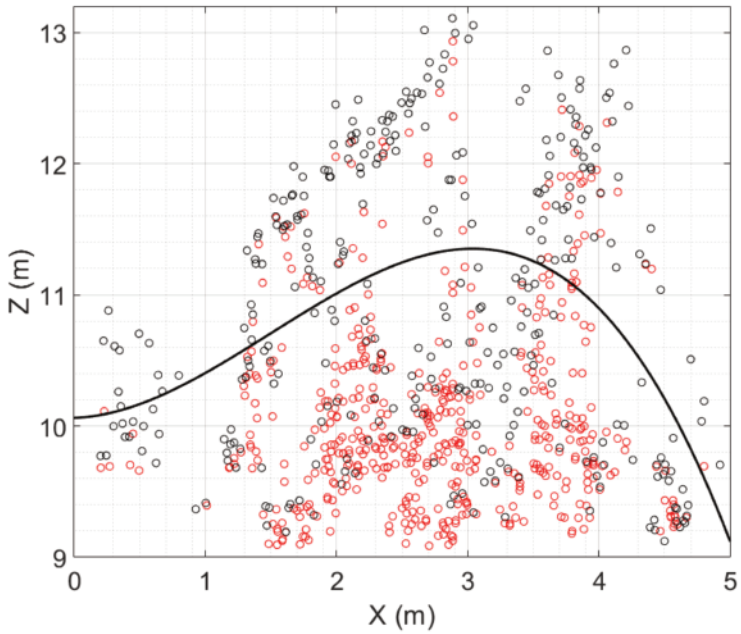



Figure 4.51: Line front form obtained from the most advanced ground points of the fire front.

In each 15 cm wide sector, only two points are selected corresponding respectively to the most advanced point named $Pts_{Front}$ (with the highest $z$ coordinate) and the less advanced point named $Pts_{Back}$ (with the smallest $z$ coordinate).

The union of the back fire points $Pts_{Back}$ and the front fire points $Pts_{Front}$ obtained by considering all the sectors of the width of the fire front corresponds to the the fire base contour. The front line and the back line of the fire front are respectively computed considering the points $Pts_{Front}$ and the points $Pts_{Back}$ and using a B-spline interpolation with a polynomial function of 3 degrees [71]. Figure 4.52 shows the front line (black color) and the back line (green color) of the fire.

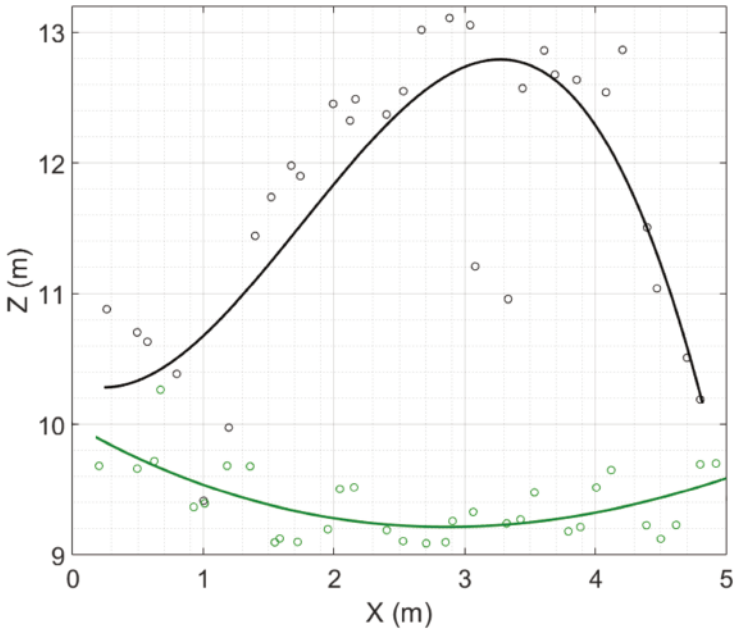



Figure 4.52: Back line (green color) and front line (black color) of the fire obtained from the $Pts_{Back}$ points and the $Pts_{Front}$ points, respectively.

### 4.6.2 Point rate of spread

The front lines of the fire front on the base plane obtained at different times are used to estimate the rate of spread of the front points. This rate of spread is the ratio of the distance between two equivalent points on two successive lines divided by the time interval between the two acquisition moments of the images from which the curves were calculated. Considering two successive front lines ($l_1$ and $l_2$) and the point $P_1$ on the line $l_1$, the equivalent point to $P_1$ named $P_2$ situated on $l_2$ is the intersection point between the normal of $l_1$ in $P_1$ and the line $l_2$.

In this thesis, for each fire front line it has been chosen one point each meter along the $x$ coordinates. Figure 4.53 shows an example of equivalent points on two successive front lines. To simplify the research of the intersection points, the front lines are approximated with a parabolic function.

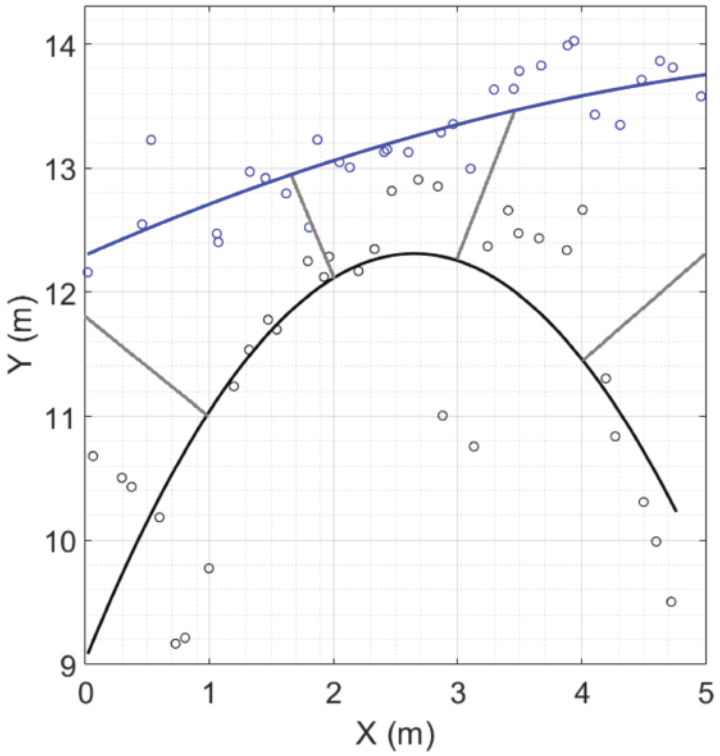



Figure 4.53: Equivalent points on two successive front lines, the black line is the first front line ($l_1$) and the blue line is the second front line ($l_2$).

The rate of spread of the front fire line $l_1$ is the average of the rates of spread of the points of this line.

### 4.6.3 Fire base perimeter and combustion surface

Researchers working on modelling forest fires consider that the fuel below a flame contributes to the combustion and approximates the surface by simple forms [134]. The surface of the vegetation in combustion is the area of the surface delimited by the perimeter of the fire.

First it is computed the alpha shape [135] of the fire ground points projected onto the $X$-$Z$ plane (Fig. 4.55), the radius alpha of the shape it is chosen as the smallest radius that form a shape that include all the points.

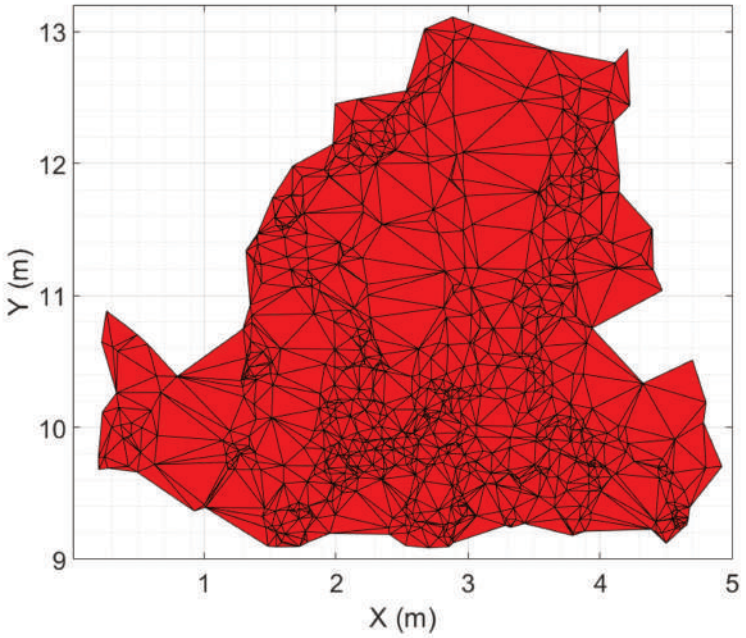



Figure 4.54: Alpha shape of the ground points.

From the alpha shape it is constructed a polygon computing the boundary points [136], and the perimeter of the fire is estimated as the sum of the euclidean distance between successive boundary points. The area of the polygon is the estimated combustion surface (Fig. 4.55).

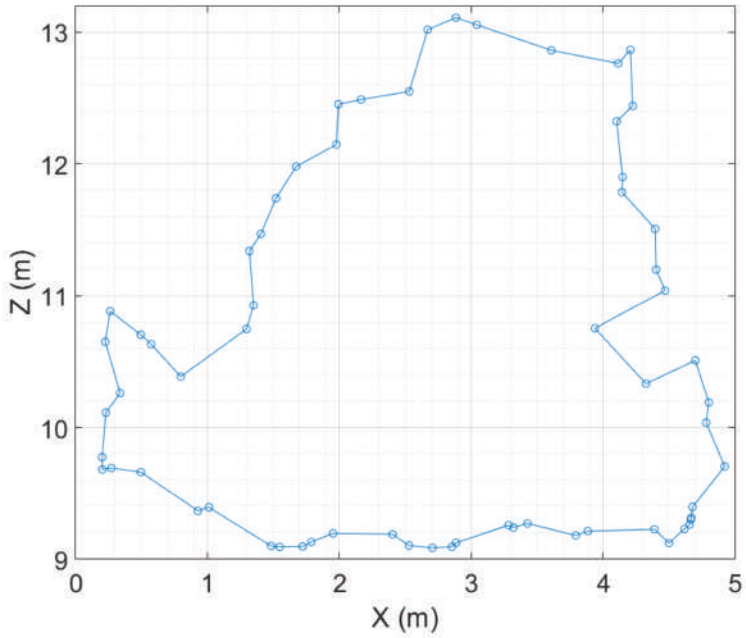



Figure 4.55: Boundary points of the alpha shape (blue circles). The blue line represents the fire ground perimeter.

### 4.6.4 Fire depth

From the $Pts_{Front}$ points set, it is computed a mean front point ($P_F$); similarly, from the $Pts_{Back}$ points set, it is computed a mean back point ($P_B$). The depth measure is the euclidean distance between $P_B$ and $P_F$ (Fig. 4.56). It is computed as follow:

$$fire\ depth = \sqrt{(P_{B_x} - P_{F_x})^2 + (P_{B_z} - P_{F_z})^2} \tag{4.53}$$

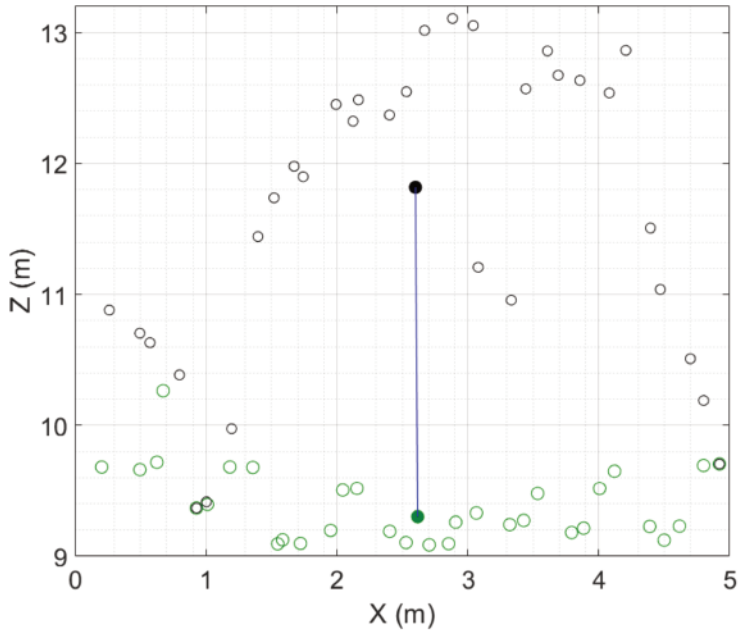



Figure 4.56: Depth of the fire front computed as the distance between the mean point of the front points and the mean point of the rear points.

### 4.6.5 Fire width

The width measure is obtained from the 3D fire ground points expressed in the Local Direction frame.

First, the point with the minimal $x$ coordinate (named $P_m$) and the point with the maximal $x$ coordinate (named $P_M$) are identified. Two lateral zones, called left zone and right zone, are defined in a similar way to that seen previously regarding the identification of the fire base. In particular, the left zone is composed by the points such that the $x$ coordinate is not more than 15 cm from $P_m$, the right zone is composed by the points such that the $x$ coordinate is not more than 15 cm from $P_M$.

Inside these two zones they are computed two average points, $\tilde{P}_m$ is the average point of the fire ground points present in the left area, and $\tilde{P}_M$ is the average point of the fire ground points present in the right area (Fig. 4.57). The width measure is the euclidean distance between $\tilde{P}_m$ and $\tilde{P}_M$, and it is computed as follow:

$$fire\ width = \sqrt{(\tilde{P}_{m_x} - \tilde{P}_{M_x})^2 + (\tilde{P}_{m_y} - \tilde{P}_{M_y})^2} \tag{4.54}$$

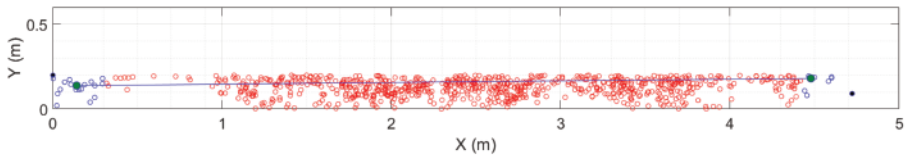



Figure 4.57: Fire width: distance from the fire ground points $\tilde{P}_m$ and $\tilde{P}_M$.

### 4.6.6 Fire height

The height of the fire front is obtained using the 3D fire points projected in the $Z$-$Y$ plane.

In order to avoid that only one point greatly influences this measure, the fire height is computed as the average height of the "top" region of the 3D points. In particular, the 3D point with the maximum $y$ coordinate $y_M$ is first identified. Starting from this point, they are selected all the 3D points such that the $y$ coordinate is not more than 30 cm below $y_M$. In this region of 3D point the average point $P_H$ is estimated; the $y$ coordinate of the point $P_H$ is the height of the fire (Fig. 4.58).

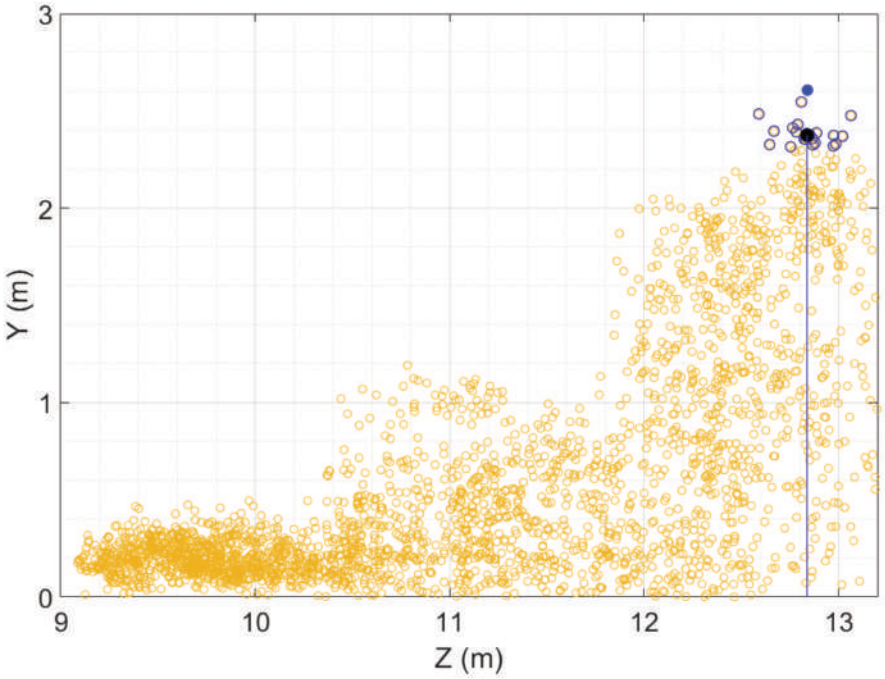



Figure 4.58: Height of a fire front considered to be the average point (black circle) of the 3D points located within the 30 cm highest of the flame (blue empty circles).

### 4.6.7 Frontal flame length and flame inclination angle

The frontal flame length is the euclidean distance between the point $P_H$ (the average point of the 3D points located within the 30 cm highest of the flame) and the point $P_F$ (the average point of the ground 3D points located in the front part of the fire) (Fig. 4.59). It is computed using the following formula:

$$flame\ length = \sqrt{(P_{H_x} - P_{F_x})^2 + (P_{H_y} - P_{F_y})^2 + (P_{H_z} - P_{F_z})^2} \tag{4.55}$$

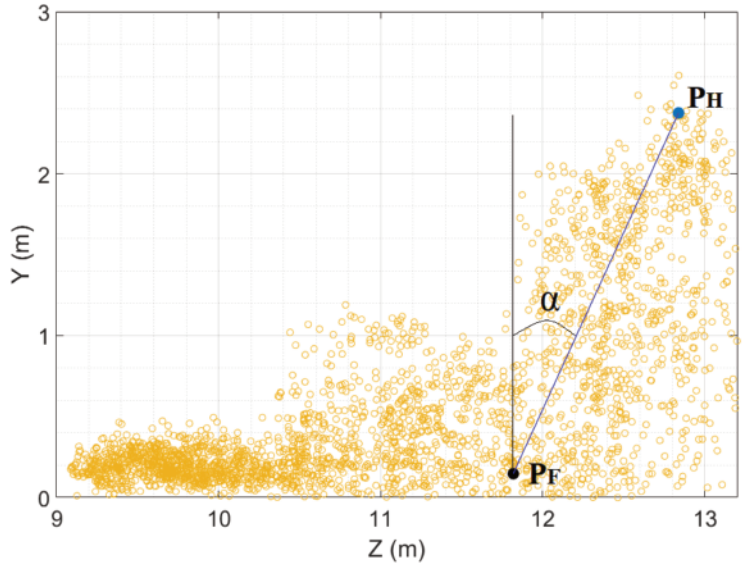



Figure 4.59: Front flame length (blue line) and inclination angle $\alpha$.

The flame inclination is defined as the angle between the segment $\overline{P_H P_F}$ and the normal of the fire base plane. Defining the fire base plane with the equation:

$$ax + by + cz + d = 0 \tag{4.56}$$

and the directing parameters ($l$,$m$,$n$) of the segment $\overline{P_H P_F}$ as:

$$l = P_{H_x} - P_{F_x} \quad m = P_{H_y} - P_{F_y} \quad n = P_{H_z} - P_{F_z}, \tag{4.57}$$

The flame inclination angle, in degree, is obtained by:

$$flame\ inclination = 90° - \frac{\mid al + bm + cn \mid}{\sqrt{a^2 + b^2 + c^2}\sqrt{l^2 + m^2 + n^2}} \tag{4.58}$$

### 4.6.8 Flame volume

The volume of the fire is estimated from the shape shown in Fig. **??**. The overall volume of the form is the sum of the volumes of the elementary tetrahedrons whose volume is estimated as follows. Let $A$, $B$, $C$, $D$ the 4 vertices of one elementary tetrahedron $i$, and $d_{AB}$, $d_{AC}$, $d_{AD}$, $d_{BC}$, $d_{BD}$, $d_{CD}$ the the six edges of a tetrahedron (Fig. 4.60).

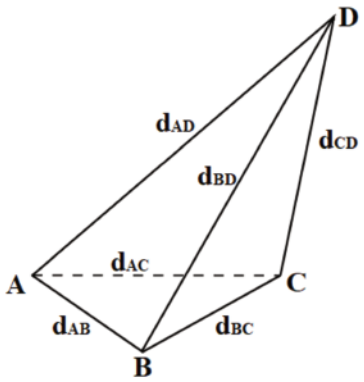



Figure 4.60: Tetrahedron identified by its 6 distance between the 4 vertices.

The volume $V$ is calculated using the Cayley–Menger determinant of third order [137]. In mathematical form, the volume of each tetrahedron is equal to:

$$V = \sqrt{\frac{1}{288}\det\begin{vmatrix} 2d_{AB}^2 & (d_{AB}^2 + d_{AC}^2 - d_{BC}^2) & (d_{AB}^2 + d_{AD}^2 - d_{CD}^2) \\ (d_{AB}^2 + d_{AC}^2 - d_{BC}^2) & 2d_{AC}^2 & (d_{AC}^2 + d_{AD}^2 - d_{BD}^2) \\ (d_{AB}^2 + d_{AD}^2 - d_{CD}^2) & (d_{AC}^2 + d_{AD}^2 - d_{BD}^2) & 2d_{AD}^2 \end{vmatrix}} \tag{4.59}$$

### 4.6.9 Surface and view factor estimation

Considering the set of triangles that form the surface of the fire, it is possible to calculate the surface that produces the heat flow in front of the fire. It is also possible to calculate the fraction of the total energy emitted by the surface of the fire and received by a target whatever its position (also called view factor).

To calculate the surface of the fire, only the surfaces of the triangles which are not masked by the others and oriented in the main direction of the fire are added.

The view factor is estimated using all the fire triangles and the 3D coordinates of the target. The method used is that described in [69]. The view factor $F_{S \longrightarrow dT}$ for the radiation, between the entire surface of the flame $S$ and the target area $dT$, is defined as the fraction of the total energy emitted by all the surfaces of the elementary triangles, $dS_i$, and received by $dT$. Let $r_i$ be the distance between the target area and the center of the triangle $i$-th of the surface $dS_i$, $\theta_{S_i}$ the angle between $r_i$ and the surface normal of the elementary triangle, and

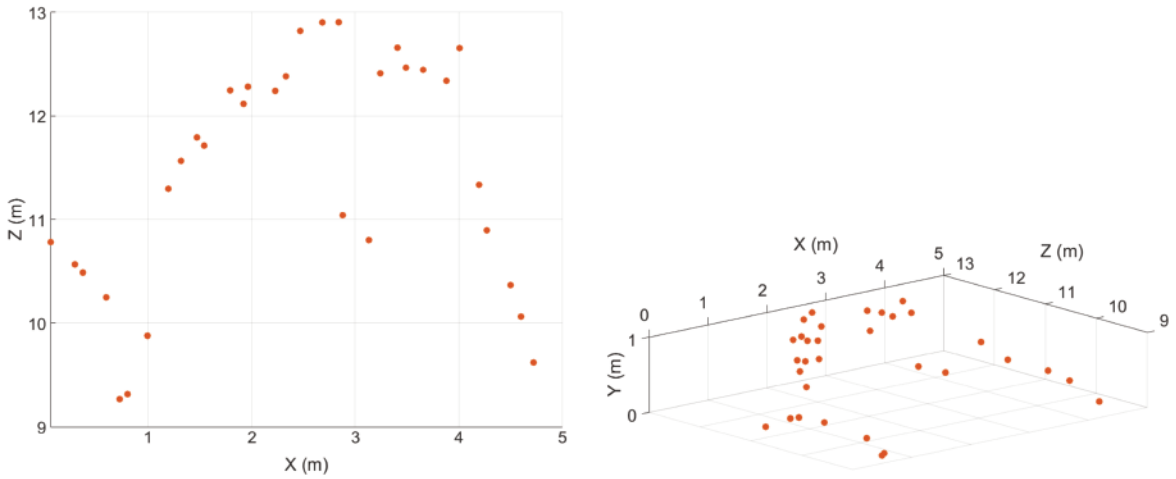
$\theta_{T_i}$ the angle between $r_i$ and the surface normal of the target. The view factor is defined by:

$$F_{S \longrightarrow dT} = \frac{dT}{S} \sum_{i=1}^{N} \frac{\cos\theta_{S_i} \cos\theta_{T_i}}{\pi r_i^2} dS_i, \tag{4.60}$$

where are considered only the angles $\theta_{S_i}$ ans $\theta_{T_i}$ between $-\pi/2$ and $\pi/2$.

### 4.6.10 GPS position of the fire front line

Using the Cartesian coordinates of the 3D fire points, the GPS position of the stereovision device, and the orientation of the compass, it is possible to calculate the positions of the 3D fire points in GPS coordinates (latitude longitude). These positions can be added on a GIS map, for example by formatting a Keyhole Markup Language (.KML) files. To obtain a simple and clear view on the map, it was chosen to calculate the GPS positions only of the points of the front line. Figure 4.61 shows the 3D fire front points in the plane (X-Z) and in 3D space.




(a) (b)

Figure 4.61: 3D fire front points. ($a$) 2D plane. ($b$) 3D space.

The 3D points are expressed in the frame L. A first transformation rotates the 3D points around the y axis using the rotation matrix $R_\delta$ with the $\delta$ angle (the angle returned by the compass); such that the z axis is oriented towards the terrestrial North. The GPS position of the stereovision system is converted into Cartesian coordinates ($x_d$, $y_d$, $z_d$) expressed in the ECEF reference system. The 3D points are then translated by the vector $[x_d, y_d, z_d]$. Figure 4.62 shows the 3D fire front points expressed in the ECEF frame.

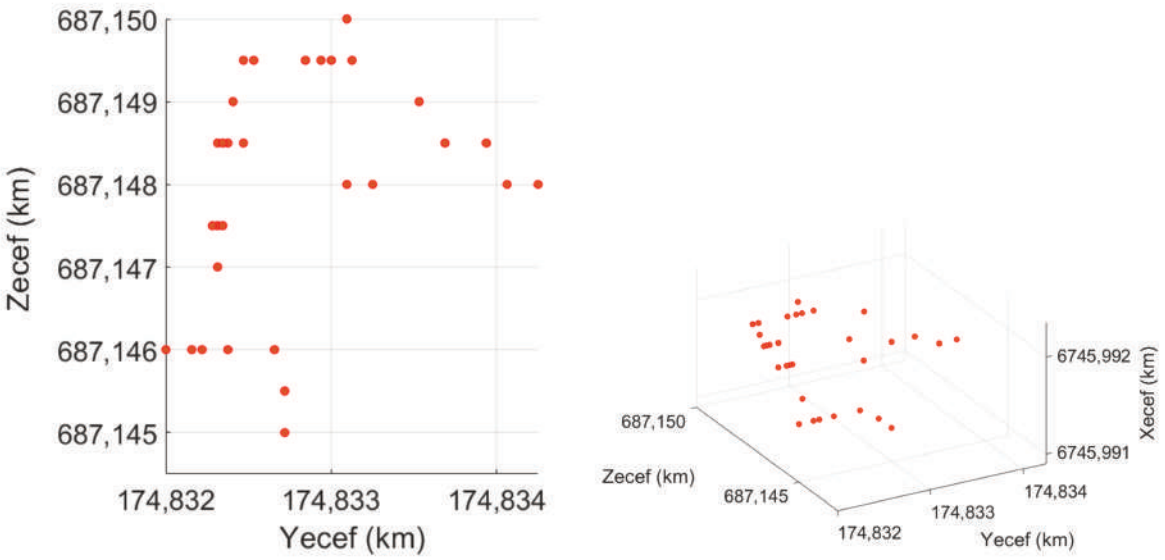



Figure 4.62: 3D fire front points expressed in the ECEF frame. (*a*) 2D space. (*b*) 3D space.

The GPS coordinates of the 3D points are computed using the Browing's method as presented in [138]. Figure 4.63 shows the GPS coordinates of the 3D fire front points over a GIS map.


Figure 4.63: Fire front points over a GIS map.

Chapter 5

# Measurement uncertainty

**Contents**



## 5.1 Introduction

Due the unpredictable and non-reproducible behaviour of the fire, it is difficult to evaluate the uncertainty of the measurements of geometric characteristics of a fire by vision. To obtain this uncertainty, the performance of the vision system was evaluated by carrying two tests. The first one using a car. Only visible images were processed and, such as the method used for fire pixel detection was not usable, the car features were selected by hand in the images. From the 3D car points, the dimensions of the car were computed and compared with the real ones.

A second test was made on a $3 \times 5$ m pseudo-static fire. The form of the fire was changing but not its position. This second experience allowed to test the entire stereovision system, using images in the LWIR and visible domain, to obtain a beta test of the complete device.

## 5.2 Car test

In this test, the UAV made a complete turn around a parked car, to acquire pictures from each car side, with a inter-distance UAV-car between 10 m and 15 m (Fig. 5.1). The position of the car was measured with a GPS sensor positioned on the roof of the vehicle, at the base of the antenna of the car.

(a) (b)

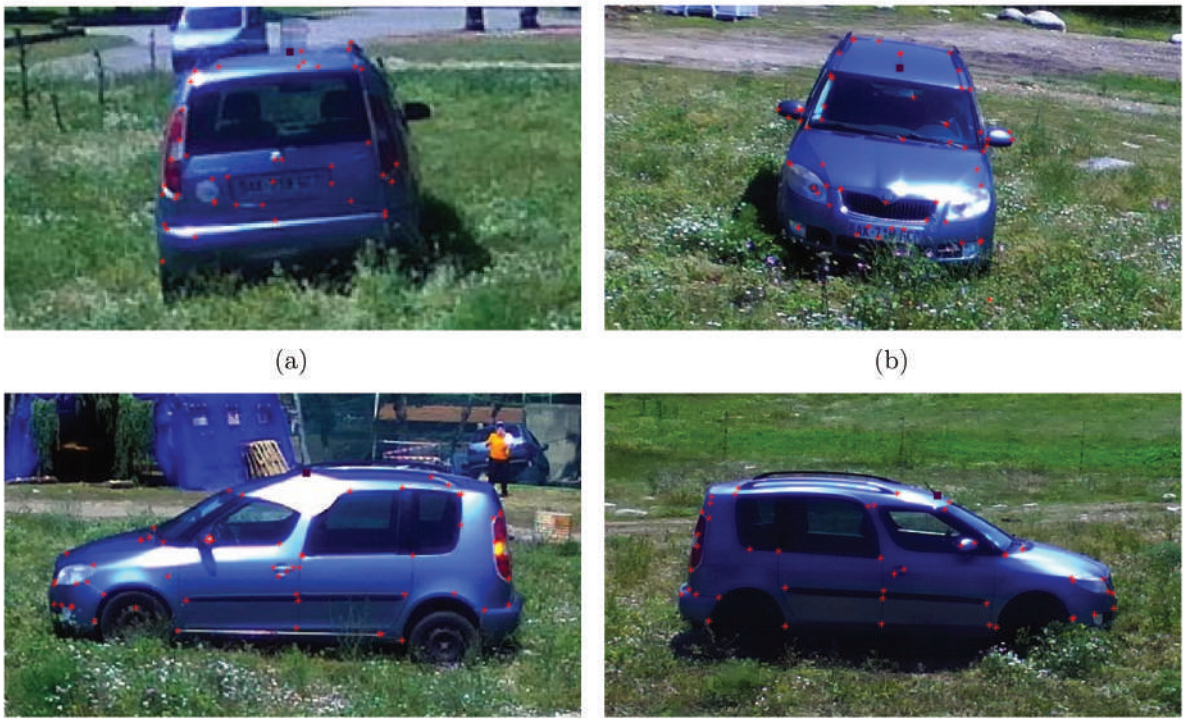
(c) (d)

Figure 5.1: Car images used for the measurement of the uncertainty. (*a*) Back side. (*b*) Front side. (*c*) Left side. (*d*) Right side.

Figure 5.2 shows the GPS position of the drone during the test, and the position of the GPS of the car.


Figure 5.2: Drone and car positions registered by GPS sensors. The light blue icons represent the positions of the drone during the test, the green icons represent the positions of the drone when the picture is captured. The dark blue icon represents the position of the car.

From each car side image, features were selected manually in order to measure the car dimensions and its shape (not to obtain a complete 3D shape): 42 for the back side, 59 for the right side, 66 for the front side and 47 for the left side. For each side, there was a point that represented the GPS sensor on the top of the car (the base of the car antenna); the Fig. 5.3 shows these points.



(a) (b)

(c) (d)

Figure 5.3: Points of the car selected by hand (red stars). The purple square represents the point where the GPS sensor was positioned. (*a*) Back points. (*b*) Front points. (*c*) Left points. (*d*) Right points.

Figure 5.4 shows the 3D car points obtained separately from the four complementary car views. From each set of points a surface was computed using the Delaunay method (Fig. 5.5). Figure 5.6 presents the complete 3D reconstruction considering all the 3D car points in the same set.

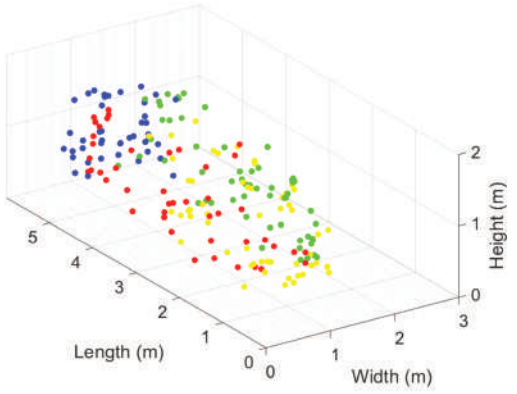



Figure 5.4: 3D points of the back side (blue color), front side (yellow color), left side (red color), and right side (green color).

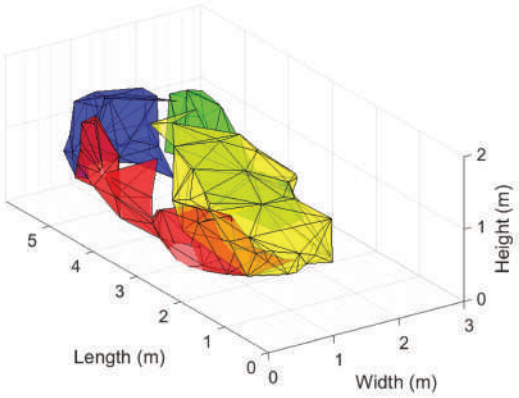



Figure 5.5: Delaunay triangulation surfacing the four sides of the car.

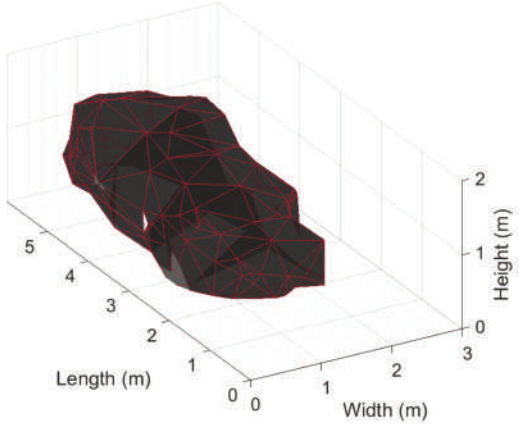



Figure 5.6: Complete 3D reconstruction of the car.

From the 3D car shape are estimated the width, height and length of the car. To estimate

the car position, the 3D points of the car antenna was estimated for each car side. Figure 5.7 shows an example of the 3D point of the car antenna, which appears in the reconstruction of the front part of the car.

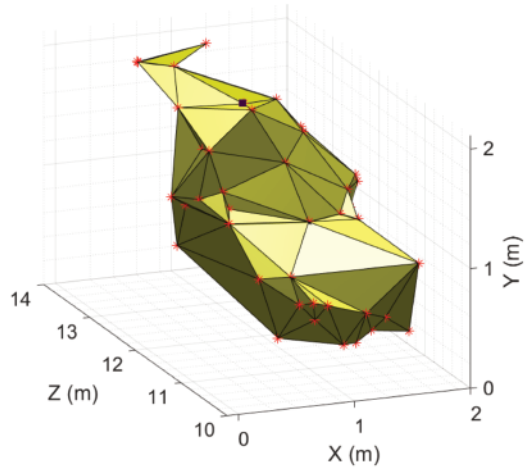



Figure 5.7: Front side of the car, the purple square represents the 3D point of the antenna.

The 3D point of the antenna was converted from Cartesian coordinates to GPS polar coordinates, and inserted into a map (Fig 5.8).

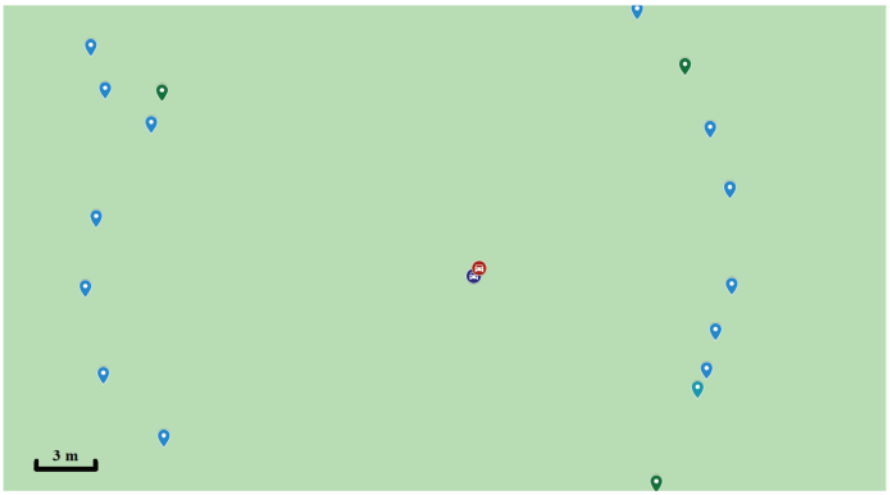



Figure 5.8: Car position obtained from the GPS sensor (blue icon), and computed from the 3D point (red icon).

Table 5.1 presents the comparison between the data estimated by stereovision and the real ones (dimensions of the car given by the manufacturer and position of the car obtained with a sensor GPS.

| | Length (m) | Width (m) | Height (m) | Position (°lat;°lon) |
|---|---|---|---|---|
| Real | 3,99 | 1,64 | 1,50 | (42,299 991 1°;9,175 529 1°) |
| Estimated | 3,96 | 1,62 | 1,48 | (42,299 992 3°;9,175 525 9°) |
| Error | 0,7% | 1,2% | 1,1% | 26 cm |

Table 5.1: Comparison between real and estimated measurements of the car.

## 5.3 Pseudo static fire test

The second test was conducted on an outdoor pseudo-static fire. Wood wool was placed on an area of $3 \times 5$ meters, and the drone was moving around the fire. The shape of the fire could change, but not its position (Figure 5.9).

Figure 5.9: Images of the pseudo-static fire at two different moments.

From the 3D fire points that are presented in Fig. 5.10, the position of the front line is estimated.

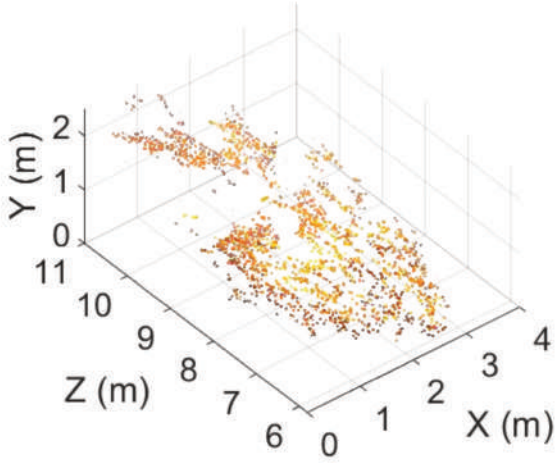



Figure 5.10: 3D fire points at a generic instant.

Figure 5.11 shows an example of 3 front lines obtained at 3 different instants. The shape of the lines changed over time, but it can be seen that the fire stays at the same position.

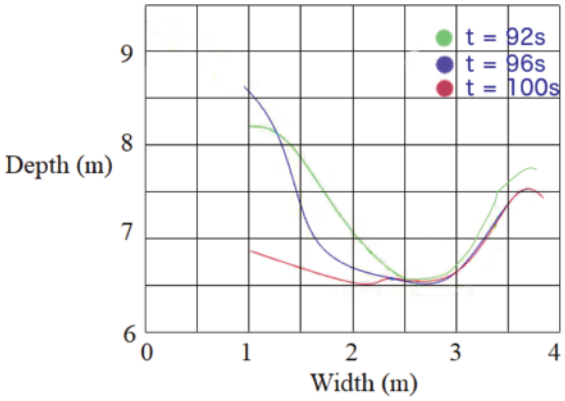



Figure 5.11: Position of the fire front at different times.

Chapter 6

# Monitoring of fire propagation with the drone vision device

Contents



## 6.1 Introduction

This chapter presents the estimated data obtained during the monitoring of an experimental fire propagation.

The experiment was conducted on the technical platform of the "Unité d'Instruction et d'Intervention de la Sécurité Civile N.5" of Corte. An area of $5 \times 10$ m consisting of a first flat zone of 2 m and a second zone with slope at 20° was filled with wood wool (Fig. 6.1).

Figure 6.1: Configuration of the fire test area.

## 6.2 Description of the experiment

The drone flew at a height of 10 m, keeping at a distance of 15 m from the fire zone. The inclination of the stereovision system was set to -3° downwards in such a way to obtain the best shooting angle at 15 m from the fire. The fire was set along the short side of the rectangle to produce propagation over the entire length of the fuel area along the long side of the rectangle (Fig. 6.2).

Figure 6.2: Fire line ignition.

Figure 6.3 shows the temporal evolution of roll and pitch angles of the stereovision device. The roll remains around 0 radiant, with a margin of more or less 0,05 radiant. The average pitch angle is 0,05 radiant.

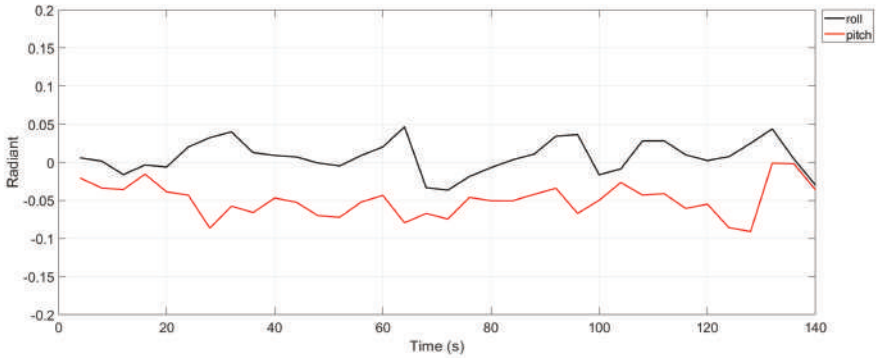



Figure 6.3: Roll and pitch angle of the stereovision device.

Given the size of the fuel area, the drone made only small movements (Fig. 6.4), and to avoid having blurred images, the drone was moved with low speed (Fig. 6.5).

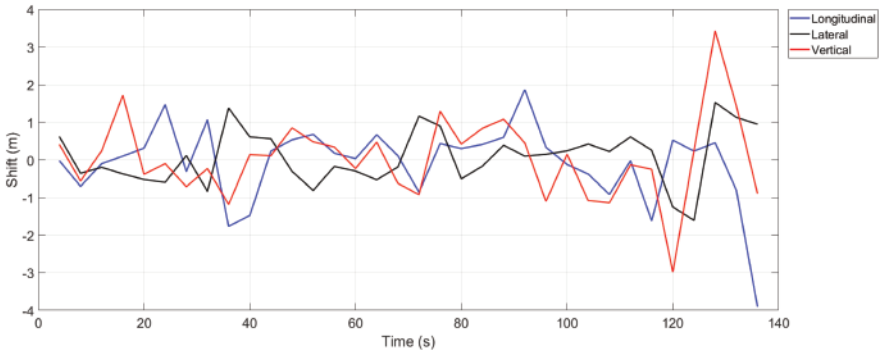



Figure 6.4: Position shift of the stereovision device.

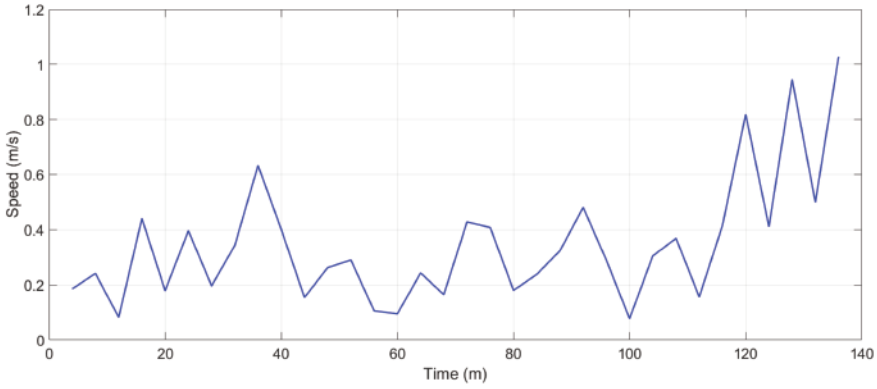
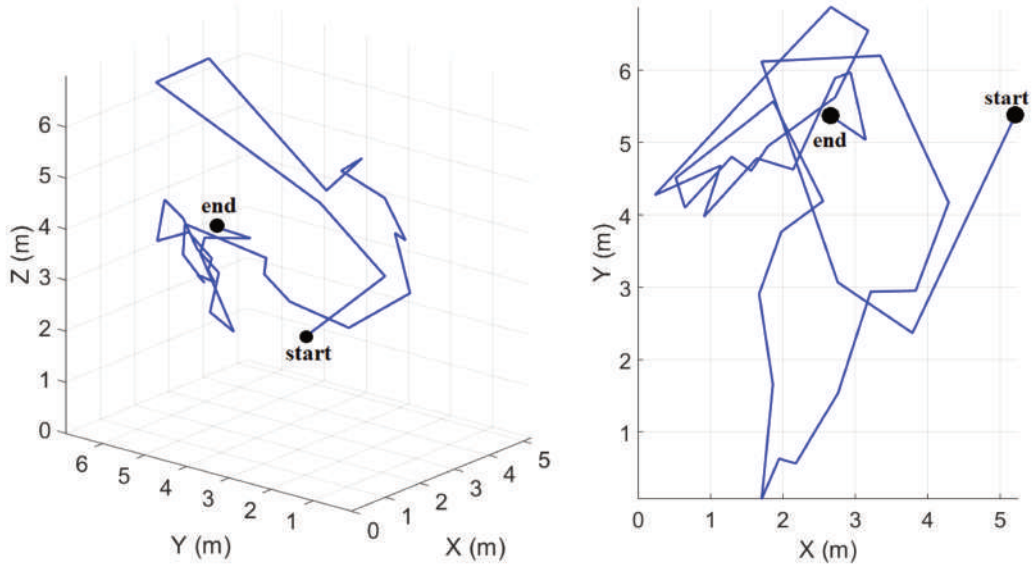



Figure 6.5: Drone speed.

Figure 6.6 shows the Cartesian positions of the stereovision system during the experimentation. These positions corresponds to the GPS positions shown in Fig. 6.7. The orientation of the drone varies in accordance with the direction of focus during the fire propagation (Fig. 6.8).




(a) (b)

Figure 6.6: Position of the stereovision device during the test. The $x$ axis is direct to East, the $y$ axis is direct to North, and the $z$ axis is direct up ($a$) 3D space. ($b$) 2D space.


Figure 6.7: GPS positions of the drone during the experimentation (blue circle). The red rectangle represents the fuel area.

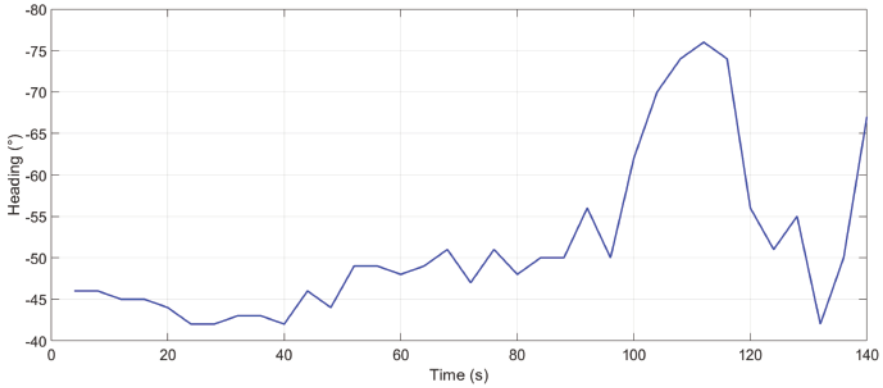



Figure 6.8: Orientation angle of the stereovision device.

## 6.3 Temporal evolution of fire geometrical characteristics

Figure 6.9 shows the temporal evolution of the fire front position obtained for this propagation.

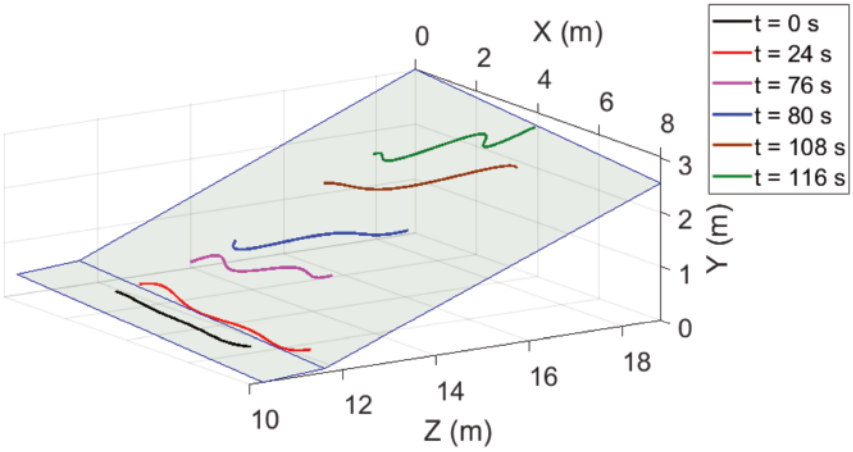



Figure 6.9: Evolution of the positions of the fire on the estimated planes.

As it can be seen, the framework models the ground with two parts, one has a longitudinal inclination equal to 0°, the second has a longitudinal inclination equal to 21°; both has a lateral inclination of 3°. These values are in accordance with the real characteristics of the ground.

The temporal evolution of the fire position is used to identify the fire regimes. A regime is characterised by a constant rate of spread. Considering the Fig. 6.10, it appears that the fire

had 3 regimes.

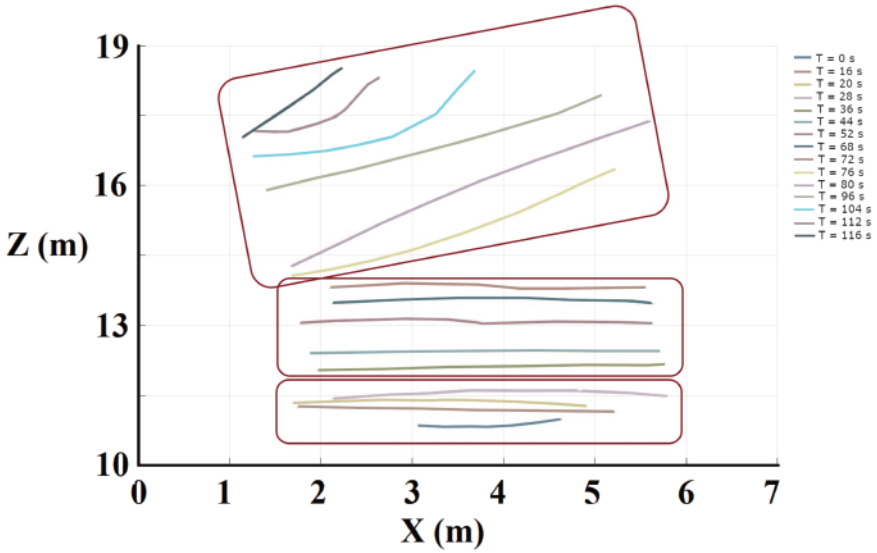



Figure 6.10: Evolution of the positions of the fire on the plane $X$-$Z$.

The first regime is between 0 and 28 seconds, characterized by a low rate of spread which results in close position lines, as can be expected from a fire that starts. The second regime is between 36 seconds and 72 seconds, characterized by an increasing of the rate of spread, this is materialized by a larger space between the position lines. This is due to the fact that the fire takes on dynamics and starts to rise along the inclined part of the terrain. The propagation direction during the first two regimes is parallel to the longitudinal direction of the focus area. The third regime is between 76 seconds and 116 seconds. In addition to a greater distance between the lines, this regime is characterized by a change of direction of the front of the fire, due to a change in wind direction.

Fig. 6.11-6.14 show the temporal evolution of the height, depth, width, flame tilt inclination, combustion surface, and rate of spread.

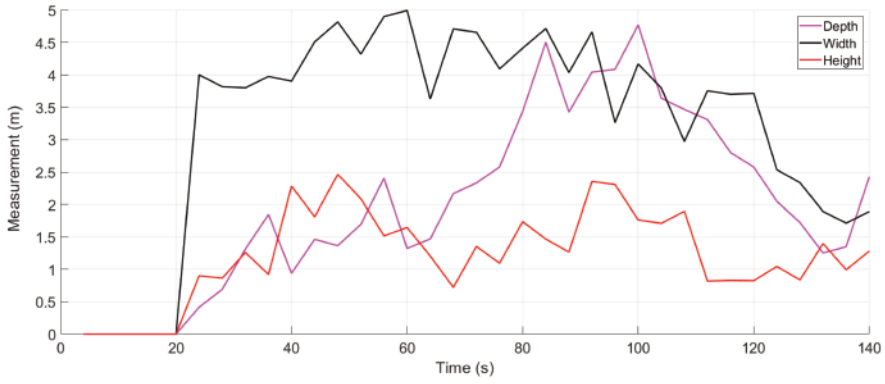



Figure 6.11: Temporal evolution of the width, height, and depth.

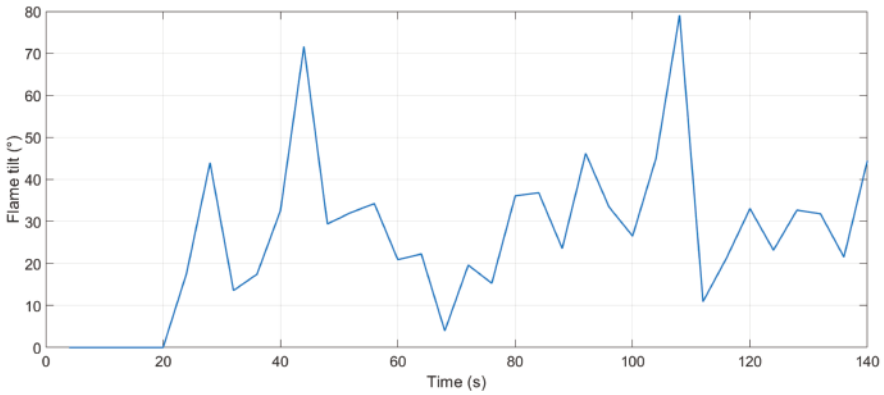



Figure 6.12: Temporal evolution of the flame tilt angle.

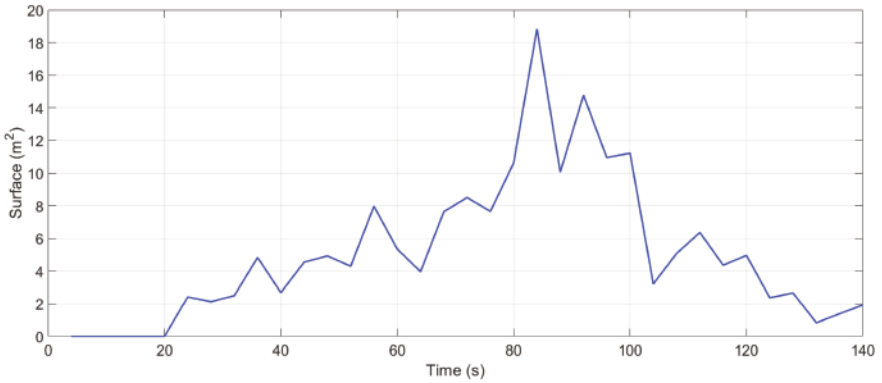



Figure 6.13: Temporal evolution of the combustion surface.

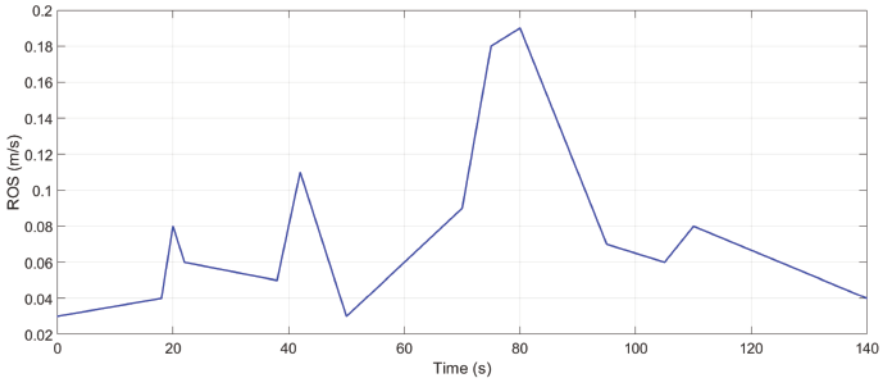



Figure 6.14: Temporal evolution of the rate of spread.

Figure 6.11 shows the temporal evolution of the width, height, and depth. The height of the fire during the experiment remains low, this is due to the fact that during the climb the flame remains close to the ground. The width is around 5 m, which is in line with the area width. The depth for the first 60 seconds grows slowly, with an average value of 1.3 m. At

about 80 seconds the depth grows faster, a sign that the line of the advanced points and the line of the backward points of the fire ground progress with different speed.

Figure 6.12 presents the flame inclination angle. The average inclination is about 25°. At about 45 seconds and 110 seconds there was a gust of wind, and the flame was inclined at 70° and 80°, respectively. At about 50 seconds the wind went down and the inclination of the flame was 5°.

Figure 6.13 shows the surface of the base participating in the combustion. The surface grows steadily starting from 20 seconds, until it reaches its maximum peak at about 80 seconds. 80 seconds is also the instant the line of advanced points and the line of backward points begin to be spaced.

Figure 6.14 shows the rate of spread during the propagation. The average speed is about 0.08 m/s, with a peak at 0.18 at 80 seconds. During the first regime (between 0 and 28 seconds), the average speed is 0.05 m/s. During the second regime (between 36 and 72 seconds) the average speed is 0.07 m/s. During the third regime (between 76 and 116 seconds) the average speed is 0.12 m/s. The reason why at 80 seconds the line of advanced points and the line of backward points move away is in fact due to an increase in the ROS of the fire front. Overall, the estimated data are consistent with what might be expected for this type of experiment.

Figure 6.15 shows the successive georeferenced fire front lines on a Google map.


Figure 6.15: Successive georeferenced fire front lines.

Chapter 7

# Conclusion and perspectives

Contents



## 7.1 Conclusion

The estimation of fire geometric characteristic is important for the researches working in the field of forest fires. The danger, the dimensions and the distances covered by the fires make difficult to measure its characteristics. A way to solve this is to use the computer vision.

As presented in Ch. 1, the vast majority vision devices used to measure wildland fires work with 2D information. In order to obtain the estimation of the geometric characteristics with these systems, it is necessary to have reference marks on the ground, fixed and easily recognizable; marks which cannot be added during a real fire, or that may be not completely visible. In addition, it is impossible to obtain detailed information such as the exact shape of the front, or the orientation of the flame towards a target. A solution, composed of several multimodal stereovision systems, allows to obtain 3D information and to estimate the geometric characteristics of fires without reference marks. However, this solution uses systems fixed on the ground, and it is difficult to consider it with large fires, spreading over large distances, or for fires occurring into inaccessible areas. The use of drones allows to obtain measurements with the same precision throughout the fire spread choose at each instant the best point of view, and in addition to and carry vision system in areas difficult to access on foot. The solution proposed in this thesis consists in the use of a multimodal stereovision module mounted on a drone. The vision system is composed of a pair of cameras in the visible spectrum and a pair of cameras in the long wave infrared spectrum. The cameras are all synchronized through the use of a Raspberry computer and the images are merged with the information of the navigation sensors (IMU, GPS, compass, and barometer). The architecture of this solution is presented in Ch. 2. This chapter describes also the study on the

wavelengths of the radiation emitted by a fire, which justifies the choice of the used spectral bands. Chapter 3 presents the theoretical aspects of stereovision, and the methodology used for the estimation of the intrinsic and extrinsic parameters of the cameras.

In Ch. 4, is described the developed methodology to estimate fire geometrical characteristics during propagation. In the first part, the fire pixel detection procedure is presented. In this method, a pre-selection fire area is obtained from infrared images selecting the fire pixels using the Otsu method. Considering only the position of the pixels pre-selected in the IR band, pixels are detected in the visible image using a performance competition between 11 state-of-the-art fire detection algorithm. From the fire detected pixels, features are selected and after the use of a matching procedure and a triangulation method, 3D fire points are obtained. This chapter presents also the methods employed to transform the 3D points in order to estimate the fire geometrical characteristics. Several frames are used. The first frame is unique for all acquisitions and positioned in front of the fire area; it is employed to obtain the temporal evolution of the fire position and the rate of spread of the fire front. The second frame has an orientation that is parallel to the orientation of the local base plan and can change during the propagation. This frame is used to obtain an estimation of the fire height and the inclination of the terrain. The third frame has an orientation that follows the instantaneous direction of the fire front during propagation. This frame is used to obtain measures such as the width, the depth, the length, the inclination angle, the base perimeter, the combustion surface, the view factor and volume.

Chapter 5 presents two experiences used to study the reliability and the accuracy of the measurements provided by the system. In the first experience are computed the measures of a motionless car, and compared with the real values. In the second experience, it is used as target a pseudo-static fire, without propagation. The images are obtained from different positions of the drone around the fire. Ch. 6 presents the fire geometrical characteristics obtained for a fire propagating on a plane of 50 $m^2$ and having different slopes.

## 7.2 Perspectives

The vision device and the methods proposed in this thesis satisfy the need of systems able to estimate fire geometrical characteristics in order to understand and model the fire behaviour. In the context of this thesis work several experiences were carried out, to reproduce a fire propagation of about 10 m over different slope. Each of these experiences required several hours of preparation, as well as a large number of people involved. For logistical problems it was not possible to use the system for a large-scale fire and in real firefighter operation.

A first improvement of this framework could be the improving of the accuracy in the position of the drone, for example using a Real-time kinematic GPS sensor. This sensor was

not initially considered given its price, but in recent years low-cost RTK-GPS sensors with good performance have been developed. A second improvement could be the possibility of using a second drone, to obtain two complementary views of the fire at the same time. This could be useful to obtain an estimation of the entire volume of a fire. In addition, for wildfires with great depth, by positioning two drones, one in front of the fire front and another in front of the back fire, it could be possible to monitor the two propagation lines the front fire propagation and the backfire propagation simultaneously. Finally, an ultimate improvement could be to make the system able to obtain the results in "real-time", unlike now that the results are produced in "post-processing". This improvement is possible, for example, by equipping the drone with a more performing on-board computer, capable of performing the calculations currently carried out by the computer on the ground, and sending the results directly to a ground station, for example via 4G connection.

Chapter 8

# French Version

**Contents**



## 8.1 Introduction

Les feux de forêt causent chaque année des pertes écologiques, économiques et humaines dans le monde entier. Annuellement, plus de 340 millions d'hectares de végétation sont détruits sur notre planète par ce phénomène [139]. Avec le réchauffement climatique, il y a une extension des zones sensibles aux feux et globalement les incendies de forêt sont de plus en plus meurtriers et destructeurs [22, 140]. Pour combattre ce risque majeur, des mesures de gestion, de prévention et de lutte sont prises. L'efficacité de ces actions dépend de la connaissance des phénomènes qui ont lieu lors du démarrage et de la propagation des feux. Depuis plus de cinquante ans, des recherches sur les incendies de végétation sont menées pour comprendre ces phénomènes, développer des modèles de propagation et de comportement et prédire l'évolution des incendies [29, 32, 141, 25, 27], le transfert de chaleur [142, 143, 28] et créer des outils d'aide à la décision pour la lutte incendie et l'aménagement du territoire [144, 145, 146, 147, 58].

Des données expérimentales obtenues lors de la propagation de feux de végétation en laboratoire et en extérieur sont nécessaires pour mieux comprendre les phénomènes et améliorer et/ou valider les modèles. Les caractéristiques d'un front de feu comme sa position, sa vitesse de propagation, sa hauteur, sa longueur, sa largeur, sa profondeur, son angle d'inclinaison, sa forme, sa surface et son volume sont importantes. Les informations tridimensionnelles donnent

plus de possibilité d'analyse que les informations 2D. Ainsi, par exemple, tous les points d'un front de feu non linéaire sont intéressants à suivre car leur vitesse traduit la déformation du front de feu qui peut être due à la topologie du terrain, le vent ou la végétation. Egalement, l'énergie transférée d'un feu vers une cible dépend de la distance entre la cible et le feu et de l'orientation de la surface du feu par rapport à la position de la cible [148]. L'obtention d'une forme tridimensionnelle de feu, de la distance feu-cible et de la surface de flamme orientée en direction de la cible permet de calculer cette énergie.

Au cours de ces dernières décennies, il y a eu un intérêt croissant dans la communauté de recherche sur les incendies pour développer des méthodes basées sur la vision par ordinateur afin de mesurer les caractéristiques géométriques d'un feu lors de sa propagation. L'étape de détection des pixels de feu dans une image est essentielle dans le processus de mesure du feu par vision car elle détermine la précision avec laquelle les caractéristiques du phénomène peuvent être estimées. Le domaine du visible est le spectre de référence pour les chercheurs travaillant sur les feux de forêt en raison de l'utilisation plus simple des caméras de ce spectre et de leur prix très abordable. Cependant, la détection des pixels de feu dans les images couleur est difficile et, dans [119], il apparaît que la performance des onze algorithmes de segmentation couleur de réference dans la littérature [61, 127, 126, 120, 121, 123, 125, 128, 122, 53, 124] (évalués sur une base de données d'images contenant plus de 500 images de feux de végétation) sont dépendants de caractéristiques de feu telles que la couleur, sa texture, la présence de fumée et la luminosité de l'environnement.

L'utilisation des images du spectre infrarouge permet de dépasser le problème des zones de feu masquées par la fumée. De nombreux travaux ont été développés en relation avec l'utilisation d'images de l'infrarouge pour la détection et la mesure des feux de végétation. Gouverneur *et al.* dans [87] présente une étude dans laquelle il apparaît que les meilleures bandes pour la détection des feux (dans le cas d'une longue distance entre le feu et la caméra) sont l'infrarouge à courtes longueurs d'onde (en anglais, Short Wave Infra-Red : SWIR 1,4 $\mu$m - 3 $\mu$m), l'infrarouge à longueurs d'onde moyenne (en anglais, Middle Wavelength Infrared : MWIR 3 $\mu$m - 5 $\mu$m) et l'infrarouge à longues longueurs d'onde (en anglais,le Long Wavelength Infrared : LWIR 8 $\mu$m - 15 $\mu$m). Toutefois, il est important de noter que compte tenu du fait que le feu émet dans différentes bandes spectrales de manière non uniforme, les zones de feu obtenues dans les images des différents spectres ne se superposent pas parfaitement. Ceci est dû notamment à des zones de gaz chauds qui ne sont pas visibles dans les images du visible [87, 82]. De ce fait, les images de l'infrarouge ne peuvent pas être utilisées seules pour la mesure des caractéristiques géométriques de feu.

Des systèmes utilisant plusieurs modalités ont été proposés : dans [131], Verstockt *et al.* présente l'utilisation d'un système qui combine des capteurs travaillant dans le visible et le LWIR et dans [149] est décrit un système utilisant la technologie TOF (Time of flight) et des capteurs travaillant dans le visible. Il est à noter que dans ces deux études, les expériences ont été limitées à des environnements contrôlés. De plus, du fait de la faible résolution des caméras travaillant dans ces modalités, les systèmes ne sont pas adaptés à de la métrologie pour les feux.

Deux familles de dispositifs ont été développées pour la mesure par vision des feux. La première utilise des caméras positionnées de sorte à obtenir des vues complémentaires du feu. Ces travaux ont été menés en intérieur [50, 44, 41, 43] et ne sont pas applicables dans un environnement non contrôlé. Martínez-de-Dios *et al.* [49] a développé un système travaillant à l'échelle du terrain qui utilise des caméras fixes positionnées à deux points de vue afin d'obtenir des visions complémentaires du feu (vues frontale et latérale). Les images frontales permettent l'estimation de la position du front de feu et de la vitesse du point le plus avancé. Les caméras latérales sont utilisées pour extraire la hauteur maximale de flamme, son inclinaison et la largeur de la base du feu. Une caméra infrarouge est également déployée en position latérale afin d'obtenir la position du feu et la largeur du feu dans le cas où de la fumée masque les zones de feu dans les images du visible. Les valeurs exprimées dans un repère monde des caractéristiques de feux sont calculées en calibrant les caméras à partir d'un modèle de terrain, de la position des caméras et de repères naturels ou artificiels posés dans la scène. Le système développé est utilisé pour estimer la position du point le plus avancé du front de feu et sa largeur. Un modèle 3D du feu est généré en utilisant les mesures calculées. Dans [51] les auteurs proposent une amélioration de cette méthode en utilisant des caméras du visible et de l'infrarouge fixées au sol ou montées sur un drone muni d'un GPS et d'une carte inertielle ainsi que des techniques de fusion de données pour intégrer les résultats issus de toutes les caméras. Le dispositif peut estimer la position du front de feu, sa vitesse et la hauteur maximale de flamme en temps réel. Une forme 3D de feu est obtenue par concaténation de triangles, chacun étant caractérisé par les paramètres estimés. Ce système estime une partie des caractéristiques géométriques requises par les chercheurs travaillant sur la propagation des incendies de forêt et la modélisation du comportement. Cependant, en utilisant des informations 2D, les informations de profondeur 3D ne peuvent pas être obtenues pour tous les points d'un front de feu et par conséquent certaines caractéristiques ne peuvent pas être estimées. De plus, son déploiement peut être difficile dans les zones d'accès difficile.

La deuxième famille de systèmes est basée sur l'utilisation de la stéréovision pour calculer l'information 3D [95]. Cette méthode permet d'obtenir des informations de profondeur à partir de deux images simultanées et décalées de la même scène sans avoir besoin de repères ou de balises artificielles. Elle a été appliquée à l'échelle du laboratoire pour mesurer les caractéristiques géométriques d'un front de feu [40, 75]. Un système a également été développé afin de travailler à l'échelle du semi-terrain [54]. Le dispositif prend en charge l'utilisation de plusieurs systèmes de stéréovision positionnés de manière à capturer des vues complémentaires du front de feu pendant sa propagation. Des paires de caméras multimodales fonctionnant à la fois dans les bandes spectrales du visible et du proche infrarouge sont employées. Chacune des caméras formant la paire de stéréovision est fixée rigidement sur une barre métallique avec une distance inter-caméra d'un mètre. Un recalage des points 3D estimés depuis les différentes vues est effectué afin d'obtenir une reconstruction 3D complète du front de feu. Ceci est effectué en utilisant les données visuelles, les positions GPS et les données inertielles. Ce dispositif permet d'obtenir l'évolution temporelle des caractéristiques géométriques d'un incendie se propageant sur une dizaine de mètres : position, vitesse, largeur de la base du feu, profondeur, périmètre, hauteur, longueur, angle d'inclinaison, forme 3D, surface, facteur de vue et volume. Cependant, il a des limites : d'abord, il faut anticiper la trajectoire du

feu pour positionner de manière optimale les caméras. Deuxièmement, comme la précision théorique de la mesure de profondeur augmente proportionnellement au carré de la distance caméra-objet [68], il y a une perte de précision lorsque le feu s'éloigne des caméras.

Cette thèse présente l'estimation de caractéristiques géométriques d'un feu se propagaeant sans limitation de distance à partir d'un système de stéréovision multimodale porté par drone. Un dispositif de vision, fixable sur drone, composé de deux caméras fonctionnant simultanément dans le visible et l'infrarouge et capable de produire des images stéréoscopiques multimodales géoréférencées a été développé. Le dispositif proposé permet d'estimer une forme 3D du front de feu, sa position au sol, sa vitesse, sa surface, son épaisseur, sa hauteur et son angle d'inclinaison pour toute configuration de terrain et sans utiliser de modèle numérique de terrain. Pour évaluer la performance du système, des essais ont été réalisés avec des incendies expérimentaux se propageant sur plusieurs dizaines de mètres sur des terrains à pente variable.

## 8.2 Matériels

La stéréovision a été choisie dans cette thèse pour obtenir des points 3D de feu à partir desquels les caractéristiques de l'incendie sont estimées.

Dans un système de stéréovision, il est possible de calculer la précision théorique de la mesure de profondeur $\Delta z$, étant donnée la longueur focale $f$ des caméras, la distance moyenne entre la caméra et la cible $z$, la longueur inter-caméras $B$, et l'imprécision de disparité $\Delta d$. La relation s'exprime par :

$$\Delta z = \left| \frac{z^2}{Bf} \right| \Delta d \tag{8.1}$$

où $z$, $\Delta z$ et $B$ sont exprimés en mètres et $f$ et $\Delta d$ en pixels.

A partir de cette équation, il peut être vu que l'erreur dans $\Delta z$ est proportionnelle au carré de la distance caméra-cible ; il est donc important que $z$ soit aussi petit que possible. Aussi, il a été décidé de transporter le système de vision par drone afin de suivre le feu tout au long de sa propagation en maintenant un bon point de vue et une distance minimale (mais sécuritaire) caméra-feu.

L'équation 8.1 montre également que pour $z$, $f$ et $\Delta d$ constants, plus la distance inter-caméras est grande et plus petite est l'erreur sur la distance estimée $z$. Les caméras du système de stéréovision considérées dans ce travail étant fixées sur un axe, la distance inter-caméra est la longueur de cet axe qui est porté par le drone. Il a été établi par un pilote de drone professionnel que pour un drone DJI S1000 la longueur maximale de cet axe est de 1 m. Avec cette valeur pour $B$, une imprécision de disparité égale à 1 pixel, et en considérant les caméras du visible utilisées, l'erreur théorique dans $z$ est inférieure à 20 cm pour une distance

feu-caméra égale à 15 m.

Le spectre de référence pour les chercheurs travaillant sur la problématique des incendies de forêt est le visible. Les zones de feu apparaissent texturées ce qui permet d'avoir des caractéristiques et des points 3D par stéréovision sur toute la zone de feu. Cependant, il peut être difficile de détecter les pixels de feu dans les images du visible en raison de la couleur et de la texture de la flamme ainsi que de la présence possible de fumée. Les images de feux obtenues dans le spectre de l'infrarouge montrent des zones de feu avec un fort contraste d'intensité avec l'environnement. Il a été décidé d'utiliser des informations multimodales à partir d'images de l'infrarouge et du visible pour segmenter les zones de feu dans les images du visible. Le spectre LWIR a été choisi car, d'une part, il est adapté pour obtenir des informations (base et flamme) à partir de feux lointains et en présence de fumée, et d'autre part, des caméras légères et de petite dimension, portables par drone, et de prix abordables sont disponibles à l'achat.

### 8.2.1 Solution proposée

Le dispositif développé est composé de deux modules de stéréovision différents, l'un fonctionnant avec des caméras du visible et l'autre avec des caméras de l'infrarouge.

#### Module de stéréovison composé de caméras du visible

Ce module contient deux caméras MIPI OV4689 fabriquées par Leopard Imaging [88] contrôlées par le processeur OV580 produit par OmniVision (Fig. 8.1).

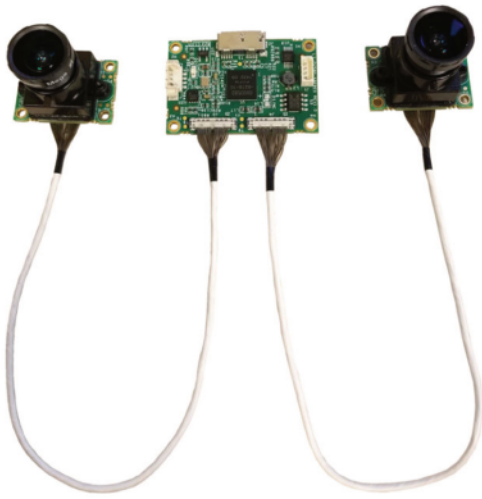


Figure 8.1: Module de stéréovision composé de deux caméras MIPI OV4689 et d'un processeur OV580.

Dans ce système, le processeur est utilisé pour fusionner les deux flux vidéos provenant de chaque caméra à 30 Hz et pour créer une seule image contenant les deux images l'une à côté de l'autre. Les caractéristiques de la caméra utilisée sont une focale de 2,8 mm, une taille de pixel de 2 $\mu$m, un champ de vision horizontal de 98°, un capteur CMOS de 1/3", et un filtre IR avec une fréquence de coupure à 650 nm. La résolution de l'image est réglée sur 2208×1242 pixels. Les caméras et le processeur sont alimentés par un câble USB3 avec une tension de 5 V et le courant requis est de 296 mA. Le poids des deux objectifs et le processeur est de 30 gr.

Les caméras sont montées sur un axe avec une distance inter-caméra de 0,9 m. Le support est un bois contreplaqué composé d'une structure fibreuse et composite qui limite ses déformations. Des trous ont été faits afin de réduire son poids. Une barre d'aluminium a été ajoutée sur le côté supérieur afin de rigidifier le système (Fig. 8.2).

Figure 8.2: Module de stéréovision composé de caméras du visible monté sur le drone DJI S1000.

La figure 8.3 présente une image du dispositif de vision porté par un drone DJI S1000 situé approximativement à 12 m à l'arrière d'un feu. La figure 8.4 montre les images stéréoscopiques couleur obtenues dans une position du drone très proche de celle présentée en figure 8.3.

Figure 8.3: Système de vision porté par le drone DJI S1000 situé approximativement à 12 m à l'arrière d'un feu.

(a) (b)

Figure 8.4: Images stéréoscopiques prises par le système de vision porté par drone. (*a*) Image de gauche. (*b*) Image de droite.

### Module de stéréovison composé de caméras de l'infrarouge

Ce module est composé de deux caméras Vue Duo Pro R produites par FLIR [89] et travaillant dans le spectre LWIR. Les caractéristiques de cette caméra sont une focale de 9 nm, un champ de vision horizontal de 69°, une résolution de 640×512 pixels, et une fréquence de fonctionnement de 30 Hz. Le poids de chaque caméra est de 113 gr. L'acquisition d'image est déclenchée par un signal PWM externe de 50 Hz avec un cycle de fonctionnement de 20 ms.

Les capteurs IR sont positionnés au-dessus des capteurs visibles tels qu'ils apparaissent sur la figure 8.5, de manière à obtenir deux paires de caméras IR-visible distantes de 0,9 m ; la distance entre le centre des lentilles est de 5 cm.

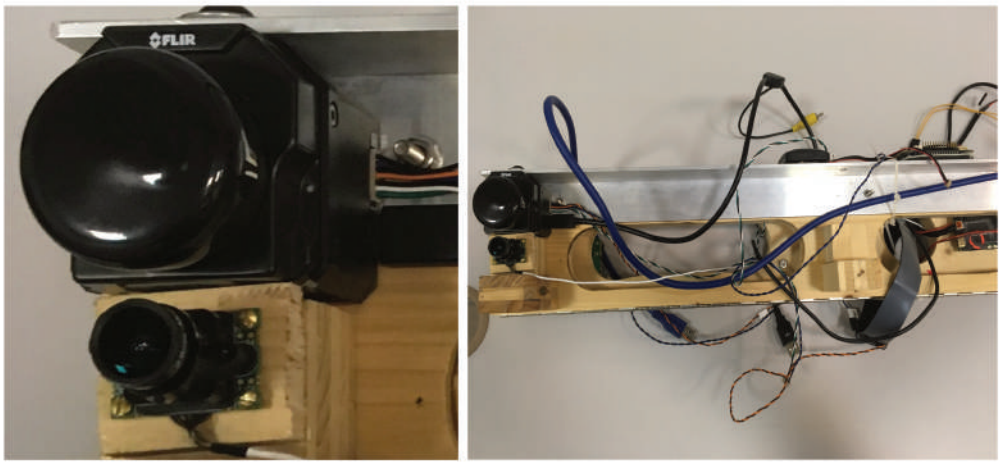



Figure 8.5: Caméras du visible et de l'IR placées l'une en dessous de l'autre.

La figure 8.6 montre le système de stéréovision multimodale, composé des caméras du visible et de celles de l'infrarouge vissées sur l'axe en contreplaqué, porté par le drone.

Figure 8.6: Système de stéréovision multimodale porté par le drone DJI S1000.

Le système de stéréovision est positionné sur le train d'atterrissage du drone dans une position fixe. La figure 8.7 montre le système d'attache.

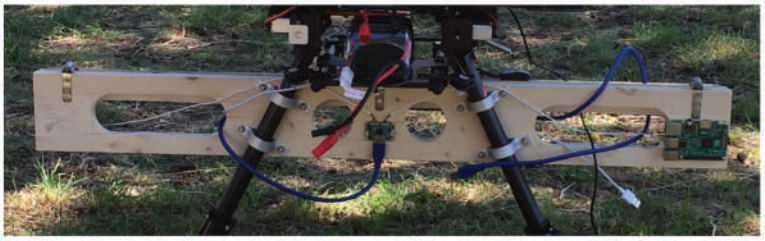


Figure 8.7: Système permettant d'attacher le dispositif de vision sur le drone.

Les caméras IR sont alimentées à partir d'une batterie délivrant une tension de 5 V nécessaire à chaque caméra qui consomme au maximum 7,8 W. Cette batterie est positionnée sur le châssis du drone.

La synchronisation des caméras IR est réalisée à partir d'un signal PWM généré par un ordinateur Raspberry. Un câble connecté au port GPIO de l'ordinateur est divisé en deux parties égales afin d'apporter le même signal aux deux caméras. Une attention particulière a été portée afin que les longueurs de câble soient très proches (différence inférieure à 1 mm) et que les soudures sur les câbles soient aussi homogènes que possible et très légères.

La figure 8.8 montre deux images stéréoscopiques de l'infrarouge acquises avec ce dispositif au même instant que les images du visible montrées en figure 8.9.

(a) (b)

Figure 8.8: Images stéréoscopiques de l'infrarouge. ($a$) Image de gauche. ($b$) Image de droite.

(a) (b)

Figure 8.9: Images du visible. ($a$) Image de gauche. ($b$) Image de droite.

**Capteurs supplémentaires et cablâge**

D'autres capteurs ont été ajoutés au dispositif de vision afin d'obtenir sa position et son orientation. En particulier, une carte inertielle, un baromètre, un capteur GPS et un compas. Les données de ces capteurs sont lues par une carte ATmega8 montée sur Arduino et envoyées au Raspberry via un câble USB.

La position du dispositif de vision est calculée par le capteur GPS et le capteur barométrique. Le capteur GPS est un "u-blox M8 GNSS" [90] avec une précision de position de 2 m. Le capteur GPS est fixé entre les deux caméras visibles en position centrale, et il donne deux valeurs angulaires appelées "latitude" et "longitude". La latitude est la distance angulaire, mesurée en degrés le long de l'arc méridien, entre l'équateur et le parallèle passant par le point considéré. La longitude est la distance angulaire, mesurée en degrés le long de l'arc parallèle, entre le "méridien de Greenwich" (méridien premier) et le méridien passant par le point considéré (figure 8.10).

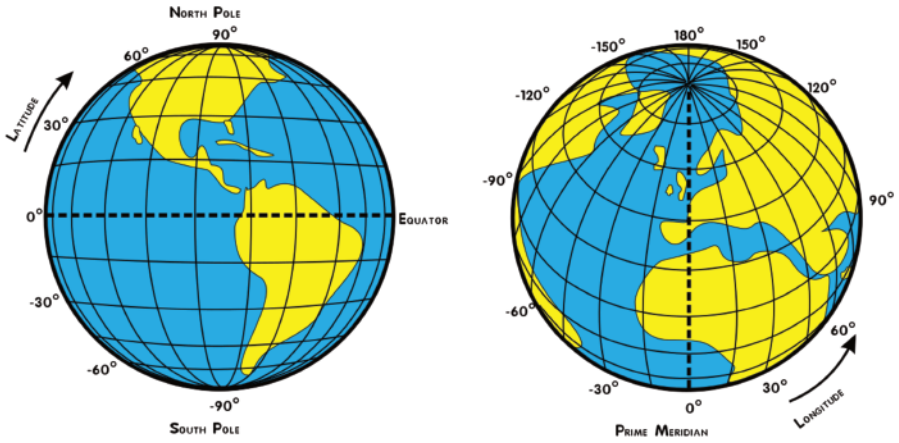



Figure 8.10: Latitude et longitude sur la Terre.

Étant donné que le capteur barométrique a une précision de 10 cm, et l'altitude donnée par les capteurs GPS a une précision de 2,5 m, pour obtenir la distance verticale de l'appareil du niveau de la mer (altitude), le capteur barométrique est préféré aux capteurs GPS.

L'orientation du dispositif de vision est donnée par la carte inertielle et par le capteur compas. La carte inertielle est utilisée pour obtenir les angles de roulis et de tangage du système de stéréovision; le capteur compas est utilisé pour obtenir l'angle de cap du dispositif (Fig. 8.11).

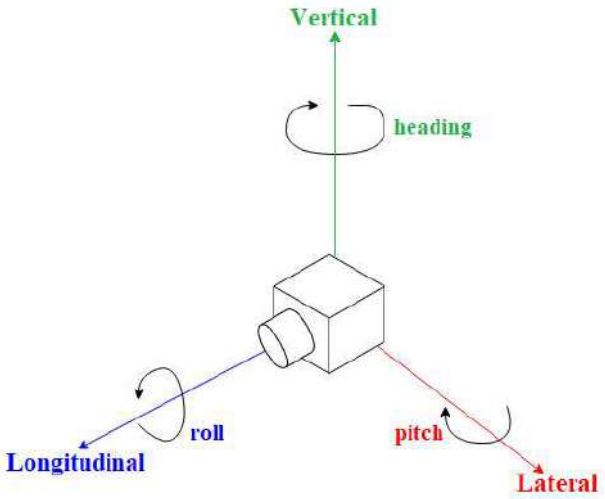



Figure 8.11: Angles de roulis, de tangage et de lacet de la caméra.

Les angles de roulis, de tangage et de lacet sont tous égaux à zéro lorsque le dispositif de stéréovision est horizontal et pointe vers le Nord. L'angle de roulis est positif si le dispositif tourne vers la droite, et il est négatif si le dispositif tourne vers la gauche. L'angle de tangage est négatif si l'appareil pointe vers le bas (l'état de l'appareil qui pointe vers le haut n'est pas pris en compte parce que les caméras sont toujours plus élevées que le feu). L'angle de cap est positif si le dispositif pointe vers l'est, négatif si le dispositif pointe vers l'ouest.

La figure 8.12 présente les connexions physiques de tous les éléments du système de stéréovision multimodale.

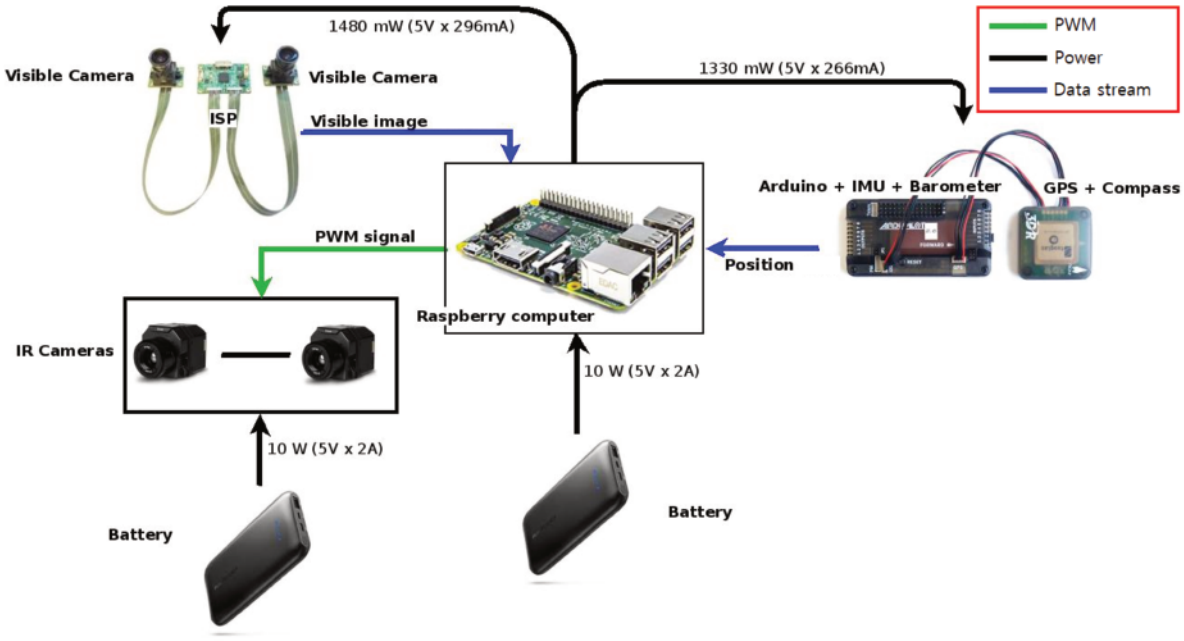



Figure 8.12: Diagramme montrant les éléments et les connexions du système de stéréovision multimodale.

Comme on peut le voir, l'ordinateur Raspberry est positionné au centre du système logique et toutes les connexions sont dirigées vers lui; il agit en tant que coordinateur des différents composants. Il est alimenté par une batterie USB délivrant une tension de 5V et une intensité 2 A. La puissance entrant dans l'ordinateur Raspberry est de 10 W. Cette valeur est suffisante pour le fonctionnement sur la carte Arduino et le dispositif stéréoscopique de vision. La puissance requise par l'Arduino est de 1 W, celle requise par les capteurs GPS/Compass est de 330 mW et celle requise par le module de stéréovision visible est de 1480 mW. Pour fonctionner correctement l'ordinateur Raspberry a besoin de 1,2 W. La consommation totale de l'ensemble composé par la carte Arduino, le GPS/Compass, le module de caméra visible et l'ordinateur Raspberry est de plus de 4 W. La puissance restante est inférieure à 6 W et elle n'est pas suffisante pour le fonctionnement du module de caméras IR. Comme mentionné ci-dessus, les caméras infrarouges ont besoin d'une charge d'énergie de 7,8 W et pour cette raison une alimentation séparée de 12,5 W est utilisée.

Comme cela peut être vu en figure 8.12, le processeur qui relie les deux caméras du visible est également connecté à l'ordinateur Raspberry. Ce composant reçoit les deux images provenant des caméras du visible et les joint en une seule image. Cette image est envoyée via un câble USB à l'ordinateur Raspberry dans un flux continu. Compte tenu de la vitesse de propagation du feu, l'ordinateur Raspberry est configuré pour lire le flux d'images toutes les 4 secondes.

La carte Arduino lit les données en continu provenant de l'accéléromètre, du GPS/Compass et du baromètre, avec une fréquence de l'ordre du milliseconde. Ces trois capteurs utilisent

trois canaux de communication distincts de la carte Arduino afin de ne pas partager le même bus et prendre trois lectures tout en gardant le temps de latence entre chacun d'eux (estimé dans l'ordre de la microseconde). Les données de l'accéléromètre, du GPS/Compass et du baromètre sont envoyées à l'ordinateur Raspberry via un deuxième câble USB en utilisant un protocole Mavlink [91].

L'ordinateur Raspberry, la carte Arduino et la batterie ont été ajoutés en tenant compte de la balance de l'axe. La charge utile totale pour le système de vison complet est de 3,2 kg. La figure 8.13 montre le dispositif de stéréovision multimodale complet avant un vol.

Figure 8.13: Dispositif de stéréovision multimodale complet avant un vol.

### 8.2.2 Méthode

Le diagramme de la méthodologie proposée est présenté dans la figure 8.14. La première étape du processus consiste à détecter les pixels de feu. Elle est menée dans le spectre du visible et du LWIR. Cette étape est suivie par la détection des points d'intérêt dans l'image du visible afin de sélectionner les points les plus importants à utiliser dans l'algorithme de correspondance stéréoscopique. A partir des points correspondants, une procédure de triangulation permet leur reconstruction tridimensionnelle. à chaque instant d'acquisition d'image, l'équation du plan local du terrain sur lequel se situe le feu est estimée à partir des points de feu 3D les plus bas. Les points de feu 3D sont alors projetés dans plusieurs repères afin de produire des résultats exprimés dans des repères de référence et intrinsèques au front de feu. La dernière étape est la reconstruction du modèle de feu 3D et l'estimation des caractéristiques géométriques du feu à partir des points 3D transformés.

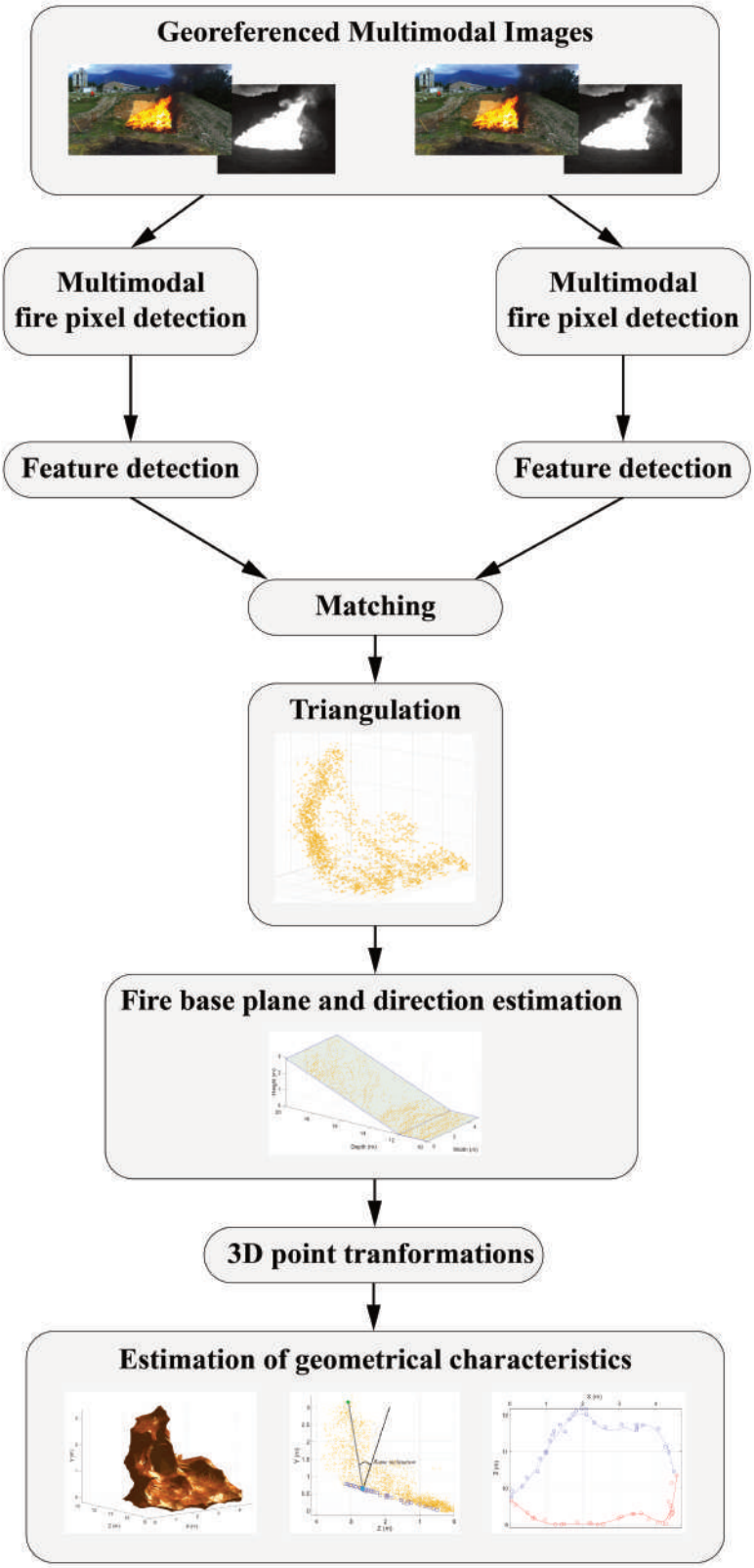



Figure 8.14: Diagramme de la méthode proposée.

#### 8.2.2.1 Détection multimodale de pixels de feu

Une méthode de détection mutlimodale de pixels de feu a été développée en utilisant des informations obtenues dans les images du spectre LWIR afin de présélectionner les pixels des images du visible à traiter. À chaque instant d'acquisition d'image, une image du spectre visible et une image du spectre LWIR sont obtenues par chaque caméra Duo Pro R. La figure 8.15 montre un exemple d'images de feu obtenues simultanément.

(a) (b)

Figure 8.15: Exemple d'images de feu obtenues simultanément dans le spectre du visible et du spectre LWIR. (**a**) Image acquise dans le visible, (**b**) Image acquise dans le spectre LWIR.

Les programmes de la toolbox Matlab développée par [129] auxquels ont été apportées de petites modifications ont été utilisés pour produire des images superposées. La procédure multimodale se déroule en deux étapes. Dans un premier temps, l'image du LWIR est traitée avec la méthode de sélection de seuil développée par Otsu [130] afin de trouver l'emplacement des pixels d'intensité la plus élevée correspondant aux pixels de feu. Seuls les pixels situés dans l'image du visible à la même position que ceux détectés dans l'image du LWIR seront considérés pour détecter les pixels de feu.

La détection des pixels de feu dans l'image du visible en ne considérant que les pixels de la zone présélectionnée par les informations de l'infrarouge se fait de la manière suivante : une interface graphique permet d'évaluer l'efficacité des onze méthodes de détection de pixels de feu de référence [61, 127, 126, 120, 121, 123, 125, 128, 122, 53, 124] à partir des premières images de la séquence d'images associée à la propagation de feu considérée. La méthode décrite dans [53] a été adaptée afin d'utiliser uniquement la deuxième étape de cette procédure (la première étape correspondant à une procédure de présélection de pixels). La procédure permettant d'obtenir la meilleure détection de pixels est utilisée pour toutes les images de la séquence.

La figure 8.16 montre les pixels de feu détectés dans l'image du visible présentée dans la figure 8.15.

Figure 8.16: Pixels de feu détéctés dans l'image du visible présentée dans la figure 8.15

.

#### 8.2.2.2 Détection de points d'intérêt, appariement et triangulation

Les points 3D de feu sont obtenus par triangulation à partir de points d'intérêt de feu appariés entre les images stéréoscopiques. Les points d'intérêt sont détéctés dans les zones de feu des images en utilisant l'algorithme de détection de Harris [105] et la procédure SURF [107]. La première méthode citée est appliquée en considérant des zones de $5 \times 5$ pixels. La seconde utilise 6 échelles de niveau et les tailles de filtre suivantes : 27×27, 51×51, 75×75, 99×99, 123×123 et 147×147. Les points obtenus par ces deux méthodes sont cumulés et les points doubles sont éliminés.

La figure 8.17 montre 3 238 points détectés par cette méthode dans l'image présentée en figure 8.16.

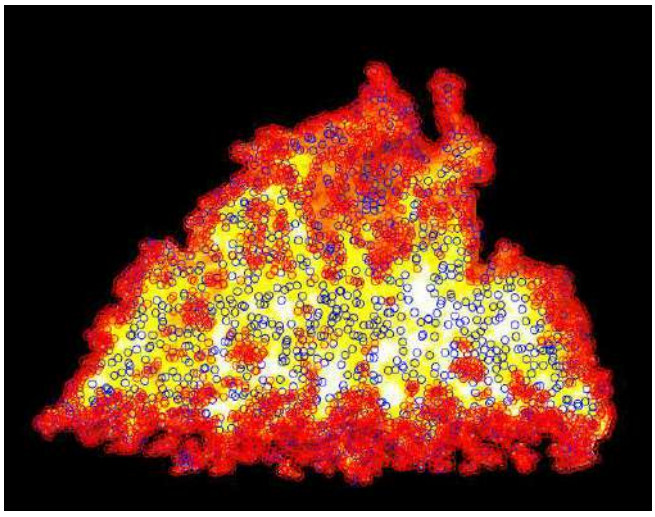


Figure 8.17: Points d'intérêt détectés dans l'image présentée en figure 8.16 (par l'algorithme de Harris (cercles rouges) et par l'algorithme SURF (cercles bleus)).

La procédure d'appariement est celle dite de "Somme des Distances au Carré Normalisées à moyenne Nulle" (Zero mean Normalised Sum of Squared Differences, ZNSSD) [74].

La figure 8.18 montre les points appariés obtenus à partir des points d'intérêt détectés dans les images stéréoscopiques dont une des images est présentée en figure 8.15.

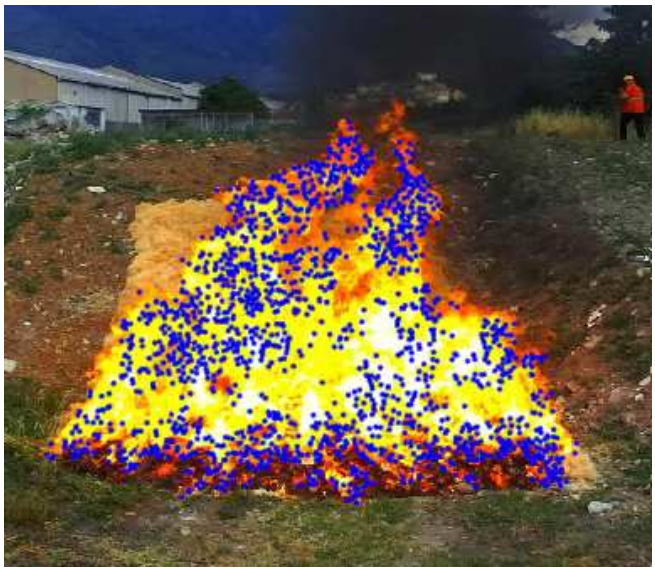


Figure 8.18: Points appariés obtenus à partir des points d'intérêt présentés en figure 8.15

.

À partir des coordonnées des pixels correspondant et en utilisant les paramètres intrinsèques et extrinsèques du système de stéréovision, les points de feu 3D sont calculés par triangulation comme décrit dans [95]. Leurs coordonnées tridimensionnelles sont données par rapport au centre de la caméra gauche du système de stéréovision. Tous les points situés loin du nuage principal de points 3D sont considérés comme des valeurs aberrantes. Si l'une des deux conditions présentées dans l'équation 8.2 est vraie, un point est identifié comme aberrant et éliminé.

$$\begin{cases} |d_1 - \mu_{d1}| \geq 3\sigma_{d1} \\ |d_2 - \mu_{d2}| \geq 3\sigma_{d2} \end{cases} \tag{8.2}$$

où $\mu_{d1}$ et $\sigma_{d1}$ qui sont respectivement la moyenne et l'écart type de la distance $d_1$ entre ce point et ses quatre voisins; $\mu_{d2}$ et $\sigma_{d2}$ qui sont respectivement la moyenne et l'écart type de la distance $d_2$ entre ce point et l'origine du repère caméra.

La figure 8.19 montre les points 3D estimés pour le feu présenté en figure 8.15.

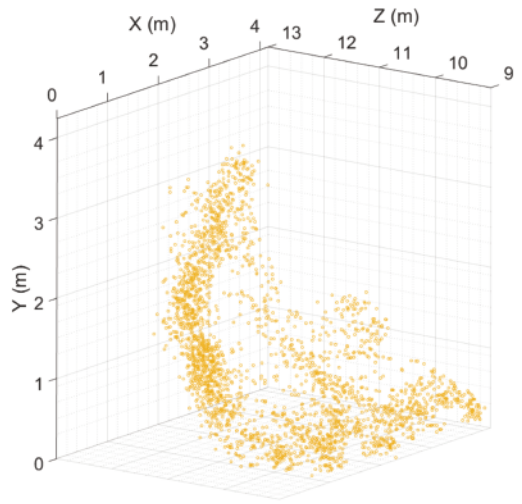



Figure 8.19: Points 3D du feu.

#### 8.2.2.3 Estimation du plan local de propagation et de la direction principale du feu

### Plan local de propagation

Comme il n'y a aucune connaissance *a priori* concernant le plan sur lequel le feu se propage, il est nécessaire de calculer l'équation du plan local de propagation afin d'estimer des informations comme la position du feu sur le sol et sa hauteur. Dans une première étape, pour chaque instant d'acquisition d'image, l'équation du plan de propagation local est obtenue en considérant les points de feu 3D les plus bas et en employant la méthode RANSAC. Dans une deuxième étape, les points 3D appartenant à des plans obtenus à des instants successifs et présentant des variations de pente de moins de 3° sont identifiés et leurs points sont agrégés pour calculer à nouveau l'équation du plan local passant par eux. Cette procédure améliore l'estimation de l'équation du plan local de propagation.

La figure 8.20 montre un exemple d'un plan local de propagation estimé à partir des points les plus bas d'un front de feu.

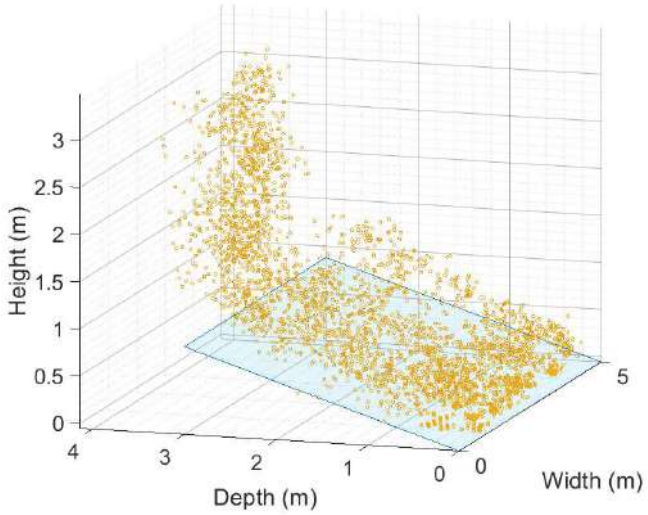
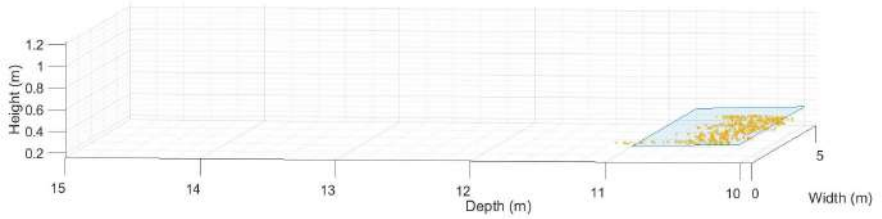


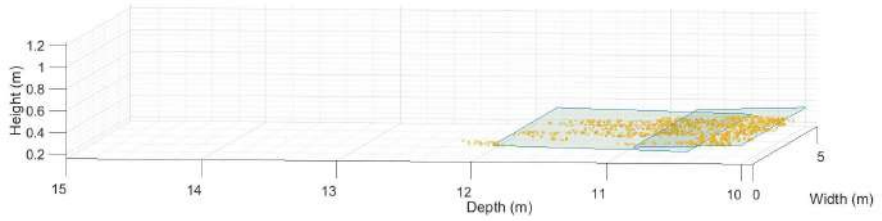

Figure 8.20: Exemple d'un plan local de propagation (couleur bleue) estimé à partir des points les plus bas d'un front de feu.

La figure 8.22 présente la topologie estimée du terrain sur lequel le feu montré en figure 8.15 s'est propagé.

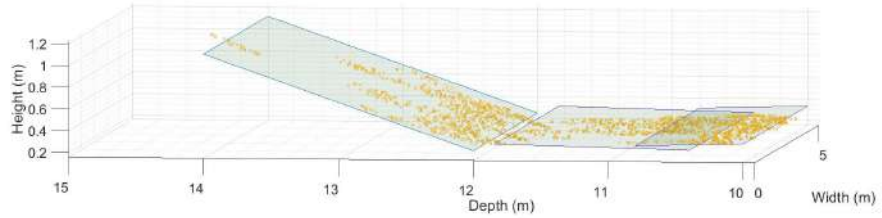
La figure 8.21 montre le résultat de la méthode d'estimation de l'ensemble des plans de propagation.



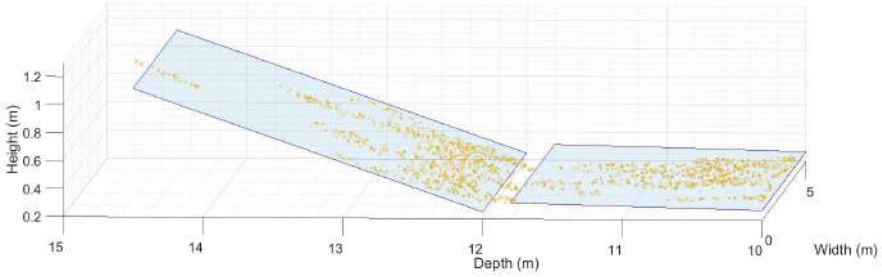

(a)




(b)




(c)




(d)

Figure 8.21: Exemple de l'estimation successive de plans de propagation, en jaune sont dessinés les points 3D utilisés pour calculer les plans de propagation, en bleu sont présents les plans calculés. *(a)* Premier plan de base. *(b)* Deux plans de base successifs ayant une inclinaison proche. *(c)* Trois plans locaux obtenus à différents instants, l'un deux a une variation d'inclinaison supèrieure à 3°. *(d)* Plan moyen obtenu à partir des points des deux premiers plans.

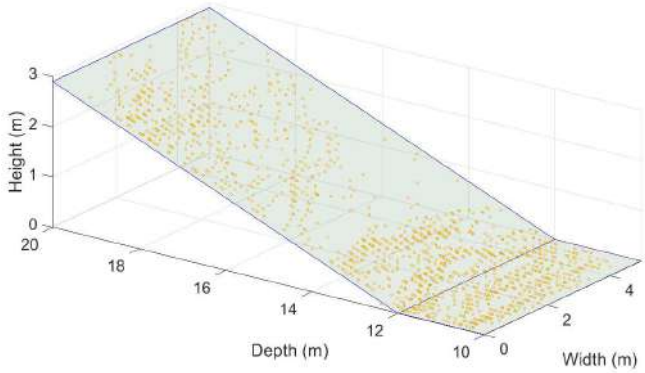



Figure 8.22: Exemple de topologie de terrain estimée.

A chaque instant, le plan local est caractérisé par son angle longitudinal ($\epsilon$) et son angle latéral ($\theta$).

### Direction principale

Le calcul de l'inclinaison, de la largeur et de la longueur d'un feu nécessite des informations sur sa direction principale. Cette direction dépend de paramètres tels que le vent et la pente du sol et peut changer au cours de la propagation. Nous considérons dans ce travail que la direction principale du feu à un instant donné est indiquée par le vecteur dont les extrémités sont le barycentre des deux ensembles successifs de points de feu au sol au moment considéré. Pour calculer cette direction, les ensembles de points de feu au sol de deux acquisitions successives sont considérés. Pour chaque ensemble de points, il est calculé son barycentre. La figure 8.23 montre un exemple de barycentre d'un ensemble de point 3D constituant la base du feu.

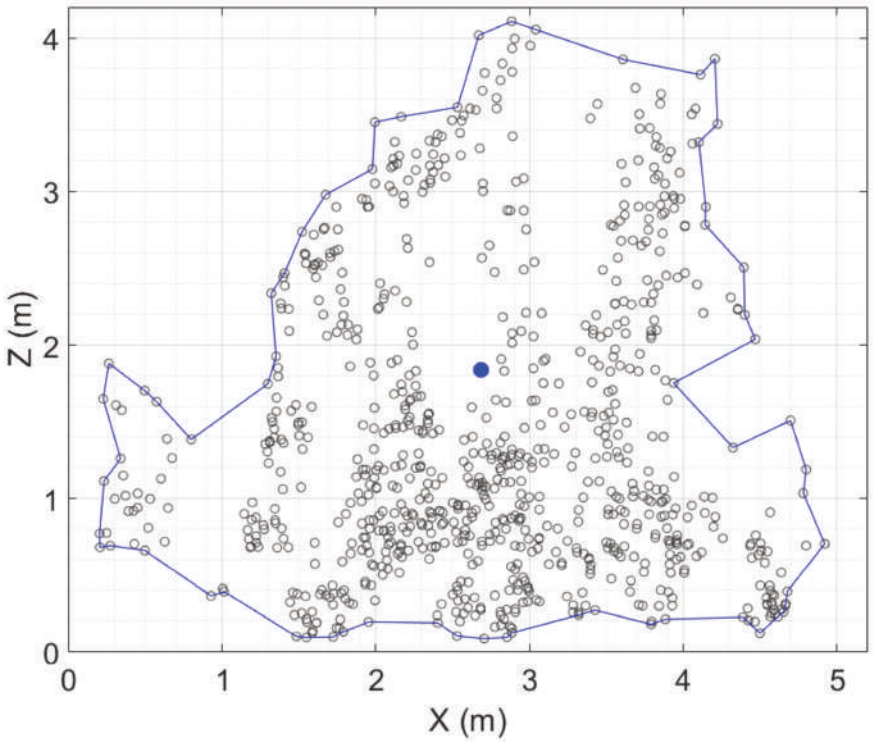



Figure 8.23: Position des points de feu au sol et de leur barycentre.

La figure 8.24 présente le barycentre de deux ensembles de points de feu au sol obtenus à deux instants successifs. $\eta$, l'angle entre la ligne qui relie les deux barycentres et l'axe $z$, représente la direction locale du feu.

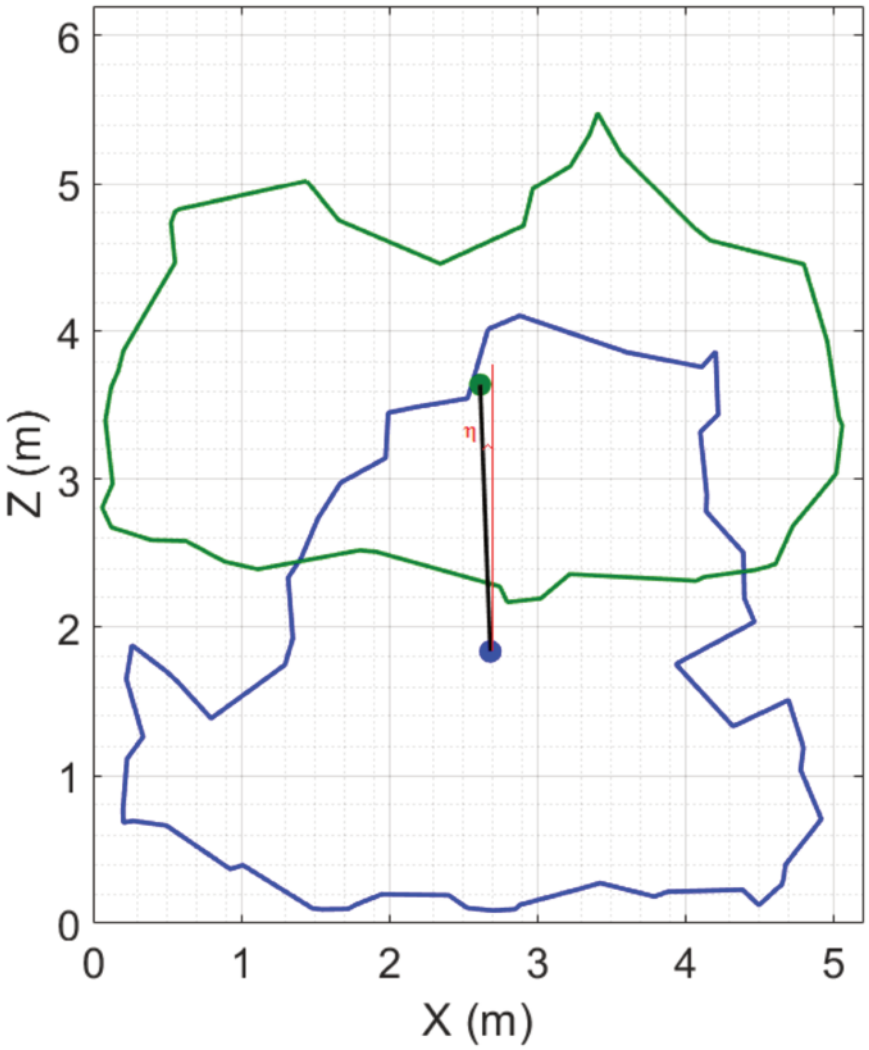



Figure 8.24: Direction principale d'un feu. L'angle $\eta$ représente la direction locale du feu.

#### 8.2.2.4 Transformations des points 3D du feu

A chaque instant d'acquisition d'image, le système de vision porté par drone a une position et une orientation différentes de celles du moment précédent. Dans la mesure où les points 3D obtenus par la stéréovision sont exprimés dans un repère positionné sur le dispositif de vision (repère de la caméra) et que ce repère se déplace dans le temps, il est nécessaire de projeter tous les points 3D dans un repère de référence global pour produire des résultats qui montrent l'évolution temporelle des caractéristiques géométriques du feu comme la position de la ligne de front.

L'origine $O_C$ du repère de la caméra est le centre optique de la caméra visible gauche, l'axe x ($X_C$) correspond à l'axe allant de la caméra gauche vers la caméra droite, l'axe z ($Z_C$) est

perpendiculaire à XC et dirigé vers l'avant des caméras, l'axe y ($Y_C$) est l'axe perpendiculaire aux deux autres de sorte que le trièdre résultant est droitier, comme on peut le voir sur la Fig. 8.25.

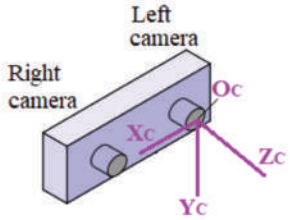



Figure 8.25: Position et orientation du repère de la caméra.

Dans ce travail, il a été choisi de prendre comme origine du référentiel global les données GPS ($O_G$) associées à la première image obtenue par le système de stéréovision avant le décollage du drone. Le drone est positionné approximativement devant la zone de propagation du feu, donc le repère a son axe x ($X_G$) parallèle à la largeur de la zone de brûlage ; son axe z ($Z_G$) correspond à la profondeur de la zone de brûlage, et l'axe y ($Y_G$) est tel que le trièdre est gaucher, correspondant à l'altitude, comme on peut le voir sur la Fig. 8.26.

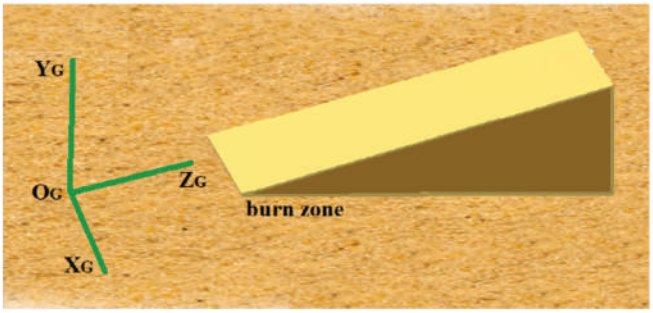



Figure 8.26: Position et orientation du repère global .

Soit $\alpha$, $\beta$ et $\gamma$ lesquels sont, respectivement, les angles de roulis, de tangage et de lacet obtenus par la carte IMU située sur le dispositif de vision avant le décollage de l'UAV et associés au repère de référence globale. Soit $R_\alpha$ (resp. $R_\beta$ et $R_\gamma$) la matrice de rotation homogène autour de l'axe x (resp. autour de l'axe y et autour de l'axe z). Soit $T$ la matrice de translation définie à l'aide des coordonnées N, E et U de la position du système de vision dans le repère de la caméra :

$$T = \begin{bmatrix} 1 & 0 & 0 & E \\ 0 & 1 & 0 & N \\ 0 & 0 & 1 & U \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{8.3}$$

N, E et U sont calculés en appliquant une transformation sur les coordonnées GPS du

repère de la caméra dans le repère de coordonnées globale. Soit Syz une matrice homogène effectuant un échange entre l'axe y et l'axe z. La matrice de transformation est définie par :

$$M_{C\longrightarrow G} = R_\gamma(-\gamma_0)\cdot R_\beta(-\beta_0)\cdot R_\alpha(\alpha_0)\cdot S_{yz}\cdot R_\alpha(-\alpha)\cdot R_\beta(\beta)\cdot R_\gamma(\gamma)\cdot T\cdot S_{yz}\cdot R_\alpha(180). \quad (8.4)$$

Les caractéristiques géométriques du feu telles que la hauteur, la longueur, l'angle d'inclinaison doivent être estimées indépendamment du plan de propagation local sur lequel le feu est situé. Pour ce faire, il est nécessaire d'appliquer une fonction de transformation sur les points de feu 3D afin de travailler comme si la pente du plan de propagation était nulle. Un repère appelé "repère de pente" est considéré ; son origine OS est égale à $O_G$, son axe x ($X_S$ est parallèle à la pente latérale du plan du sol, l'axe z ($Z_S$) est parallèle à la pente longitudinale du plan du sol et l'axe y ($Y_S$) est parallèle à la normale du plan de masse. Ce repère est considéré à chaque instant d'acquisition d'image (et de calcul de points 3D), et chaque plan de propagation local (caractérisé par ses angles moyens $\epsilon$ et $\theta$) impose une orientation du repère S, comme il est possible de le voir sur la figure 8.27.

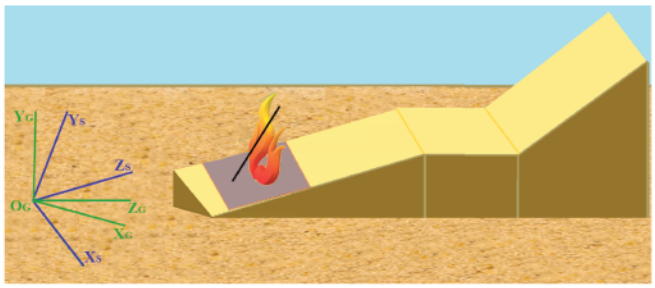



Figure 8.27: Repère de pente (couleur violette) et repère global (couleur verte). La ligne noire représente la normale du plan local.

Les coordonnées des points 3D exprimées dans le repère global sont transformées à l'aide d'une matrice $M_{G\longrightarrow S}$. Soit $\epsilon$ et $\theta$, respectivement, les angles longitudinal et latéral du plan de propagation local obtenus à un instant donné d'image d'acquisition. Soit $R_\epsilon$ et $R_\theta$, respectivement, les matrices de rotation homogènes de l'angle ($\epsilon$) autour de l'axe $X_G$ et de l'angle $\theta$ autour de l'axe $Z_G$. La matrice de transformation $M_{G\longrightarrow S}$ est définie par :

$$M_{G\longrightarrow S} = R_\epsilon \cdot R_\theta. \quad (8.5)$$

Enfin, comme le feu peut changer de direction au fil du temps, pour calculer des caractéristiques géométriques telles que la largeur, l'inclinaison de la flamme et la longueur, il est nécessaire de faire pivoter les points 3D, de telle sorte que l'axe de profondeur du repère utilisé pour les exprimer correspond à la direction instantanée. Soit $\eta$ l'angle entre la direction du feu instantané et l'axe z du repère de pente, les points de feu 3D tournent alors autour de l'axe Y de l'angle $\eta$.

#### 8.2.2.5 Estimation des caractéristiques géométriques du feu

A partir des points de feu 3D transformés, on estime les caractéristiques géométriques du feu. Tous les points sont utilisés pour la reconstruction 3D du feu et le calcul de sa surface et du facteur de forme. Seuls les points au sol sont pris en compte pour l'estimation de la position avant, de la surface de la base, de la largeur et de la profondeur.

### Forme et volume

La méthode de triangulation de Delaunay, avec un rayon égal à 0,35 est appliquée aux points 3D et fournit un ensemble de tétraèdres. Ceux pour lesquels la projection du centre dans l'image segmentée ne correspond pas aux pixels du feu sont éliminés. La figure 8.28 montre la reconstruction 3D du feu illustré dans la figure 8.15.

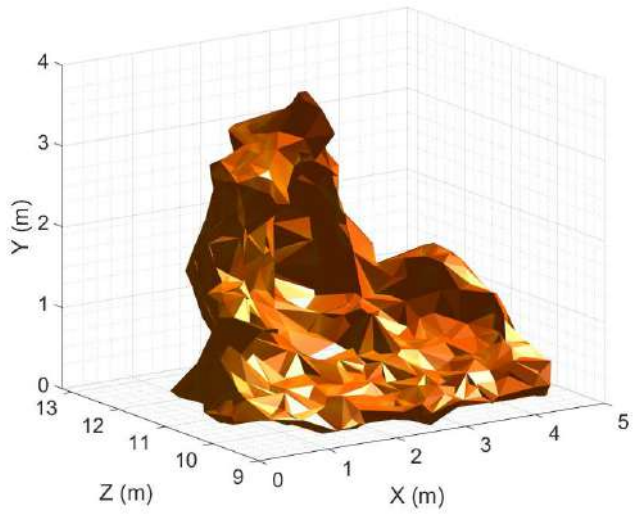



Figure 8.28: Reconstruction 3D du feu avec Delaunay.

Le volume du feu est estimé en additionnant le volume des tétraèdres sélectionnés.

### Surface et facteur de forme

Compte tenu de l'ensemble des triangles formant la surface du feu, il est possible de calculer la partie de la surface qui produit le flux de chaleur dirigé face au front du feu ainsi que la fraction de l'énergie totale émise par la surface du feu et reçue par une cible quelle que soit sa position (également appelée facteur de forme).

Pour calculer la surface du feu, seules les surfaces des triangles qui ne sont pas masquées par d'autres et orientées dans la direction principale du feu sont ajoutées.

Le facteur de forme est estimé en tenant compte de tous les triangles du feu et des coordonnées 3D de la cible. La méthode utilisée est celle décrite dans [69]. Le facteur de forme $F_{S \longrightarrow dT}$, pour l'irradiation entre la totalité de la surface de la flamme $S$ et la zone cible $dT$, est défini comme la fraction de l'énergie totale émise par toutes les surfaces élémentaires du triangle $dS_i$ et reçue par $dT$. Soit $r_i$ la distance entre la zone de cible et le centre du triangle i-ème $dS_i$. Soit $\Theta_{S_i}$ et $\Theta_{T_i}$, respectivement l'angle entre $r_i$ et la normale de la surface du triangle élémentaire et l'angle entre $r_i$ et la normale de la surface cible. Ainsi, il est possible d'établir la formule suivante :

$$dF_{S \longrightarrow dT} = \frac{dT}{S} \sum_{\substack{i=1 \\ \cos\theta_{S_i}>0 \\ \cos\theta_{T_i}>0}}^{N} \frac{\cos\theta_{S_i} \cos\theta_{T_i}}{\pi\, r_i^2}\, dS_i. \tag{8.6}$$

### Position, vitesse de propagation et profondeur

La figure 8.29 montre les points transformés de la base du feu de la figure 8.19.

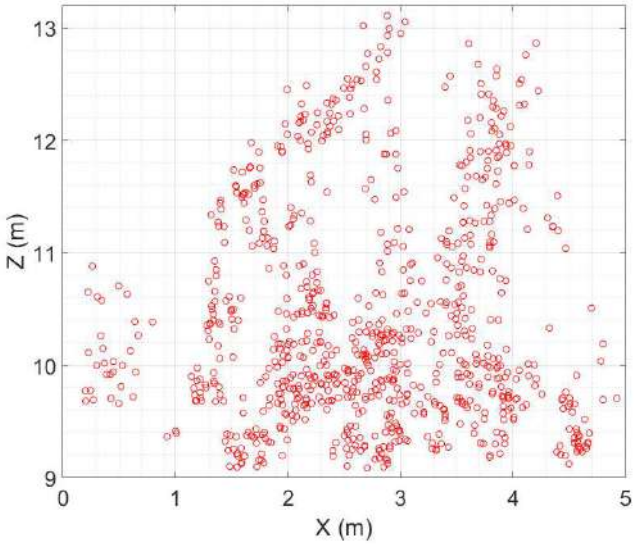



Figure 8.29: Les points 3D de la base du feu.

Un feu est délimité et localisé par sa ligne avant et sa ligne arrière. Pour les retrouver, les points sont traités par secteurs de 15 cm de large. Dans chaque secteur, les points les plus avancés et les moins avancés sont sélectionnés. La ligne avant du feu est obtenue en utilisant une interpolation B-spline avec une fonction polynomiale passant par les points avancés. Par

analogie, la ligne arrière du feu est calculée en utilisant les points arrière. C'est ce que montre la Fig. 8.30.

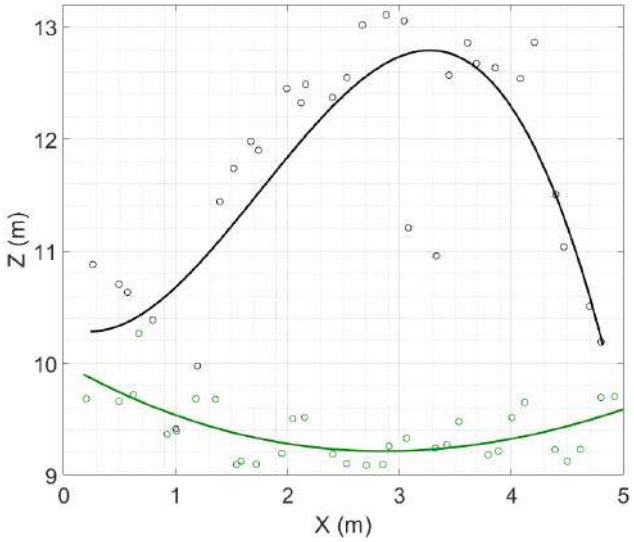



Figure 8.30: Ligne avant (couleur rouge) et ligne arrière (couleur verte) du feu.

Pendant la propagation du feu, les lignes bougent et se déforment. Ce phénomène peut être modélisé en considérant la vitesse de chaque point des lignes. Les chercheurs travaillant sur la modélisation de la propagation du feu s'intéressent à la vitesse de propagation de points caractéristiques situés au centre et sur les côtés de la ligne de feu. Cette vitesse est égale au rapport de la distance entre deux points équivalents sur deux lignes successives divisées par l'intervalle de temps entre les deux moments d'acquisition des images à partir desquelles les courbes ont été calculées, comme il est décrit dans [69]. En considérant deux lignes successives et un point donné de la première ligne, un point équivalent est le point d'intersection entre la normale de la première ligne passant par le point donné et la deuxième ligne.

La profondeur du front de feu est estimée en calculant la distance entre le point moyen obtenu à partir des points les plus avancés et le point moyen obtenu à partir des points les moins avancés.

### Surface de combustion

Les chercheurs travaillant sur la modélisation des incendies de forêt considèrent que le combustible sous la flamme est le seul à contribuer à la combustion et font une approximation de la surface de la base du feu par des formes simples [134]. À partir des points du feu au sol, un polygone est obtenu en utilisant la méthode décrite dans [135] (Fig. 8.31). Il correspond à la base du feu et à la surface du combustible en combustion. Sa superficie est calculée en

additionnant la surface des triangles contenus dans le polygone.

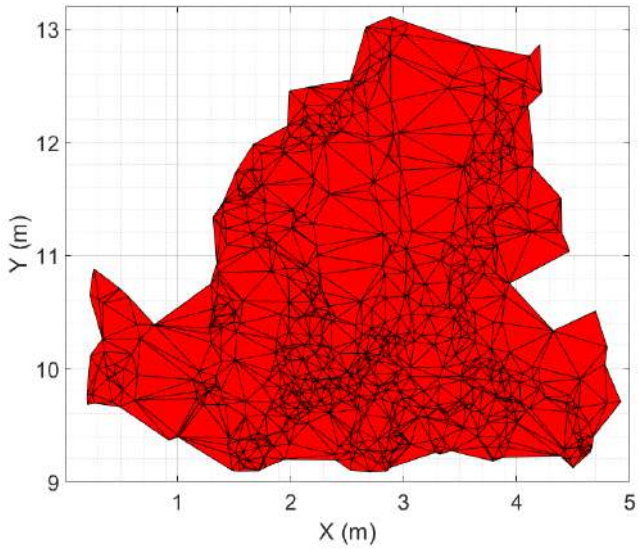



Figure 8.31: Surface du combustible à la base du feu.

### Largeur, hauteur, longueur et angle d'inclinaison

Le calcul de la largeur, de la hauteur, de la longueur et de l'angle d'inclinaison du feu est effectué à partir des points 3D transformés, de telle sorte que l'axe de profondeur du repère utilisé pour les exprimer correspond à la direction instantanée du feu.

L'estimation de la largeur du feu se fait en trois étapes. Dans un premier temps, les deux points qui ont une coordonnée x extrême parmi les points de la base sont identifiés. Dans un deuxième temps, les points dont la coordonnée x se situe au maximum à 15 cm des extrêmes sont utilisés pour calculer deux points moyens. Dans un troisième temps, la largeur est calculée comme la distance euclidienne entre les deux points moyens.

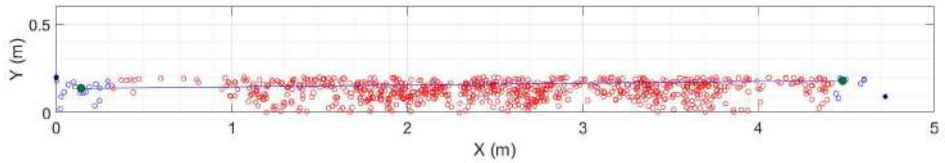



Figure 8.32: Les points de la base du feu dans le plan $X$-$Y$

La distance entre le plan de base et chaque point de la partie supérieure du feu correspond à la hauteur du point. La hauteur de feu moyenne est égale à la hauteur du point moyen calculé en utilisant les points 3D situés au maximum à 30 cm du point le plus élevé.

La longueur du feu est définie par les chercheurs travaillant sur la modélisation du comportement du feu, comme étant la distance entre le sommet de l'avant du feu et le point le plus avancé de la base du feu. Dans ce travail, la longueur est calculée comme la distance euclidienne entre le point moyen ($P_H$) des points 3D situés au maximum à 30 cm du sommet de la flamme et le point moyen ($P_F$) des points 3D au sol les plus avancés (Fig. 8.33).

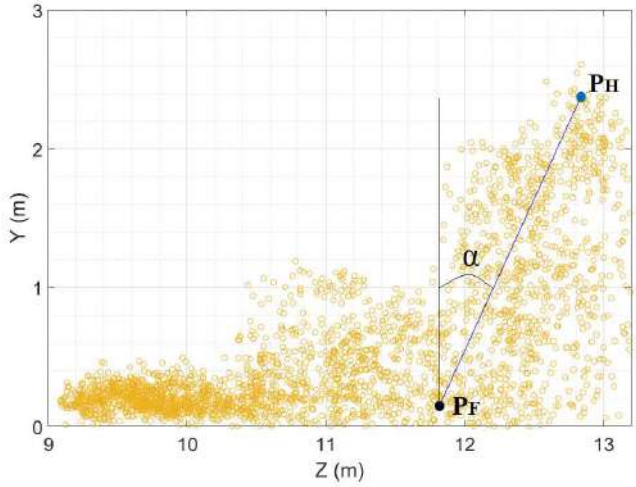



Figure 8.33: Longueur et angle d'inclinaison ($\alpha$) du feu.

L'angle d'inclinaison du feu est égal à l'angle entre le segment $\overline{P_H P_F}$ et la normale du plan de base du feu.

## 8.3 Résultats

En raison du comportement imprévisible et non reproductible du feu, il est difficile d'évaluer l'incertitude de la solution proposée. Les performances du système ont été évaluées à l'aide de deux expériences : l'une utilisant une voiture et l'autre utilisant un feu pseudo-statique. La première expérience consiste en un tour complet effectué par l'UAV autour d'une voiture garée, pour acquérir des images de chaque côté de la voiture. L'inter-distance drone-voiture est entre 10 m et 15 m.

La figure 8.34 montre l'image d'un côté de la voiture prise par le drone.

Figure 8.34: Image d'un côté de la voiture prise par l'UAV.

Seules les images visibles ont été traitées et, comme la méthode choisie pour la détection des pixels de feu n'était pas utilisable, les pixels de la voiture ont été sélectionnées à la main dans les images. À partir des points 3D reconstruits, les dimensions de la voiture ont été calculées et comparées aux valeurs réelles.

La figure 8.35 montre la reconstruction 3D du véhicule obtenu à partir des quatre vues latérales de la voiture. On peut observer que la forme générale du véhicule est compatible avec l'image montrée à la Fig. 8.34.

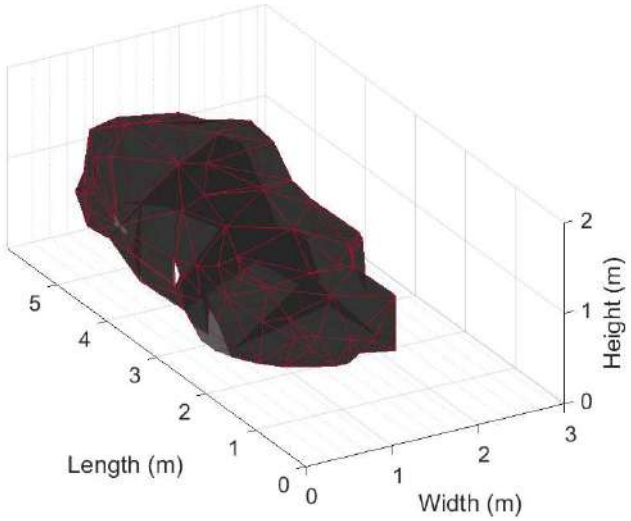



Figure 8.35: Reconstruction 3D d'une voiture.

La fiche technique du véhicule précise que la longueur, la largeur et la hauteur sont respectivement de 3,99 m, 1,64 m et 1,5 m. Les données estimées présentées dans le tableau 8.1 sont : longueur égale à 3,96, largeur égale à 1,62 m et hauteur égale à 1,48 m. L'erreur est donc inférieure à 1,2% dans cette expérience.

| | Length (m) | Width (m) | Height (m) |
|---|---|---|---|
| Real | 3,99 | 1,64 | 1,5 |
| Estimated | 3,96 | 1,62 | 1,48 |
| Error | 2% | 0% | 0% |

Table 8.1: Longueur, largeur et hauteur réelles et estimées, et les erreurs associées.

La différence entre la position GPS de la voiture et celle mesurée est de 0,25 m.

La deuxième expérience a été menée sur un feu pseudo-statique extérieur. La laine de bois a été mise en place sur une surface de 3×5 m. La forme du feu changeait mais pas sa position, avec l'UAV qui se déplaçait en face. La figure 8.36 montre les lignes du front obtenues à trois instants différents. La forme des lignes évolue avec le temps mais reste dans la même position, qui correspond à la réalité.

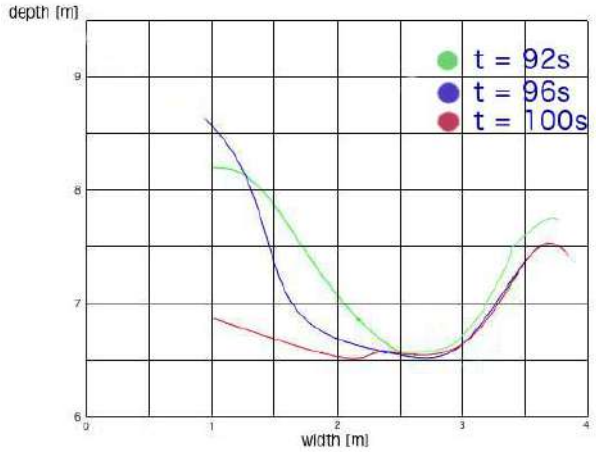
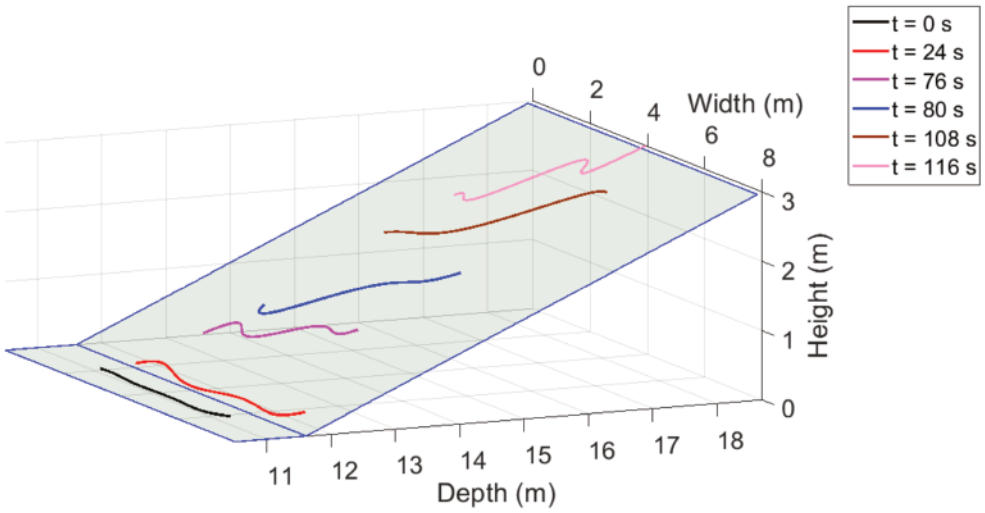


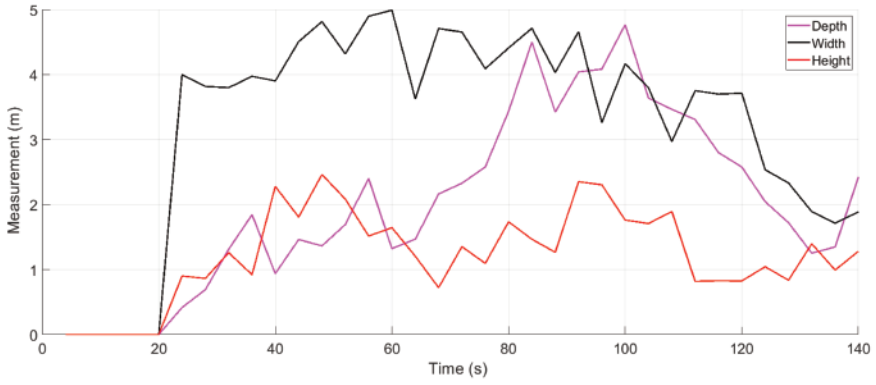

Figure 8.36: Lignes du front estimé du feu pseudo-statique, obtenu à trois instants différents.

Enfin, plusieurs tests ont été menés à l'extérieur sur des zones de brûlage contrôlés. Dans cette section, sont présentées les données estimées correspondant à une propagation du feu sur une zone de 5×10 m, composée d'une première zone plane de 2 m et d'une seconde zone inclinée à 20°.

La figure 8.37 présente l'évolution temporelle de la ligne du front du feu, la profondeur, la largeur et la hauteur obtenues lors de la propagation.




(a)




(b)

Figure 8.37: Evolution temporelle des données estimées. *(a)* Front line. *(b)* Depth, width and height.

## 8.4 Conclusions

Ce document propose un module de stéréovision multimodale adapté aux drones, pour mesurer les caractéristiques géométriques des feux de forêt. Il est composé d'un dispositif de vision, intégrant deux caméras multimodales fonctionnant simultanément dans les spectres LWIR et visible, et un drone. À partir d'images stéréoscopiques multimodales géo référencées, des points de feu 3D sont obtenus, à partir desquels les caractéristiques géométriques du feu sont estimées. Cette solution est capable d'estimer l'évolution temporelle des caractéristiques géométriques d'un feu se propageant sur une distance illimitée. La solution proposée satisfait le besoin

de systèmes capables d'estimer les caractéristiques géométriques du feu, pour comprendre et modéliser son comportement. Elle contribue également aux recherches sur l'utilisation des drones pour combattre les feux de forêt.

Chapter 9

# Appendix

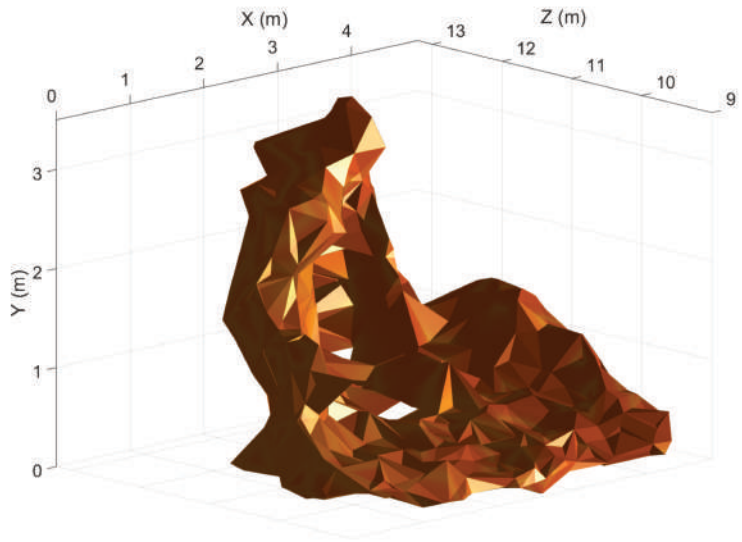
## Introduction

This section aims to summarize some results obtained during this thesis work, but not used for the final purpose of the project. These results may be the starting point for new research in the field of fire.

## 3D points in high resolution

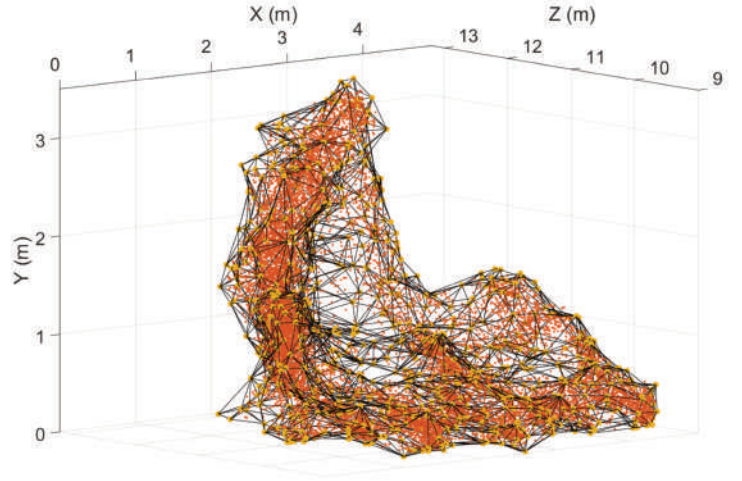
To increase the accuracy of the calculation of the fire geometrical characteristic, an interpolation procedure allowing the increase of the 3D points number is carried out in two steps.

In a first step, a Delaunay triangulation method [150] is applied to the 3D points. Only the tetrahedrons whose radius value of the inscribed sphere is less than 35 cm are kept.



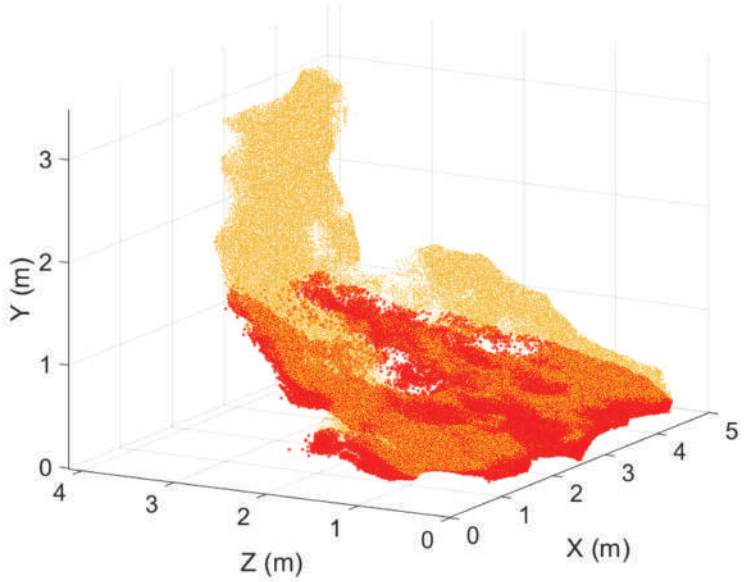

Tetrahedrons produced by the Delaunay triangulation method applied to the 3D fire points.

In a second step, for each tetrahedron it is computed the centroid. The centroid that are inside the 3D shape [151] are added to the set of 3D fire points.

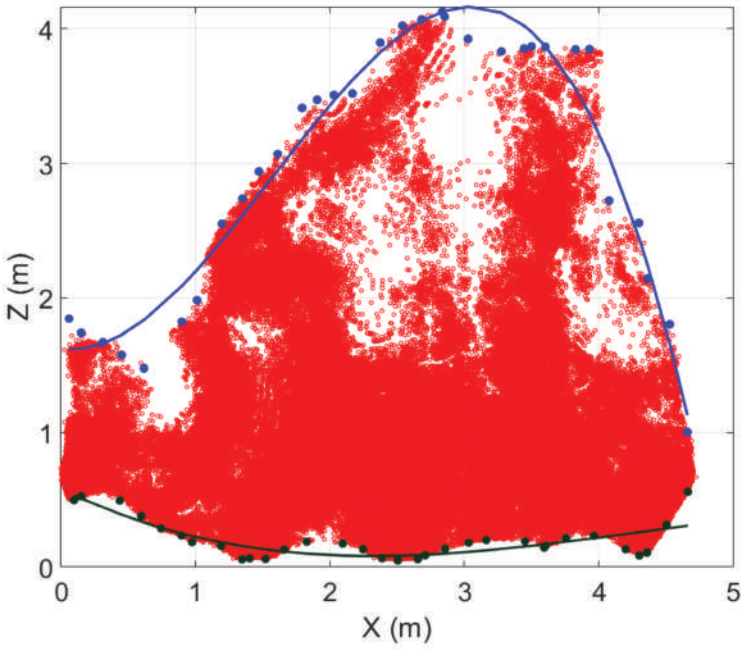



Tetrahedron of the 3D fire points (yellow points), and their centroid (brown points).

This step is repeated 3 times, to obtain a very large resolution of the 3D fire points.



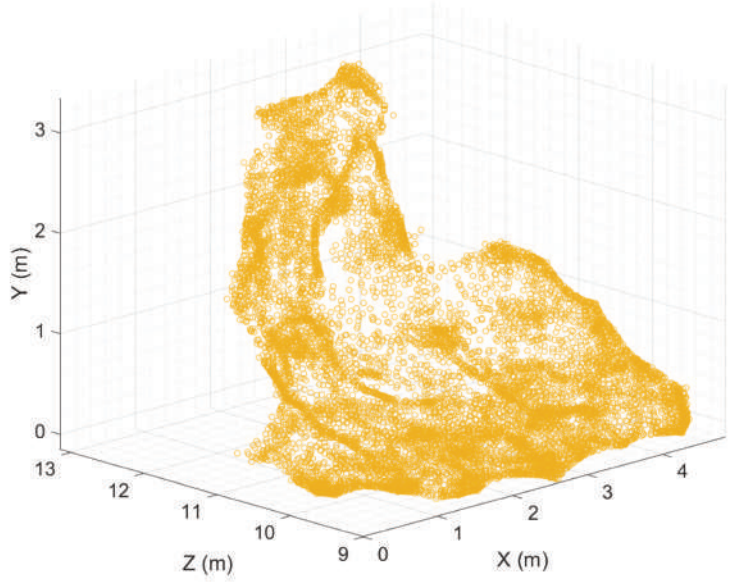

3D fire points in high resolution, composed by 943 415 3D points (red points compose the fire base points).




Fire ground points in high resolution, the blue points are the most advanced points, the green points are the less advanced points.

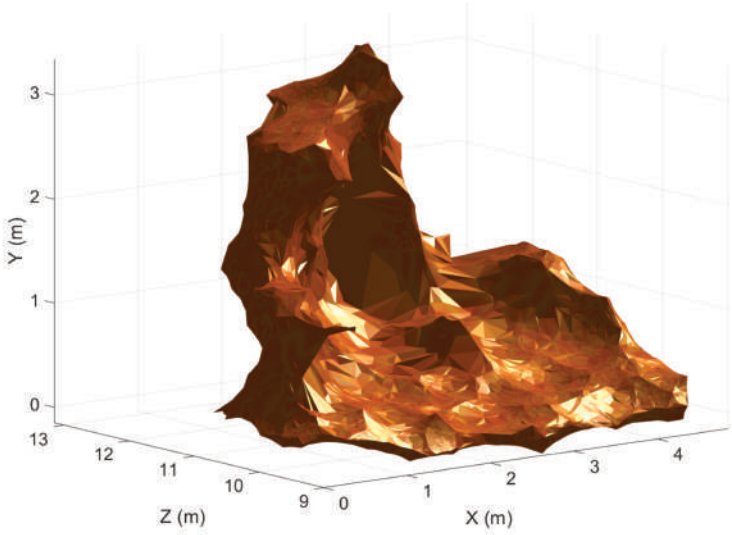
## Virtual fire points to improve the volume computation

The elimination of the outlier points could creates some small empty zone in the shape of the 3D fire points, that influences the volume computation; these zones are also created by matching algorithm, in particular in the zones where the texture is too homogeneous. A method to refill these zone it is applied. First it is used the method to increase the number of 3D points.



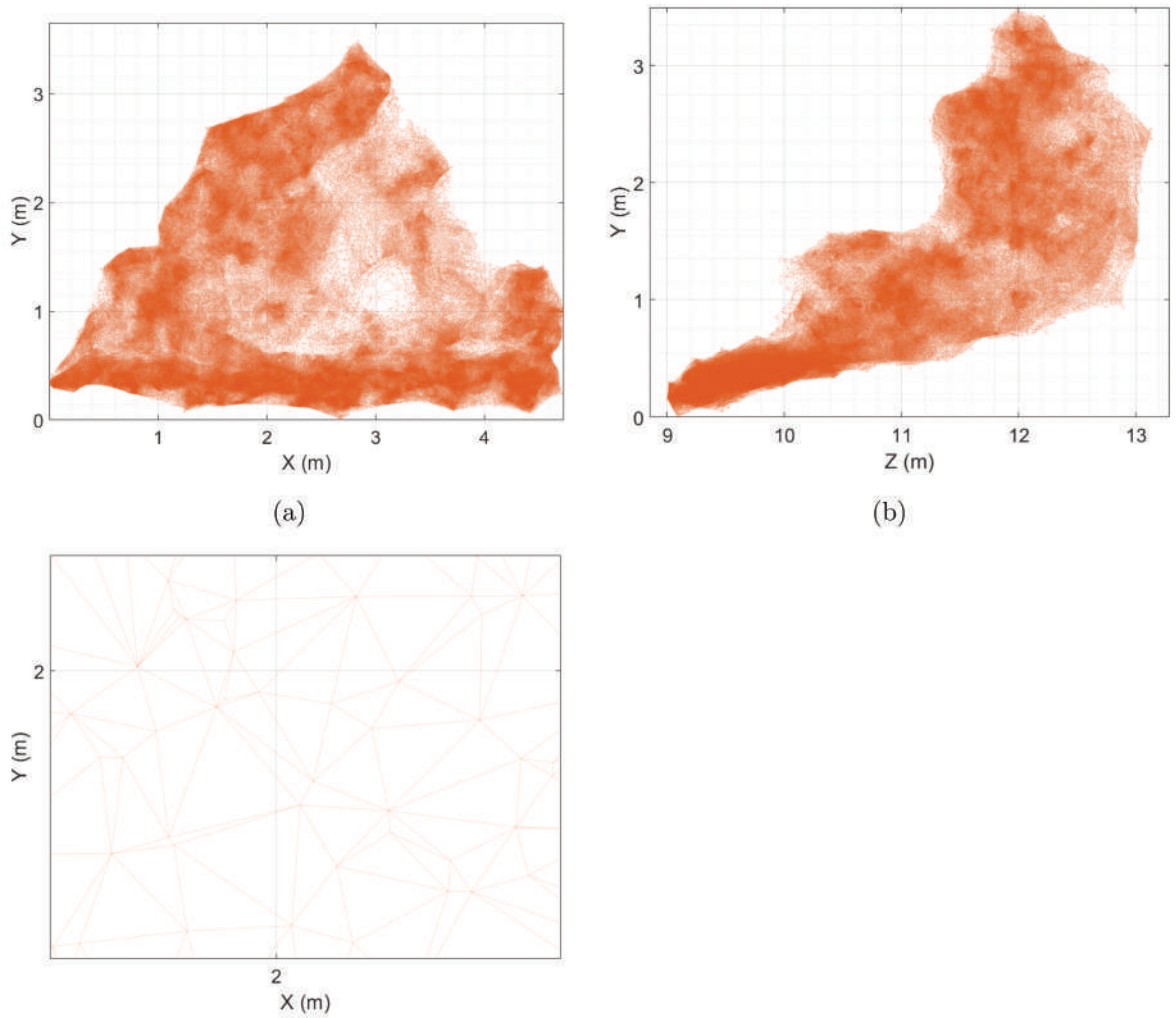

3D fire points in high resolution.

The technique to increase the resolution of the 3D points can only fill the empty zone that are inside the 3D shape, but the 3D shape resulted is too fine.




Alpha shape of the 3D fire points in high resolution.

The second step computes two projections of the 3D points, one on the plane $X$-$Y$, and one on the plane $Z$-$Y$. On these 2 projections are computed the alpha shape, and on these shapes a Delaunay triangulation.

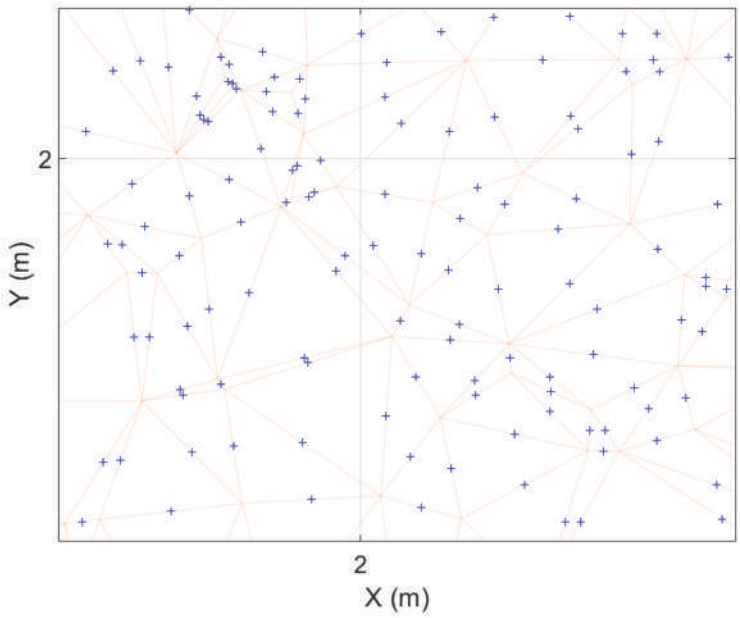



(a) (b)

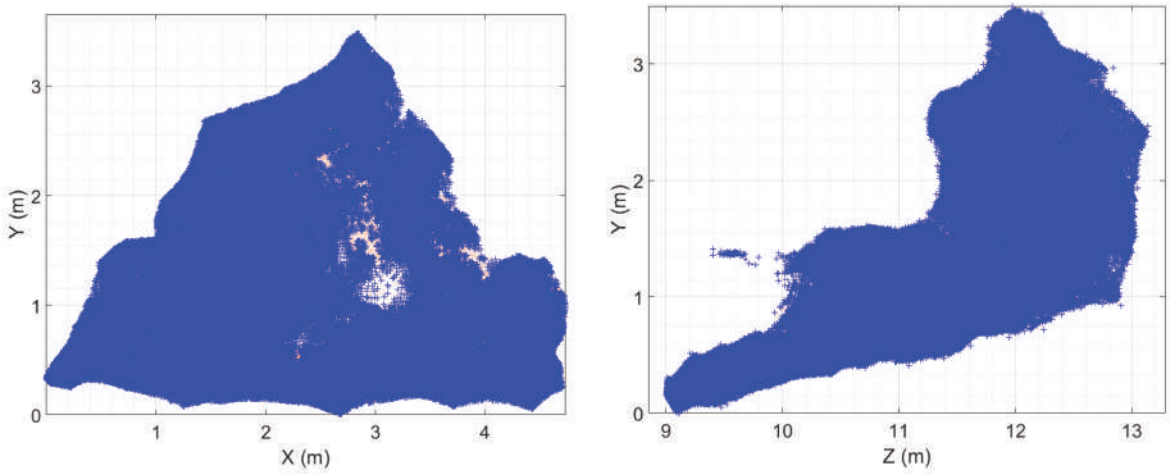
(c)

Delaunay triangulation applied on the alpha shape of the 3D points in high resolution. ($a$) Projection of the triangles on the $X$-$Y$ plane. ($b$) Projection of the triangles on the $Z$-$Y$ plane. ($c$) Zoom of the image ($a$).

In the third step, for each each triangles in each projection, they are computed the middle points of the edges.



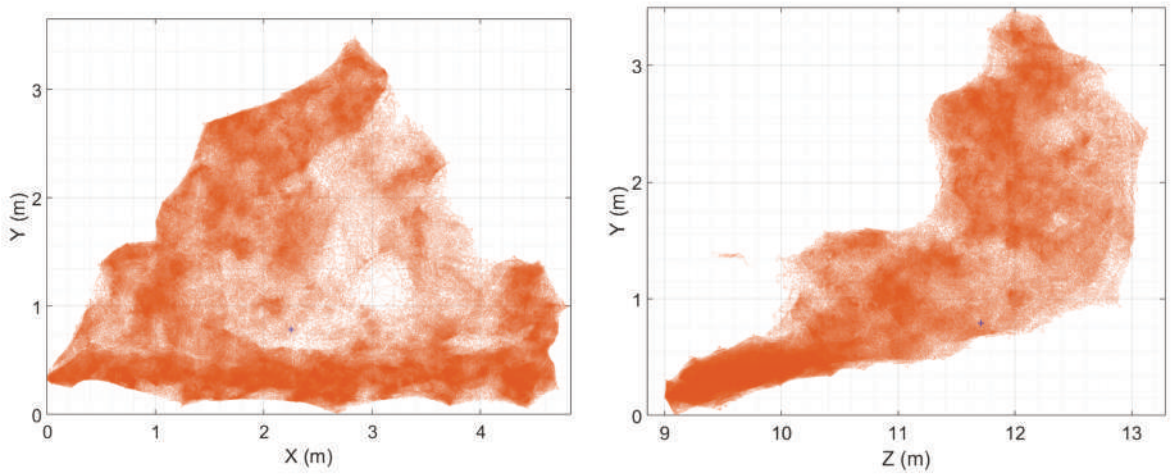

Middle points of the triangles edges (blue crosses).




(a) (b)

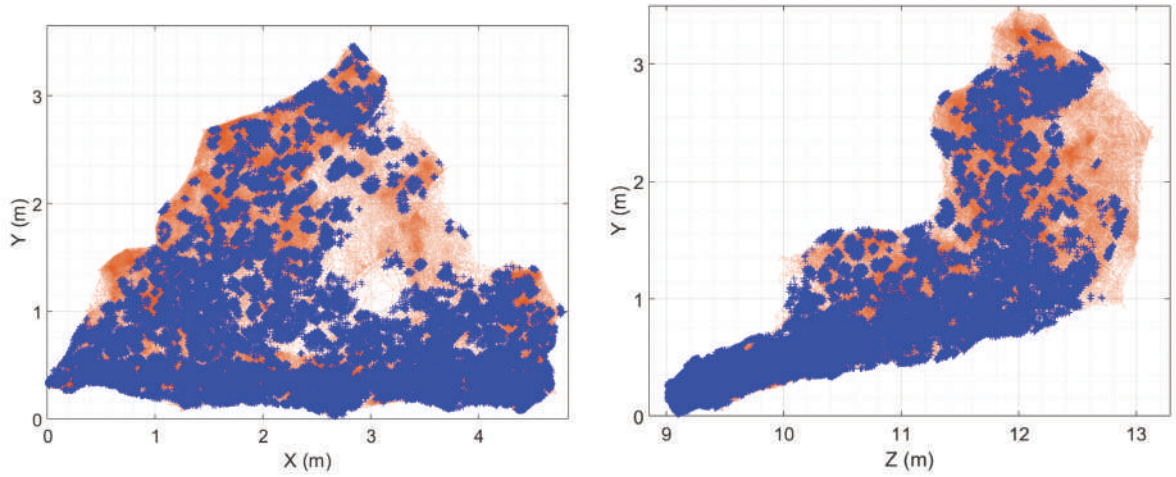
Triangles on the $X$-$Y$ plane and on the $Z$-$Y$ plane where the blue crosses represent the middle points of the triangles edges. ($a$) Projection on the $X$-$Y$ plane. ($b$) Projection on the $Z$-$Y$ plane.

In the fourth step, for each point in the $X$-$Y$ plane, $(x_i,\ y_i)$ it is verified if exist an equivalent point in the $Z$-$Y$ plane with the $y_i$ coordinate.




(a) (b)

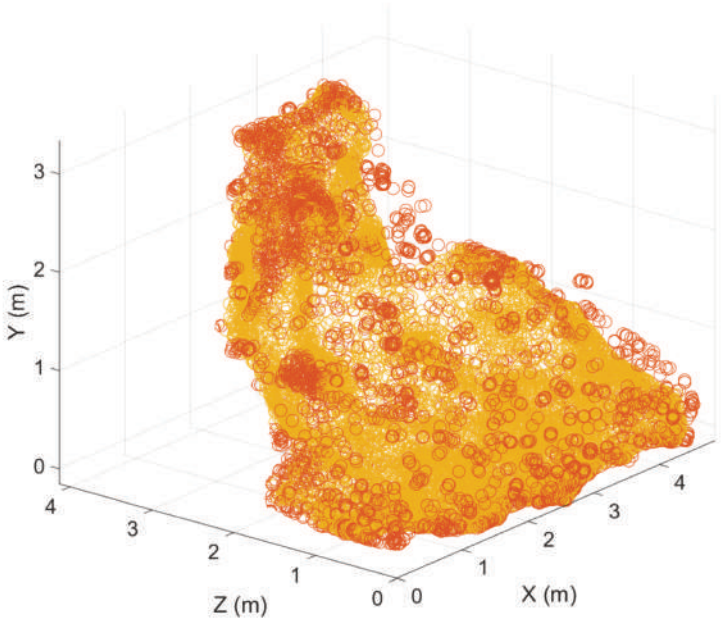
Example on one equivalent point (blue cross). (*a*) Equivalent point on the *X*-*Y* plane. (*b*) Equivalent point on the *Z*-*Y* plane.




(a) (b)

Equivalent points in the two projections (blue crosses). (*a*) Equivalent points on the *X*-*Y* plane. (*b*) Equivalent points on the *Z*-*Y* plane.

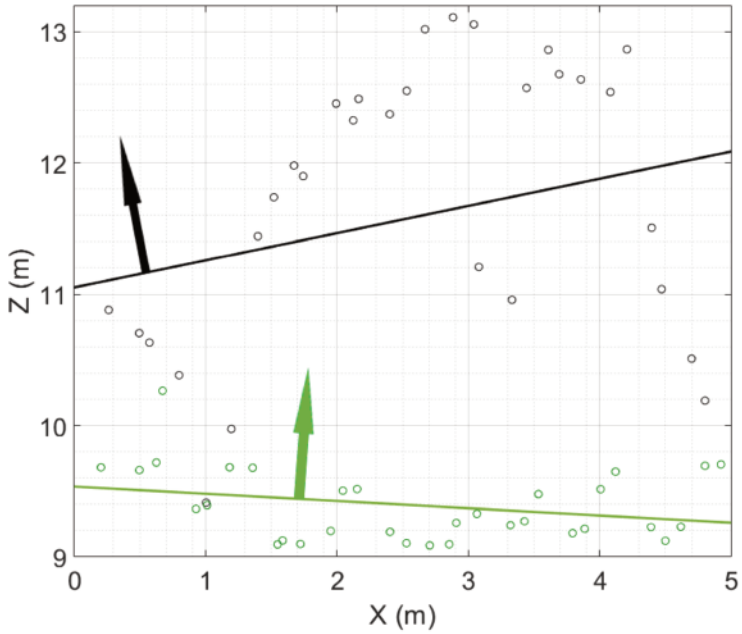
For each equivalent point it is possible to obtain the three coordinates ($x_i$, $y_i$, $z_i$), and in the final step, these points are added to the 3D point in high resolution.




3D fire points in high resolution (yellow circles), and equivalent points (brown circles).

## Fire front line and fire back line directions

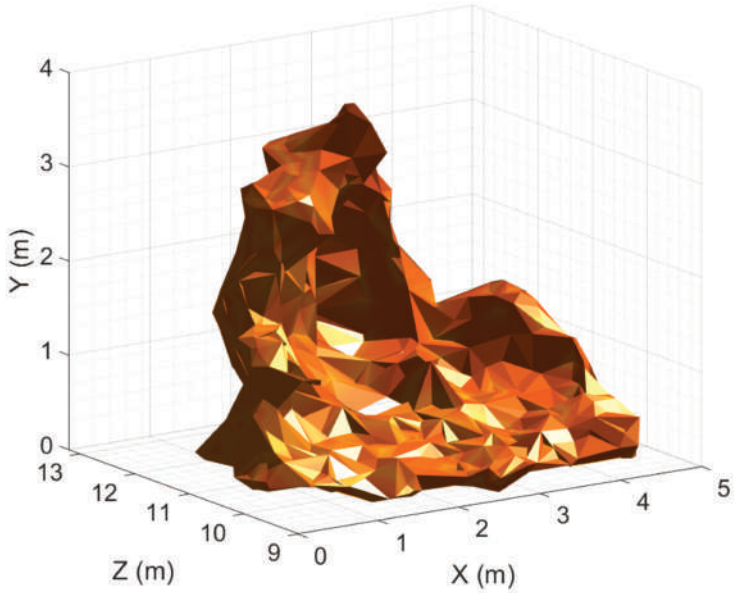
To calculate this direction, they are considered the sets of the fire ground points. The angular coefficient of the straight line that interpolates $Pts_{Front}$ is the local direction of the front fire, and similarly the angular coefficient of the straight line that interpolates $Pts_{Back}$ is the local direction of the back fire.




Fire local direction computed as the normal vector of straight line of the front fire points (black line) and straight line of the back fire points (green line).

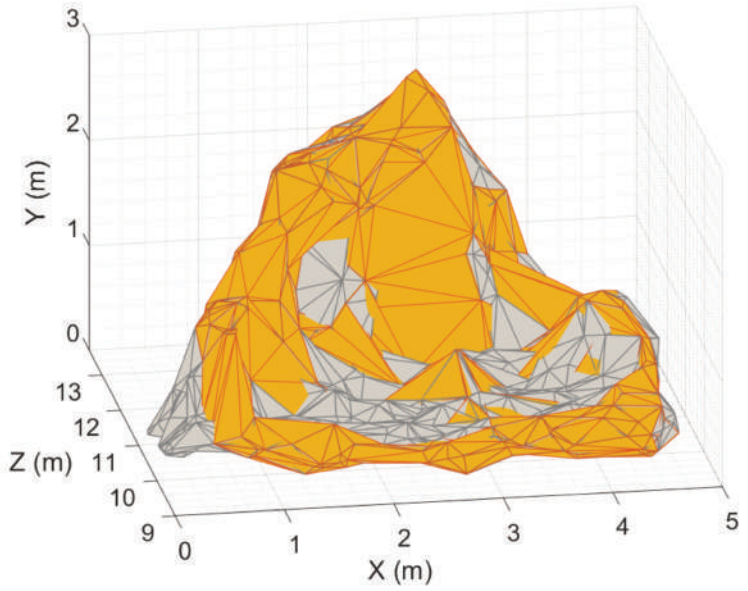
## Frontal flame surface

It is possible to compute the surface of the front fire flame and the back fire flame.

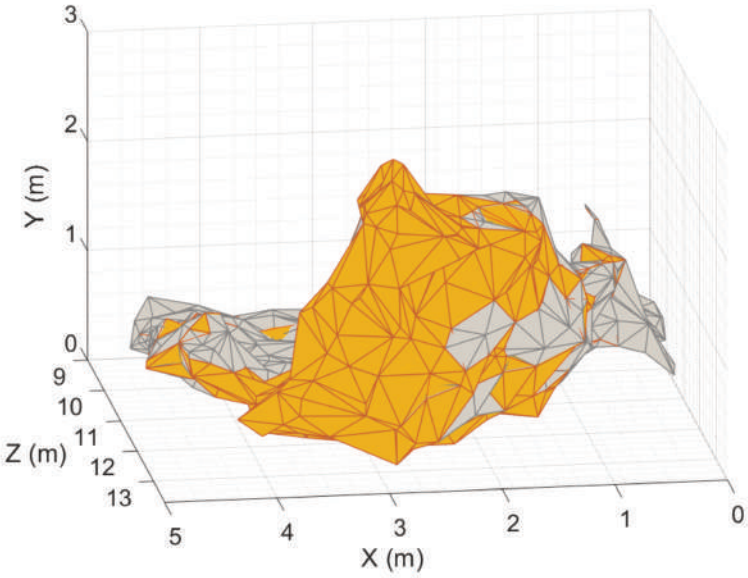



Surfacing of the 3D fire points using a Delaunay triangulation.

In order to obtain the surface of the fire give them this information, only the triangles of the fire surfaces that are not masked by the others and which their normal are colinear to the $z$ axis of the Local direction frame constitute the front part of the front from which the area calculation is carried out. These triangles have to satisfy the property that the half-line collinear with the axis of the depth and whose origin is the center of the triangle does not intercept any other triangle. Figure **??** shows the triangles, in orange color, used for the calculation of the frontal surface if the target is positioned in front of the fire, along its propagation. In this case the rear part of the fire, in gray color, is not taken into account.



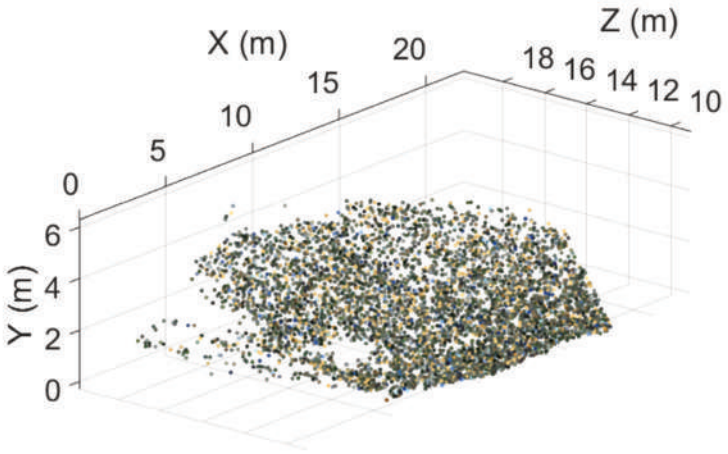

Triangles selected to compute the area of the back fire flame .



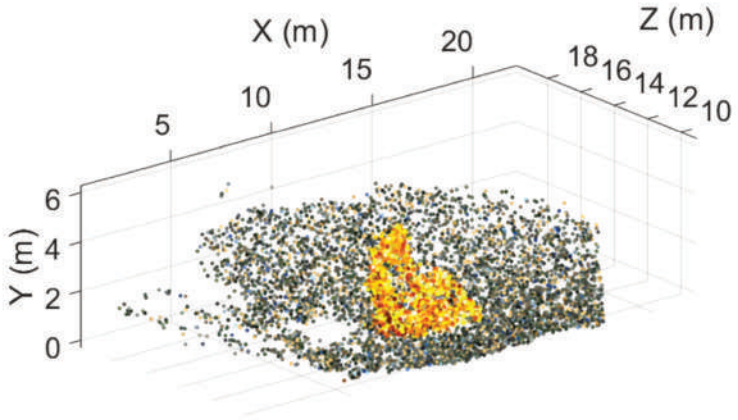

Triangles selected to compute the area of the front fire flame.

## 3D reconstruction of the background

It was possible to obtain 3D reconstruction of the environment around the fire, by performing a feature point detection on the inverse of the segmented image. This results in a draft of a DEM map, on which 3D reconstructions of the fire can be placed. In the future it will be possible to consider this kind of reconstruction to apply measurements to the vegetation around the fire as well. This type of reconstruction could be useful in case you want to create a virtual environment with a fire inside.



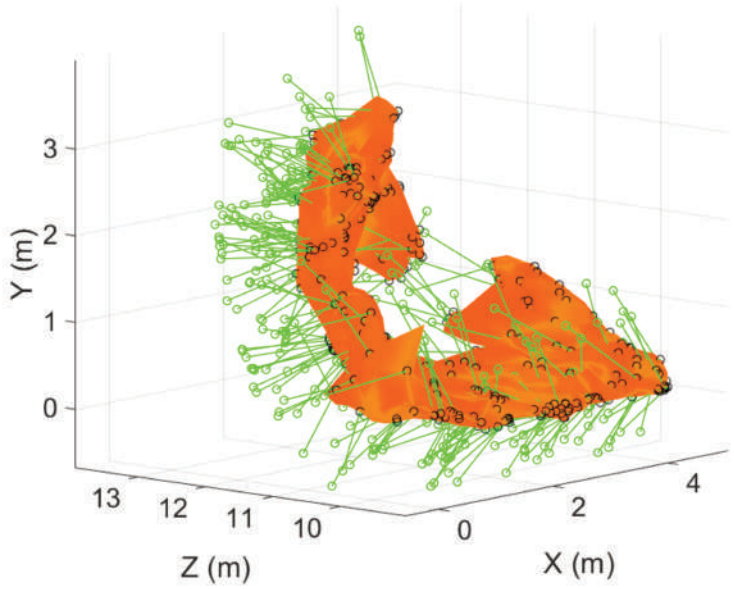

3D reconstruction of the fire combustible area.



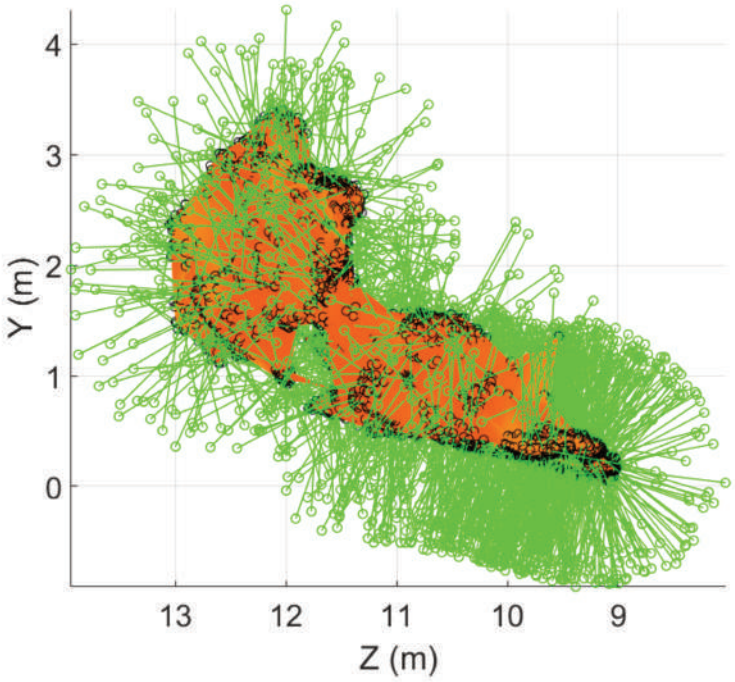

3D fire reconstruction over the background reconstructed.

## Surface normals

It is possible to calculate the normals to the fire surface through all the 3D points. This could be the direction of the energy emitted by the flame at that point. It could be assumed that knowing the point temperature of the fire, it could be possible to calculate the amount of heat emitted in each direction on each 3D fire points.




3D fire reconstruction with a selection of the normal of the surface through the 3D points (green line).




3D fire reconstruction with all the normal of the surface through the 3D points.

# Bibliography


[1] M. E. Alexander. *Proposed revision of fire danger class criteria for forest and rural areas in New Zealand. Second edition [reprint with corrections]*. National Rural Fire Authority, 2008.

[2] *Enlisting communities in wildfire prevention*. URL: http://www.fao.org/news/story/en/item/41230/icode/ (visited on 07/12/2019).

[3] *Forest-fires*. URL: https://www.nrcan.gc.ca/our-natural-resources/forests-and-forestry/wildland-fires-insects-and-disturbances/forest-fires/13143 (visited on 07/03/2019).

[4] *EFFIS*. URL: http://effis.jrc.ec.europa.eu/ (visited on 07/05/2019).

[5] S. Lahaye et al. "What are the drivers of dangerous fires in Mediterranean France?" In: *Int. J. Wildland Fire* 27.3 (2018), p. 155.

[6] *Prométhée : Liste des incendies*. URL: http://www.promethee.com/default/incendies (visited on 07/03/2019).

[7] *California state wildfire map*. URL: https://sf.curbed.com/maps/california-wildfire-map-2018-fall-camp-fire-november (visited on 07/02/2019).

[8] *He tried to plug a wasp nest. He ended up sparking California's biggest wildfire*. URL: https://www.nytimes.com/2019/06/11/us/california-ranch-wildfire-wasp-nest.html (visited on 10/07/2019).

[9] C. R. Wootson Jr. "The deadliest, most destructive wildfire in California's history has finally been contained". In: *Washington Post* (2018).

[10] *List of wildfires*. In: *Wikipedia*. URL: https://en.wikipedia.org/w/index.php?title=List_of_wildfires&oldid=901848711 (visited on 07/01/2019).

[11] *2018 was California's worst year of fire ever, federal report confirms*. URL: https://www.latimes.com/local/lanow/la-me-ln-california-fires-record-report-20190309-story.html (visited on 07/01/2019).

[12] W. S. Folkman. *Fire prevention in Butte County, California... evaluation of an experimental program*. Tech. rep. 1973.

[13] *Sweden battles forest fires north of Stockholm - BBC News*. URL: https://www.bbc.com/news/av/world-europe-44850836/sweden-battles-forest-fires-north-of-stockholm (visited on 07/02/2019).

[14] *Sweden's 'chaotic' response to historic wildfires criticized - The Local*. URL: https://www.thelocal.se/20190208/swedens-chaotic-response-to-historic-wildfires-criticized (visited on 07/02/2019).

[15] *Wildfire risk in Sweden as spring returns - The Local*. URL: https://www.thelocal.se/20190311/wildfire-risk-in-sweden-as-spring-returns (visited on 07/02/2019).

[16] *Scandinavie: les premiers feux de forêt font craindre un nouvel été de cendres.* URL: http://www.lefigaro.fr/international/scandinavie-les-premiers-feux-de-foret-font-craindre-un-nouvel-ete-de-cendres-20190424 (visited on 07/15/2019).

[17] *Portugal fires: wildfire rages as Portugal deals with 46C heat.* URL: https://www.express.co.uk/news/world/999672/portugal-fires-algarve-evacuation-zones-maps-latest-warnings (visited on 07/01/2019).

[18] *Wildfires kill dozens on Greek coast.* URL: https://www.bbc.com/news/world-europe-44932366 (visited on 07/03/2019).

[19] *France Hit by fires in south as heatwave reaches peak.* URL: https://www.bloomberg.com/news/articles/2019-06-29/france-hit-by-fires-in-south-as-june-s-heatwave-reaches-peak (visited on 07/03/2019).

[20] *'Carelessness' blamed for Corsica's winter wildfires.* URL: https://www.thelocal.fr/20190226/carelessness-blamed-for-corsicas-winter-wildfires (visited on 07/03/2019).

[21] E. Rigolot. "Impact du changement climatique sur les feux de forêt". In: *Association Forêt Méditerranéenne, 14 rue Louis Astouin, 13002 MARSEILLE, France* (2008).

[22] C. Chatry et al. *Changement climatique et extension des zones sensibles aux feux de forêts.* Tech. rep. Rapport de mission interministérielle, 2010.

[23] *IAWF Issue paper: extreme fires.* URL: https://www.iawfonline.org/article/iawf-issue-paper-extreme-fires/ (visited on 07/03/2019).

[24] J. H. Balbi et al. "A 3D physical real-time model of surface fires across fuel beds". In: *Combustion Science and Technology* 179.12 (), pp. 2511–2537.

[25] J. H. Balbi et al. "Physical modeling of surface fire under nonparallel wind and slope conditions". In: *Combustion Science and Technology* 182.7 (), pp. 922–939.

[26] J. L. Rossi et al. "An analytical model based on radiative heating for the determination of safety distances for wildland fires". In: *Fire Safety Journal* 46.8 (2011), pp. 520–527.

[27] J. H. Balbi et al. "Modelling of eruptive fire occurrence and behaviour". In: *Journal of Environmental Science and Engineering* 3 (2014), pp. 115–132.

[28] F. J. Chatelon et al. "A convective model for laboratory fires with well-ordered vertically-oriented fuel beds". In: *Fire Safety Journal* 90 (2017), pp. 54–61.

[29] A. G. McArthur. "Weather and grassland fire behavior." In: *Forestry and Timber Bureau.* Leaflet.100 (1966).

[30] F. A. Albini. "A model for fire spread in wildland fuels by-radiation". In: *Combustion Science and Technology* 42.5-6 (1984), pp. 229–258.

[31] P. M. Fernandes, W. R. Catchpole, and F. C. Rego. "Shrubland fire behaviour modelling with microplot data". In: *Canadian Journal of Forest Research* 30.6 (2000), pp. 889–899.

[32] R. C. Rothermel. "A mathematical model for predicting fire spread in wildland fuels". In: *Res. Pap. INT-115. Ogden, UT: US Department of Agriculture, Intermountain Forest and Range Experiment Station. 40 p.* 115 (1972).

[33] A. M. Grishin. *Mathematical modeling of forest fires and new methods of fighting them.* House of the Tomsk University, 1997.

[34] J. M. C. Mendes-Lopes, J. M. P. Ventura, and J. M. P. Amaral. "Flame characteristics, temperature-time curves, and rate of spread in fires propagating in a bed of pinus pinaster needles". In: *Int. J. Wildland Fire* 12.1 (2003), p. 67.

[35] E. Pastor et al. "Mathematical models and calculation systems for the study of wildland fire behaviour". In: *Progress in Energy and Combustion Science* 29.2 (2003), pp. 139–153.

[36] D. X. Viegas. "On the existence of a steady state regime for slope and wind driven fires". In: *Int. J. Wildland Fire* 13.1 (2004), p. 101.

[37] R. J. Sneeuwjagt and W. H. Frandsen. "Behavior of experimental grass fires vs. predictions based on Rothermel's fire model". In: *Can. J. For. Res.* 7.2 (1977), pp. 357–367.

[38] N.P. Cheney, J.S Gould, and W.R. Catchpole. "The influence of fuel, weather and fire shape variables on fire-spread in grasslands". In: *International Journal of Wildland Fire* 3.1 (1993), pp. 31–44.

[39] N.P. Cheney and J.S. Gould. "Fire growth in grassland fuels". In: 5.4 (1995), pp. 237–247.

[40] W. B. Ng and Y. Zhang. "Stereoscopic imaging and reconstruction of the 3D geometry of flame surfaces". In: *Exp Fluids* 34.4 (2003), pp. 484–493.

[41] J. R. Martinez-de Dios et al. "Laboratory fire spread analysis using visual and infrared images". In: *Int. J. Wildland Fire* 15.2 (2006), p. 179.

[42] L. Rossi et al. "Estimating the surface and volume of laboratory-scale wildfire fuel using computer vision". In: *IET Image Processing* 6.8 (2012), pp. 1031–1040.

[43] S. Verstockt et al. "FireCube: a multi-view localization framework for 3D fire analysis". In: *Fire Safety Journal* 46.5 (2011), pp. 262–275.

[44] E. Pastor et al. "Computing the rate of spread of linear flame fronts by thermal image processing". In: *Fire Safety Journal* 41.8 (2006), pp. 569–579.

[45] P. S. Mason et al. "Estimating Thermal Radiation Fields from 3D Flame Reconstruction". In: *Fire Technol* 45.1 (2009), pp. 1–22.

[46] C. Pinto et al. "Fire whirls in forest fires: An experimental analysis". In: *Fire Safety Journal* 87 (2017), pp. 37–48.

[47] J. M. Johnston et al. "Flame-Front Rate of Spread Estimates for Moderate Scale Experimental Fires Are Strongly Influenced by Measurement Approach". In: *Fire* 1.1 (2018), p. 16.

[48] L. Rossi et al. "Advanced stereovision system for fire spreading study". In: *Fire Safety Journal* 60 (2013), pp. 64–72.

[49] J. R. Martinez-de Dios et al. "Computer vision techniques for forest fire perception". In: *Image and Vision Computing* 26.4 (2008), pp. 550–562.

[50] H. B. Clements, D. E. Ward, and C. W. Adkins. "Measuring fire behavior with photography". In: *Photogrammetric Engineering and Remote Sensing* 49.2 (1983), pp. 213–217.

[51] J. R. Martinez-de Dios et al. "Automatic forest-fire measuring using ground stations and unmanned aerial systems". In: *Sensors* 11.6 (2011), pp. 6328–6353.

[52] L. Merino, J. R. Martinez-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.

[53] L. Rossi, M. Akhloufi, and Y. Tison. "On the use of stereovision to develop a novel instrumentation system to extract geometric fire fronts characteristics". In: *Fire Safety Journal* 46.1 (2011), pp. 9–20.

[54] T. Toulouse et al. "A multimodal 3D framework for fire characteristics estimation". In: *Meas. Sci. Technol.* 29.2 (2018).

[55] M. E. Alexander. "Calculating and interpreting forest fire intensities". In: *Canadian Jornal of Botany* 60.4 (), pp. 349–357.

[56] W. Anderson et al. "Evaluating models to estimate flame characteristics for free-burning fires using laboratory and field data". In: *Forest Ecology and Management* 234 (2006), S77.

[57] J. R. Howell, M. P. Menguc, and R. Siegel. *Thermal radiation heat transfer*. CRC press, 2015.

[58] J. L. Rossi, F J. Chatelon, and T. Marcelli. "Fire Intensity". In: *Encyclopedia of Wildfires and Wildland-Urban Interface (WUI) Fires* (2019).

[59] E. D. Breejen et al. "Infrared measurements of energy release and flame temperatures of forest fires". In: *Associacao para o Desenvolvimento da Aerodinamica Industrial* (1998).

[60] T. W. Ridler and S. Calvard. "Picture Thresholding Using an Iterative Selection Method". In: *IEEE Transactions on Systems, Man, and Cybernetics* 8.8 (1978), pp. 630–632.

[61] W. Phillips III, M. Shah, and N. Da Vitoria Lobo. "Flame recognition in video". In: *Pattern Recognition Letters* 23.1 (2002), pp. 319–327.

[62] J. R. Martinez-de Dios and A. Ollero. *A Multiresolution Threshold Selection Method Based on Training*. Springer, 2004.

[63] J. R. Martinez-de Dios and A. Ollero. "A Multiresolution-Fuzzy Method for Robust Threshold Selection in Image Segmentation". In: *Intelligent Automation & Soft Computing* 12.4 (2006), pp. 419–430.

[64] A. R. Conn, N. Gould, and P. L. Toint. "Large-scale Nonlinear Constrained Optimization: a Current Survey". In: *Algorithms for Continuous Optimization*. Springer, 1994, pp. 287–332.

[65] T. Coleman and Y. Li. "An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds". In: *Journal on Optimization* 6.2 (1996), pp. 418–445.

[66] L. Rossi et al. "A 3D vision system for the measurement of the rate of spread and the height of fire fronts". In: *Measurement Science and Technology* 21.10 (2010).

[67] L. Rossi and M. Akhloufi. "Dynamic fire 3D modeling using a real-time stereovision system". In: *Technological Developments in Education and Automation*. 2010.

[68] E. Trucco and A. Verri. *Introductory techniques for 3-D computer vision*. Prentice Hall Englewood Cliffs, 1998.

[69] L. Rossi et al. "Measurement of the geometric characteristics of a fire front by stereovision techniques on field experiments". In: *Measurement Science and Technology* 22.12 (2011).

[70] T. Toulouse. "Estimation par stéréovision multimodale de caractéristiques géométriques d'un feu de végétation en propagation". PhD thesis. 2015.

[71] Pierre Bézier. *L'utilisation des courbes et surfaces en CAO*. Hermès, 1988.

[72] C Bradford Barber et al. "The quickhull algorithm for convex hulls". In: *Transactions on Mathematical Software* 22.4 (1996), pp. 469–483.

[73] *JAI industrial 2-sensor prism-based color/NIR area scan camera*. URL: `//www.jai.com/products/ad-080-ge` (visited on 08/12/2019).

[74] M. Brown, R. Szeliski, and S. Winder. *Multi-image matching using multi-scale oriented patches*. Tech. rep. Microsoft, 2004.

[75] L. Rossi et al. "Estimation of spreading fire geometrical characteristics using near infrared stereovision". In: *Three-Dimensional Image Processing (3DIP) and Applications 2013*. International Society for Optics and Photonics. 2013.

[76] A. E. Ononye, A. Vodacek, and E. Saber. "Automated extraction of fire line parameters from multispectral infrared images". In: *Remote Sensing of Environment* 108.2 (2007), pp. 179–188.

[77] R. Paugam, M. J. Wooster, and G. Roberts. "Use of Handheld Thermal Imager Data for Airborne Mapping of Fire Radiative Power and Energy and Flame Front Rate of Spread". In: *IEEE Transactions on Geoscience and Remote Sensing* 51.6 (2013), pp. 3385–3399.

[78] C. Yuan, Y. Zhang, and Z. Liu. "A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques". In: *Canadian journal of forest research* 45.7 (2015), pp. 783–792.

[79] J. R. Matinez-de-Dios et al. "Multi-UAV experiments: application to forest fires". In: *Multiple Heterogeneous Unmanned Aerial Vehicles*. Springer, 2007, pp. 207–228.

[80] A. Ollero and L. Merino. "Unmanned aerial vehicles as tools for forest-fire fighting". In: *Forest Ecology and Management* 234.1 (2006), S263.

[81] J. M. Robinson. "Fire from space: Global fire evaluation using infrared remote sensing". In: *International Journal of Remote Sensing* 12.1 (1991), pp. 3–24.

[82] Y. Billaud et al. "Determination of woody fuel flame properties by means of emission spectroscopy using a genetic algorithm". In: *Combustion Science and Technology* 185.4 (2013), pp. 579–599.

[83] Y. Le Maoult et al. “Fire detection: a new approach based on a low cost CCD camera in the near infrared”. In: *Process Safety and Environmental Protection* 85.3 (2007), pp. 193–206.

[84] S. Briz et al. “Reduction of false alarm rate in automatic forest fire infrared surveillance systems”. In: *Remote Sensing of Environment* 86.1 (2003), pp. 19–29.

[85] *MIVIM*. URL: `http://mivim.gel.ulaval.ca/dynamique/` (visited on 08/29/2019).

[86] J. Mendes Lopes et al. “Fire Star: a decision support system for fuel management and fire hazard reduction in mediterranean wildland-urban interfaces. Fire Star behaviour model of wildland fire. Experimental values of basic parameters”. In: *Fire Star: a decision support system for fuel management and fire hazard reduction in mediterranean wildland-urban interfaces.* (2003).

[87] B. Gouverneur et al. “Archeological treasures protection based on early forest wildfire multi-band imaging detection system”. In: *9th Electro-Optical and Infrared Systems: Technology and Applications*. International Society for Optics and Photonics. 2012.

[88] *LI-OV580-STEREO*. URL: `https://www.leopardimaging.com/uploads/LI-OV580-STEREO_datasheet.pdf` (visited on 09/05/2019).

[89] *FLIR Vue Pro R*. URL: `https://www.flir.com/products/vue-pro-r/` (visited on 08/28/2019).

[90] *u-blox M8*. URL: `https://www.u-blox.com/sites/default/files/products/documents/u-blox8-M8_ReceiverDescrProtSpec_%5C%28UBX-13003221%5C%29_Public.pdf` (visited on 07/03/2019).

[91] *Introduction · MAVLink Developer Guide*. URL: `https://mavlink.io/en/` (visited on 09/30/2019).

[92] *DJI – Le leader mondial des drones civils et de la technologie d'imagerie aérienne*. URL: `https://www.dji.com/fr/spreading-wings-s1000` (visited on 09/06/2019).

[93] K. J. Åström and T. Hägglund. *PID controllers: theory, design, and tuning*. Instrument society of America Research Triangle Park, NC, 1995.

[94] M. Akhloufi, T. Toulouse, and L. Rossi. “Multiple spectrum vision for wildland fires”. In: *7th International Conference on Image Processing Theory, Tools and Applications*. IEEE. 2017.

[95] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2003.

[96] J. G. Fryer and D. C. Brown. “Lens distortion for close-range photogrammetry”. In: *Photogrammetric engineering and remote sensing* 52.1 (1986), pp. 51–58.

[97] R. Y. Tsai and T. S. Huang. “Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces”. In: *IEEE Transactions on pattern analysis and machine intelligence* 1 (1984), pp. 13–27.

[98] D. W. Marquardt. “An algorithm for least-squares estimation of nonlinear parameters”. In: *Journal of the society for Industrial and Applied Mathematics* 11.2 (1963), pp. 431–441.

[99] Z. Zhang. "A flexible new technique for camera calibration". In: *IEEE Transactions on pattern analysis and machine intelligence* 22.11 (2000), pp. 1330–1334.

[100] H. C. Longuet-Higgins. "A computer algorithm for reconstructing a scene from two projections". In: *Nature* 293.5828 (1981), pp. 133–135.

[101] D. Nistér. "An efficient solution to the five-point relative pose problem". In: *IEEE transactions on pattern analysis and machine intelligence* 26.6 (2004), pp. 756–770.

[102] P. Stefanovic. "Relative orientation–a new approach". In: *ITC Journal* 3 (1973), pp. 417–448.

[103] R. Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

[104] G. Gales. "Mise en correspondance de pixels pour la stéréovision binoculaire par propagation d'appariements de points d'intérêt et sondage de régions". PhD thesis. 2011.

[105] C. G. Harris and M. Stephens. "A combined corner and edge detector." In: *Alvey vision conference*. Citeseer. 1988.

[106] E. Rosten and T. Drummond. "Machine learning for high-speed corner detection". In: *European conference on computer vision*. Springer. 2006.

[107] H. Bay, T. Tuytelaars, and L. Van Gool. "Surf: Speeded up robust features". In: Springer. 2006.

[108] J. Matas et al. "Robust wide-baseline stereo from maximally stable extremal regions". In: *Image and vision computing* 22.10 (2004), pp. 761–767.

[109] J. Shi. "Good features to track". In: *Conference on computer vision and pattern recognition*. IEEE. 1994.

[110] F. Ali et al. "A comparison of FAST, SURF, Eigen, Harris, and MSER features". In: *International Journal of Computer Engineering and Information Technology* 8.6 (2016), p. 100.

[111] J. A. Noble. "Finding corners". In: *Image and vision computing* 6.2 (1988), pp. 121–128.

[112] D. G. Lowe. "Distinctive image features from scale-invariant keypoints". In: *International journal of computer vision* 60.2 (2004), pp. 91–110.

[113] K. Mikolajczyk and C. Schmid. "Scale & affine invariant interest point detectors". In: *International journal of computer vision* 60.1 (2004), pp. 63–86.

[114] F. Schaffalitzky and A. Zisserman. "Multi-view matching for unordered image sets, or "How do I organize my holiday snaps?"" In: *European conference on computer vision*. Springer. 2002.

[115] C. Strecha et al. "LDAHash: Improved matching with smaller descriptors". In: *IEEE transactions on pattern analysis and machine intelligence* 34.1 (2011), pp. 66–78.

[116] J. Van De Weijer and C. Schmid. "Coloring local feature extraction". In: *European conference on computer vision*. Springer. 2006.

[117] S. Chambon and A. Crouzil. “Similarity measures for image matching despite occlusions in stereo vision”. In: *Pattern Recognition* 44.9 (2011), pp. 2063–2075.

[118] J. Y. Bouguet. *Camera calibration toolbox for matlab*. URL: `http://www.vision.caltech.edu/bouguetj/calib%5C_doc/index.html` (visited on 02/11/2020).

[119] T. Toulouse et al. “Benchmarking of wildland fire color segmentation algorithms”. In: *IET Image Processing* 9.12 (2015), pp. 1064–1072.

[120] T. Celik et al. “Fire detection using statistical color model in video sequences”. In: *Journal of Visual Communication and Image Representation* 18.2 (2007), pp. 176–185.

[121] B. C. Ko, K. H. Cheong, and J. Y. Nam. “Fire detection based on vision sensor and support vector machines”. In: *Fire Safety Journal* 44.3 (2009), pp. 322–329.

[122] J. F. Collumeau et al. “Fire scene segmentations for forest fire characterization: A comparative study”. In: *2011 18th IEEE International Conference on Image Processing*. IEEE. 2011.

[123] T. Celik and H. Demirel. “Fire detection in video sequences using a generic color model”. In: *Fire safety journal* 44.2 (2009), pp. 147–158.

[124] S. Rudz et al. “Investigation of a novel image segmentation method dedicated to forest fire applications”. In: *Measurement Science and Technology* 24.7 (2013).

[125] T. Celik. “Fast and efficient method for fire detection using image processing”. In: *ETRI journal* 32.6 (2010), pp. 881–890.

[126] W. B. Horng, J. W. Peng, and C. Y. Chen. “A new image-based real-time flame detection method using color analysis”. In: *Proceedings. 2005 IEEE Networking, Sensing and Control, 2005.* IEEE. 2005.

[127] T. H. Chen, P. H. Wu, and Y. C. Chiou. “An early fire-detection method based on image processing”. In: *2004 International Conference on Image Processing*. 2004.

[128] A. Z. Chitade and S. K. Katiyar. “Colour based image segmentation using k-means clustering”. In: *International Journal of Engineering Science and Technology* 2.10 (2010), pp. 5319–5325.

[129] *Stereo Camera Calibration under Different Resolution*. URL: `https://github.com/balcilar/Calibration-Under_Different-Resolution` (visited on 02/19/2020).

[130] N. Otsu. “A threshold selection method from gray-level histograms”. In: *IEEE transactions on systems, man, and cybernetics* 9.1 (1979), pp. 62–66.

[131] S. Verstockt et al. “Multi-sensor fire detection by fusing visual and non-visual flame features”. In: *International Conference on Image and Signal Processing*. Springer. 2010.

[132] J. R. Martinez-de Dios and A. Ollero. “Wavelet applications to forest-fire monitoring and measurement”. In: *Proceedings of the 5th Biannual World Automation Congress.* IEEE. 2002.

[133] R. C. Bolles and M. A. Fischler. “A RANSAC-Based Approach to Model Fitting and Its Application to Finding Cylinders in Range Data.” In: *7th International Joint Conference on Artificial Intelligence*. IJCAI. 1981.

[134] B. Moretti. “Modélisation du comportement des feux de forêt pour des outils d’aide à la décision”. PhD thesis. 2015.

[135] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel. “On the shape of a set of points in the plane”. In: *IEEE Transactions on information theory* 29.4 (1983), pp. 551–559.

[136] T. J. Cholewo and S. Love. “Gamut boundary determination using alpha-shapes”. In: *Color and Imaging Conference*. Society for Imaging Science and Technology. 1999.

[137] D. M. L. Y. Sommerville. *Introduction to the Geometry of N Dimensions*. Courier Dover Publications, 2020.

[138] G. P. Gerdan and R. E. Deakin. “Transforming cartesian coordinates X, Y, Z to geographical coordinates $\phi$, $\lambda$, h”. In: *Australian surveyor* 44.1 (1999), pp. 55–63.

[139] *Global land cover change - wildfires*. URL: `http://stateoftheworldsplants.org/2017/report/SOTWP_2017_8_global_land_cover_change_wildfires.pdf` (visited on 03/25/2020).

[140] S. H. Doerr and C. Santın. “Global trends in wildfire and its impacts: perceptions versus realities in a changing world”. In: *Philosophical Transactions of the Royal Society B* 371.2016 ().

[141] D. Morvan and J. L. Dupuy. “Modeling the propagation of a wildfire through a Mediterranean shrub using the multiphase formulation”. In: *Combustion and flame* 138.3 (2004), pp. 199–210.

[142] J. F. Sacadura. “Radiative heat transfer in fire safety science”. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 93.1 (2005), pp. 5–24.

[143] J. L. Rossi et al. “Simplified flame models and prediction of the thermal radiation emitted by a flame front in an outdoor fire”. In: *Combustion Science and Technology* 182.10 (2010-09-22), pp. 1457–1477.

[144] M. A. Finney. “FARSITE: Fire area simulator-model development and evaluation”. MA thesis. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 1998.

[145] R. Linn et al. “Studying wildfire behavior using FIRETEC”. In: *International Journal of Wildland Fire* 11 (2002), pp. 233–246.

[146] C. Tymstra et al. “Development and structure of Prometheus: the Canadian wildland fire growth simulation Model”. MA thesis. Edmonton, 2009.

[147] P. A. Bisgambiglia et al. “DIMZAL: a software tool to compute acceptable safety distance”. In: *Open Journal of Forestry* A.7 (2017), pp. 11–33.

[148] R. Siegel and J. R. Howell. In: *Thermal Radiation Heat Transfer*. Hemisphere Publishing Corporation, 1994.

[149] S. Verstockt et al. “A multimodal video analysis approach for car park fire detection”. In: *Fire Safety Journal* 57 (2013), pp. 9–20.

[150] B. Delaunay. “Sur la sphere vide”. In: *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk* 7.793-800 (1934), pp. 1–2.

[151] M. Galetzka and P. O. Glauner. “A simple and correct even-odd algorithm for the point-in-polygon problem for complex polygons”. In: (2012).

# Publications and distinctions

## Publications

Ciullo V. "Detection, suivi et measure de feux de foret par drone". *Tribune des chercheurs de Corse*, 2017.

Ciullo V., Rossi L., Toulouse T., Pieri A. "Fire Geometrical Characteristics Estimation Using a Visible Stereovision System Carried by Unmanned Aerial Vehicle". *15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, 18-21 November 2018 (Singapore)

Ciullo V., Rossi L. "Monitoring the geometric characteristics of a vegetation fire by drone". *Journée GT UAV (Unmanned Aerial Vehicles)*, 21 June 2019 (Laboratoire ISM UMR 7287, Marseille, France)

Ciullo V., Rossi L. "Wildfire measurements with UAV multimodal stereovision" *Remote Sens.*, submitted 12 May 2020.

## Awards

Winner of the first prize at the event "my innovation is..." for the best innovation project in the south-east of France in 2017.

Selected in the 20 best research poster of the University of Corsica at the "Journée des doctorants 2017"

## Patents

DroneVisionForFire: conceptualization, Ciullo V., Rossi L. and Pieri A.; methodology, Ciullo V., Rossi L. and Pieri A.; validation, Ciullo V. and Rossi L.; formal analysis, Ciullo V. and Rossi L.; programming, Ciullo V.; investigation, Ciullo V., Rossi L. and Pieri A.; resources,

Rossi L., writing-original draft preparation, Ciullo V. and Rossi L.

# Acknowledgements

The work presented in this document is the result of five years of research in Corsica, in Corte. I was lucky to be surrounded by people I want to thank them.

First, I warmly thank my supervisor Dr. Lucile Rossi, without whom this thesis would not have seen the day and for having welcomed me in her team. Thank you for her investment, kindness, her support and for her meticulous re-reading of this manuscript.

I thank my Thesis Commission for its encouragement and help with relevant advice during the end-of-year meeting throughout this thesis. I also thank it for his relevant advice while throughout this thesis.

Thanks to Pr. Yvon Voisin and Dr. Simon Lacroix for accepting to be examiners of this PhD thesis. I also thank Dr. Paul-Antoine Bigambiglia and Pr. Paul Bisgambiglia for being part of the jury and for their welcome in the computer science department of the University of Corsica.

I would also like to thank Pr. Paul-Antoine Santoni, director of the Joint Unit 6134 Sciences For the Environment CNRS - University of Corsica for welcoming me to his laboratory.

I thank the people of Corsica who contributed to the smooth running of this thesis in starting with Antoine Pieri, a colleague of the fire project, for his great help during the experiments and for the construction of the device. Thanks to Thierry Carlotti, member of the information systems department, for his technical assistance during the preparation of the room for the good success of the video conferences.

I also thank Cedric Savignoni, Xavier Carriera and David Savignoni of ECALIS association for their help during the experiments in the field of drones and for their good mood; they have produced the first model of the drone that allowed me to start developing my studies the first year. I thank the members of the company Photos-Videos-Reporters for their collaboration and assistance during the execution of experimental tests; in particular the support provided by Xavier Carreira as a drone pilot.

Thanks to the SAT SUD-EST office, for their contribution and help in the development of a patent during these thesis years, in particular to Margot and Bertrand for their help for the preparing of the presentation of my project at the event "My innovation is...", which allowed me to get the first prize.

I thank David Celio of the information systems department for his availability regarding the implementation of computer equipment. Thanks also to the student Elie El Rif that did his End of Study Project of Engineering school with us and worked on the latest version of

the stereo-vision system we developed.

I thanks in particular Toulouse Tom for his welcome from the day one, for his nice company during my first two years of thesis and for all the work that was useful for me to get into the perspective of my thesis, other more for his advice as "doctor".

I also thank all the doctoral students of the University of Corsica (some of whom are now doctors) with whom I sympathized during these years: Andreï, Christelle, Gauthier Lapa, Jean-Baptiste Tramoni, Laetitia, Louis-Felix, Marie, Mathilde, Mohamed, Melissa, Raphael, Romain, Lara, Paul-Henry, Nicolas, Guillame, Karina, Carmen, Margaux, Jean-Baptiste Filippi, and also Wani, with whom I could taste his very precious beer.

I also thank the Unit of Instruction and Intervention of Civil Security No. 5 of Corte for making possible the fire experiments carried out during this thesis and for ensuring our safety during these burns. Thanks also to all our partners.

I also want to thank people from the University of Brindisi, Perugia and Freiburg, where my passion for computer vision and robotics began. Thanks to Siralab Robotics and Skyrobotic for improving my ability to develop that helped me with some aspects of this thesis. I thank Maelle Kernel and Camille Belle for their help during my final writing period. And thank you also to all my friends from Corsica who accommodated me: Romain Haond, Audrey, Brahim, Romain Menegat, Zoé, Emilie, Jean Baptiste, Gauthier, Patrick, Maria Stella, Claire, Geraldine.

I also thank the Cyrnea bar for accommodation in the city of Corte and Henry "the barman" for his sympathy.

Finally, thanks to my father, my mother my brother and my step-sister for always having me supported in what I undertook also if I am always far away.

Thanks to my partner Emilie, for her great help especially in the last period of writing, especially for the chapter of the French version.

Thanks to all for the beautiful moments that I was able to spend in their company during these four years.

This research was funded by the Corsican Region, the French Ministry of Research and the CNRS under Grant CPER 2014-2020