



The
University
Of
Sheffield.

University of Sheffield

**The development and utilization of
low-cost near infrared imagers for
fire detection and diagnostics**

Xiao Wang

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Faculty of Engineering
Department of Mechanical Engineering

Submission Date
Spring 2021

This thesis is dedicated to
my wife, my parents and my son

Acknowledgements

First and foremost, I am deeply grateful to my wife and parents for their constant support and understanding. They give me strength to weather the hardships of life. They bring the most sincere joy to my life.

I am grateful to my supervisor Prof. Yang Zhang for his guidance and advice. The valuable discussion shed lights on the subject matter which helped me proceed with confidence.

I would thank my second supervisor Dr. Jon Willmott for his generosity of sharing lab resources, and his help on the experimental collaborations.

I would like to thank Dr. Thomas C Wilkes for his help on the modification of the camera sensors, and the discussion on the various matter of sensor hardware.

I would like to extend my gratitude to my good friend and colleague Ahmed ARR Albadi for the discussion on just about anything. With similar thoughts, propensity and curiosity, discussions and share of ideas have never been more enjoyable.

I would like to thank my good friend and colleague, Yufeng Lai for the collaboration and share of ideas for experiments.

Finally, I would like to thank other group members for their support and companionship.

Declaration

I, the author, confirm that the Thesis is my own work. I am aware of the Universitys Guidance on the Use of Unfair Means. This work has not been previously presented for an award at this, or any other, university.

Publications

1. Xiao Wang and Yang Zhang. The Potential of utilizing Near Infrared Spectrum for fire detection. *27th ICDERS conference*, page 7, 2019
2. Yufeng Lai, Xiao Wang, Thomas B.O. Rockett, Jon R. Willmott, Hangxu Zhou, and Yang Zhang. The effect of preheating on fire propagation on inclined wood by multi-spectrum and schlieren visualisation. *Fire Safety Journal*, 118:103223, December 2020

Nomenclature

Abbreviations

<i>BCFA</i>	Bayer colour filter array
<i>CCD</i>	Charge-coupled device
<i>CFA</i>	Colour filter array
<i>CMOS</i>	Complementary metal oxide semiconductor
<i>CNN</i>	Convolutional neural network
<i>DPGMM</i>	Gaussian mixture model with Dirichlet process prior
<i>FOV</i>	Field of view
<i>FPGA</i>	Field-programmable gate array
<i>GLCM</i>	Gray level co-occurrence matrix
<i>GMM</i>	Gaussian mixture model
<i>HgCdTe</i>	Mercury cadmium telluride
<i>HOG</i>	Histogram of oriented gradient
<i>InGaAs</i>	Indium gallium arsenide

<i>InSb</i>	Indium antimonide
<i>IR</i>	Infrared
<i>ISO</i>	Camera sensitivity
<i>LMSE</i>	Least mean square error
<i>LWIR</i>	Long wavelength infrared
<i>MOSFET</i>	metaloxidesemiconductor field-effect transistor
<i>MWIR</i>	Mid-wavelength infrared
<i>NIR</i>	Near infrared
<i>PSD</i>	Power spectral density
<i>ROI</i>	Region of interest
<i>ROIC</i>	Readout integrated circuits
<i>SVM</i>	Support vector machine
<i>SWIR</i>	Short-wavelength infrared
<i>UAV</i>	Unmanned aerial vehicle
<i>UV</i>	Ultraviolet
Greek letters	
α	learning rate
γ	non-linearity constant in radial basis function
$\mathcal{N}(x_i \mu_k, \Sigma_k)$	Multivariate Gaussian distribution

μ	mean
μ_k	mean of k^{th} component
$\sigma(\cdot)$	Activation function
σ^2	Variance
Σ_k	Covariance matrix of k^{th} component
$\varphi(\cdot)$	Nonlinear transformation function
ζ_i	Slack variable
w_k	Weight of Gaussian component

Common symbols

ΔI	Change in intensity
\hat{Y}_i	Predicted class label
μm	Micrometer
μs	Microsecond
b	bias term
C	Regularization constant
d	Margin distance
f_2	Focal length of the second mirror
fps	Frame per second
H_0	Optimal decision hyperplane

H_1	Marginal hyperplane 1
H_2	Marginal hyperplane 2
I	Spectral intensity
J	Objective function
K	Gladstone-Dale constant
$Loss_{CE}$	Cross-entropy loss
ms	Millisecond
$x_{i,j}$	Feature vector
Y_i	Class label
\mathbf{w}	Weight vector

Abstract

This research explores the use of near infrared spectrum for video fire detection and combustion diagnostics. The near infrared spectrum is somehow very under-explored in these fields. A low-cost silicon-based backside illuminated CMOS image sensor was modified to become a monochrome sensor exposing its full spectral sensitivity. The sensor was tested to show its improved spectral sensitivity. The multi-spectrum fire detection combines stereo cameras with NIR only and NIR-visible spectrum respectively for robust fire detection. Flame image properties in both conditions are extensively studied, whereby the NIR-only channel corresponds to ROI extraction and texture feature extraction; the NIR-visible channel give rise to a unique colour model for false positive classification rejection. Machine learning algorithms were employed for fire recognition. Practical considerations of designing a fire detection system have been discussed in terms of sensor selection, feature extraction as well as choice of classification algorithms. Dual-band stereo video fire detection not only showed great potential for robust fire detection, but also for vision-based automated firefighting. In combustion diagnostics, low-cost NIR sensors were used for the imaging of combustion products of fuel-lean premixed flames, which demonstrated its effectiveness in potential flame instability related diagnostics. Moreover, the thermal

capability of NIR camera sensor were applied in conjunction with visible and schlieren imaging, to study the mechanism of fire propagation on solid fuel combustion. Finally, a summary was made with additional suggestions and speculations on the subject as a whole.

Contents

1	Introduction	1
1.1	Background	1
1.2	Aim and objectives	5
1.3	Thesis outline	6
2	Literature Review	7
2.1	Traditional fire detection	7
2.1.1	Smoke detectors	7
2.1.2	Optical gas detectors	9
2.2	Fire detection based on radiation emission	13
2.2.1	Single or multiple photosensitive diodes	14
2.2.2	Focal plane array: video fire detection	15
2.3	Video fire detection	17
2.3.1	ROI extraction and feature extraction	17
2.3.2	Colour detection	18
2.3.3	Moving region detection and dynamic features	22
2.3.4	Decision-making: machine learning methods	25
2.3.4.1	Support vector machine	26
2.3.4.2	Neural networks in general	29
2.3.4.3	Convolutional neural network (CNN)	34

2.4	Video fire detection in IR spectrum	39
2.4.1	Long Wavelength IR fire detection	40
2.4.2	Choice of IR detectors for fire detection	42
2.5	Sensor hardware: focal plane array	46
2.5.1	CMOS vs CCD	46
2.5.2	Infrared detectors	50
2.5.2.1	Photon detectors	50
2.5.2.2	Thermal detectors	51
2.6	Fire detection in different scenarios	54
2.7	Passive NIR instrumentation in literature	56
2.8	Literature review summary	60
3	Raspberry Pi camera system for NIR fire detection	63
3.1	Background	63
3.1.1	The near infrared spectrum: active imaging	63
3.1.2	The near infrared spectrum: passive imaging	65
3.2	Back-illuminated CMOS sensor: Raspberry Pi cameras	68
3.3	Turning Pi camera into a monochrome sensor	71
3.4	Advantages of NIR for fire detection	74
3.4.1	NIR simplifies ROI extraction	80
3.4.2	Using NIR false colour for fire detection	84
3.5	Summary	84
4	Multi-spectral fire detection system	86
4.1	ROI extraction stage	88
4.2	Feature extraction stage	91

4.3	Decision-making stage	99
4.3.1	Classification using SVM based on extracted features	99
4.3.2	GMM colour model to eliminate false positives	104
4.3.2.1	The cause of false colour in NIR image	108
4.3.2.2	Colour distribution modelling	112
4.3.3	CNN for classification	116
4.4	Summary	119
5	NIR combustion diagnostics	121
5.1	Introduction	121
5.2	NIR spectrum of a flame	124
5.3	Passive NIR imaging of fuel-lean hydrocarbon flames	126
5.3.1	Fuel-lean premixed flames imaging in 300-1100 nm	126
5.4	Effect of Bayer CFA on hydrocarbon flame colour	131
5.5	Premixed flame with NIR imaging	139
5.5.1	Premixed methane and hydrogen flame pulsation with NIR imaging	140
5.5.2	Premixed methane and hydrogen flame evolution with NIR imaging	145
5.6	Usage of NIR imaging in solid fuel combustion	151
5.6.1	Background	151
5.6.2	Fire propagation on wood with combined imaging system . . .	156
5.6.2.1	Some observations and discussion	160
5.6.3	Fire propagation on wood with combined imaging system with fan	165
5.7	Summary	171

6	Conclusion and future work	173
6.1	Conclusion	173
6.2	Future work	179
A	Fire detection algorithm	182
	Bibliography	183

List of Figures

1.1	Signal sources from a fire	3
1.2	Different fire detection scenarios	5
2.1	Ionization smoke detector	8
2.2	Gas emission and absorption spectrum bands	9
2.3	Optical gas detectors	10
2.4	Detector layout	13
2.5	Hydrocarbon flame spectral emission	14
2.6	Video fire detection	15
2.7	Colour model	19
2.8	Flame flicker frequency spectrum	24
2.9	Image energy change in the presence of smoke	24
2.10	Illustration of kernel transformation	26
2.11	Illustration of optimal hyperplane	28
2.13	Illustration of the structure of a perceptron	30
2.12	Illustration of feedforward network	30
2.14	Nonlinear activation functions	31
2.15	Illustration of logistic regression	32
2.16	Illustration of multi-class logistic regression	33
2.17	Illustration of backpropagation	34

2.18	Illustration of LeNet as a CNN structure	36
2.19	Long Wavelength IR images	39
2.20	Long Wavelength IR images	40
2.21	Blackbody radiation plot at low temperatures	43
2.22	Atmospheric transmission	44
2.23	Illustration of charge integration and shifting in CCD	47
2.24	Magnified cross-sectional view of CMOS image sensor pixel	48
2.25	Photon absorption depth and energy in silicon	49
2.26	General spectral response: photon detectors vs thermal detectors	50
2.27	Sensing mechanisms: photon detectors vs thermal detectors	50
2.28	IR detector materials and operating temperature	52
2.29	Illustration of microbolometer pixel structure	53
2.30	Illustration of wireless sensor network (WSN)	55
2.31	Radiance of CO_2 and H_2O at various temperature	58
2.32	Visualization of radiation heat loss and local extinction in IR	60
3.1	Spectral response of Si-CMOS and InGaAs cameras	65
3.2	NIR image and Visible image comparison	66
3.3	Illustration for Mie and Rayleigh scattering	66
3.4	Comparison of visible and NIR images	67
3.6	How Bayer colour filter array works	69
3.5	Pi cameras V1 and V2	69
3.7	Two versions of Pi cameras image comparison: visible only vs visible + NIR	70
3.8	Removal of BCFA and bespoke lens	71

3.9	DN (digital number) counts comparison of original and de-Bayered V1 Pi Camera	72
3.10	de-Bayered Pi camera linearity tests	73
3.11	De-Bayered Pi camera spectral response.	74
3.12	Solar spectral irradiance visible and NIR	75
3.13	Stereo NIR and visible image comparisons: smouldering charcoals and a flame	76
3.14	NIR image scene test with 950 nm cut-on filter	78
3.15	Integrated atmosphere IR spectral transmission	78
3.16	Thresholding for ROI extraction	79
3.17	The failure of Otsu’s method of thresholding in scene test	81
3.18	Modelled spectral irradiance from the sun by SMARTS radiant transfer model	83
3.19	NIR false colour vs common visible colour of a fire	84
4.1	Multi-spectral fire detection system	87
4.2	Morphological operation to combine blobs into patch	88
4.3	Dilation with different filter size	89
4.4	Second stage adaptive crop	90
4.5	Computation of co-occurrence frequency matrices	92
4.6	Illustration of texture difference between fire and non-fire objects . . .	93
4.7	GLCM texture homogeneity vs sample number (fire and non-fire objects) at different orientations	94
4.8	GLCM texture homogeneity vs sample number (fire and non-fire objects) with different pixel pair distances	95
4.9	Entropy image of fire and non-fire objects	97

4.10	Variance plot of fire and non-fire objects	98
4.11	SVM visualization in feature space	100
4.12	Effect of C on SVM decision plane	101
4.13	Effect of γ in RBF on SVM decision plane	102
4.14	Misclassified examples in training	103
4.15	Training confusion matrix	103
4.16	Test case: 5m	105
4.17	Test case: 10m	105
4.18	Test case: 15m	106
4.19	Test case: 20m	106
4.20	Substance concentration in lateral position of a diffusion flame	108
4.21	Typical spectral response of a RGB camera	109
4.22	Diffusion methane flame colour space visualization	110
4.23	Colour histogram of diffusion flame captured with 760nm cut-on filter	111
4.24	Bayer CFA spectral response up to 1000 nm	111
4.25	Colour distribution: NIR+Visible vs NIR cut-on at 760 nm	112
4.26	Colour distribution: NIR cut-on at 950 nm at two exposures	113
4.27	Visualization of Dirichlet distribution	114
4.28	Visualization of fitting GMM with EM algorithm	115
4.29	Visualization of fitting GMM with Dirichlet process prior	116
4.30	Illustration of CNN model structure	117
4.31	CNN and SVM detection rates vs distance to camera	118
4.32	Training graphs of CNN	119
5.1	Emission relative intensity of hydrocarbon flames	122
5.2	Line absorption strength of H_2O , CO_2 and CO at 1500K	122

5.3	Lean premixed methane flame in the complete sensitive spectrum of Si-based CMOS sensor	128
5.4	Narrow band imaging with lean-premixed methane flames	129
5.5	Short band NIR imaging with lean premixed hydrocarbon flames . . .	130
5.6	Diffusion methane flame imaging VIS vs VIS+NIR and hue channel distribution	132
5.7	Diffusion methane flame blue and red channel histograms: VIS vs VIS+NIR	133
5.8	Premixed methane flame imaging in fuel-lean condition: VIS vs VIS+NIR	134
5.9	Premixed methane flame colour shift w.r.t equivalence ratio	135
5.10	Premixed methane flame colour distribution w.r.t equivalence ratio .	136
5.11	Camera Bayer CFA spectral response in visible spectrum	137
5.12	Premixed methane flame R,G,B histograms at equivalence ratio 0.8 .	138
5.13	Premixed NIR flame experiment layout	140
5.14	Illustration for pixel value extraction lines	141
5.15	PSD of NIR intensity variation in premixed methane flame	142
5.16	PSD of NIR premixed methane flame pulsation	143
5.17	PSD of NIR intensity variation premixed hydrogen flame	144
5.18	PSD of NIR premixed hydrogen flame pulsation	145
5.19	Colour contour of premixed methane flames in fuel-lean condition . .	147
5.20	Colour contour of premixed methane flames in fuel-lean condition . .	148
5.21	Colour contour of premixed hydrogen flames in fuel-lean condition . .	149
5.22	Time averaged methane flames in fuel-lean condition	150
5.23	Illustration for biomass and wood constituents	151
5.24	Molecular structure of cellulose polymer	152

5.25	The loop of self-sustained wood combustion	154
5.26	Radiation energy versus flame size for pool fire	155
5.27	Schematic diagram of smouldering in wood	156
5.28	Schematic diagram of the experiment setup	157
5.29	Layout of combined imaging system	158
5.30	Combined imaging with NIR, visble and schlieren for $\theta = 30^\circ$	161
5.31	Visualization of different regimes of wood combustion	162
5.32	Illustration for the dominance of convection or radiation	163
5.33	Schematic diagram of the experiment setup with fan	165
5.34	Combined imaging at different fan speeds	167
5.35	Illustration of deflected light rays caused by density change in schlieren system	169
5.36	Illustration of patch intensity change from schlieren image	170
5.37	Histogram of pixel intensities in the extracted patch underneath the wood sample	171

Chapter 1

Introduction

1.1 Background

Fire outbreak poses great threat to human lives and assets. Great efforts have been made to improve fire safety, in terms of new regulations for different constructions enforcing new fire resistant materials to be used in building or compliance fire safety design, etc. However, in the fire alarming domain, the status quo seems to have not been improved much over the years. The most common fire alarms uses may contain one or more point detectors, such as smoke detector, heat detector, *CO* detector, which are usually effective in the small spaces. In a large indoor or outdoor open space, their effectiveness dramatically decreases with respect to distance. In addition, limited by the working principles of these systems, false alarm rate is high due to many environmental disturbances. The reduction of false alarms of the traditional fire alarm systems is difficult given the complex nature of aerosol substance properties. Even the detection is successful, it provides very limited information of the fire in terms of location, size, spread rate, etc. Oftentimes, manual check is entailed to eliminate potential fire event. One crucial fact is that the time taken for the manual check would be quite enough for the fire to develop, especially in large indoor space with complex structure. Moreover, the manual check may be subjected to human errors.

For example, the tragedy of Notre-Dame de Paris fire in 2019, was caused by delayed human check because of no available information other than the alarm was set off [1]. The guard was sent to check but went to the wrong place, the attic of the adjoining sacristy. By the time of the fire was discovered, fifteen minutes had passed leading to a well-developed fire.

The lack of information is one of the vital drawbacks in conventional fire alarm systems, not to mention the delayed response in large space. Video based fire detection offers a promising addition, or may be an alternative choice for fire safety. Notwithstanding the tremendous amount of surveillance cameras deployed in urban areas, it is not an easy job to keep an eye for the event of a fire. The development of camera sensors, hardware computability and machine vision algorithms promotes new openings for fire detection. Video fire detection offers no time delay in receiving signals comparing to conventional detectors. Real-time detection is achievable. In addition, vision based technique can provide extra valuable information to operators before firefighting. Furthermore, vision based methods are potential candidates for future automated robots for fire detection, localization and suppression.

The main challenge in video fire detection is to reduce false alarm rate and timely fire recognition at the earliest stage of a fire breakout. Simple as the aim states, however, the implementation is rather challenging. Not only is fire a complex phenomenon rendering its behaviour not easily predictable, but the dependence on surrounding environment poses a great barrier for the robustness of a fire detection system. Typical fuels in an event of fire contain carbon (C), hydrogen (H), oxygen (O), fluorine (F), nitrogen (N), chlorine (Cl) and bromine (B), which usually are subjected to incomplete combustion [2]. The combustion process generates a range of products, such as CO_2 , CO , H_2 , hydrogen cyanide (HCN), hydrogen

chloride (HCl), hydrogen fluoride (HF) and hydrogen bromide (HBr) and carbon (*soot*). Moreover, the radiation from a fire has overlapping spectrums from different combustion products or intermediates, but usually is dominated by soot, which gives a wide spectrum spanning from UV to far infrared.

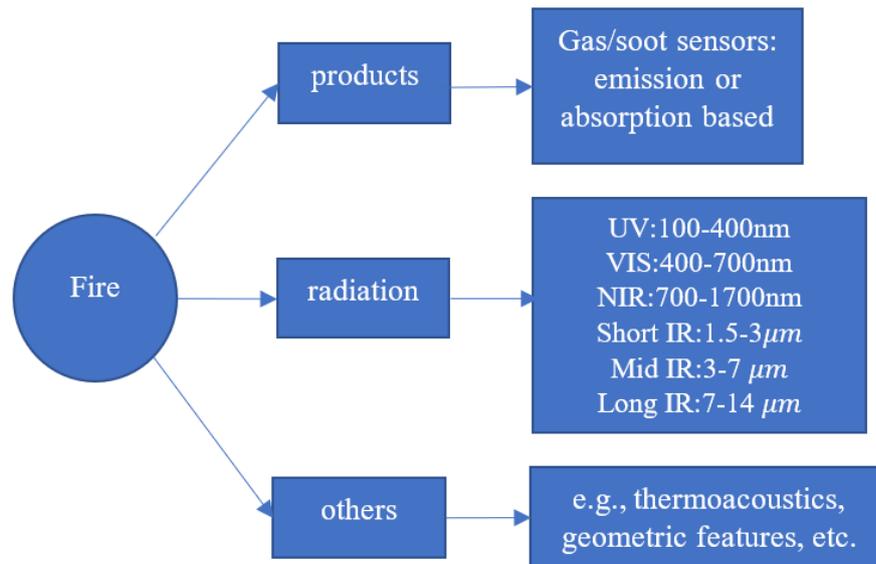


Figure 1.1: Signal sources from a fire

With respect to these fundamental physical properties of a fire, many fire detection systems try to tackle the problem by utilizing the fire signals, as shown in Figure 1.1. The detection strategy varies greatly according to the system design. However, in practice, it is usually a compromise in system integration. For the most obvious reasons, the added sensors and processing units will raise the cost significantly. It may not be an issue for some bespoke high-end situation, but the cost-effectiveness is the driven factor for both household and mass use for public safety. Secondly, more data throughput and processing not only increase the computation load but makes the decision making of the system more complicated, therefore robustness may be compromised. Finally, the dependence on scenarios will require different

bias on extracting appropriate information from the fire, as shown in Figure 1.2. For example, the fire detection in shipment compartment is mostly enclosed and free from outside light source, where a simple thresholding with IR camera would achieve great performance [3]. By contrast, in an outdoor shopping area, many interference exist to degrade the performance of a system, for example car lights, advertisement display, reflections of sunlight from glasses and floors, etc. In the more complex scenario, the detection will involve lots of computation.

In addition to fire detection, the understanding of fire propagation is also instructive on structure design and fire retardation, which is subjected to many factors, such as material, geometric layout, ambient air flow condition, etc. Visible RGB cameras are the dominant choice in fire detection literature that its robustness is questionable. There has been very limited study on video fire detection with IR cameras, all of which used low resolution, high cost microbolometer long wavelength IR cameras. The sensitivity of common silicon based CMOS image sensors in the near infrared spectrum somehow has always been overlooked, which offers the capability of near infrared imaging at a very low cost. In combination with cheap on-board computatbility, multi-spectrum fire detection, flame diagnostics and fire propagation will be investigated.

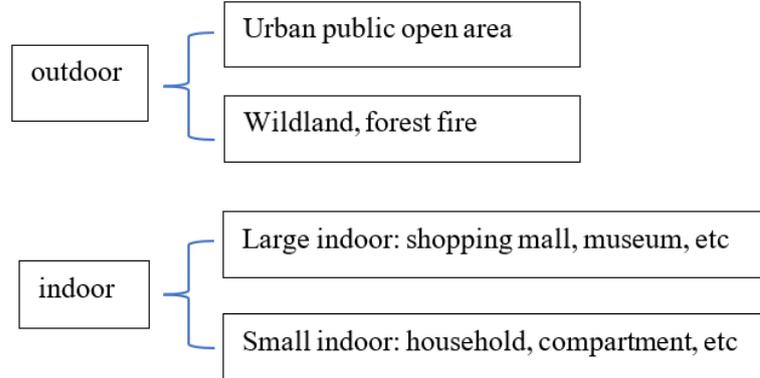


Figure 1.2: Different fire detection scenarios

1.2 Aim and objectives

To promote fire safety in the domains where traditional fire detection technology failed, the visual-based video fire detection is an effective method. However, most existing video fire detection researches rely solely on the visible spectrum. The exploration beyond the visible spectrum is lacking, which can expand the horizon of the fire detection technology.

The aim is to explore the potential of the near infrared capability of low-cost CMOS image sensors in regards to fire detection and combustion diagnostics. Both fire detection and passive imaging combustion diagnostics, in near infrared spectrum, are very limited in literature. Gaining understanding of flames at this spectrum as well as its fit into video fire detection will promote its application with great cost-effectiveness. To that end, the objectives are listed below:

- To study flame properties in visible and near infrared spectrum for diffusion, premixed hydrocarbon flames and wood fires.
- To develop multi-stage machine learning based fire detection algorithm by

exploring the fire image property using low-cost CMOS image sensors.

- To compare the difference between popular machine learning algorithms in the field of video fire detection in terms of practical application and considerations.
- To modify image sensor for low-cost flame chemiluminesce and combustion products instrumentation.
- To apply near infrared imaging for solid fuel combustion and fire propagation.
- To study fire propagation on inclined wood by means of multi-spectrum in combination with schlieren imaging.

1.3 Thesis outline

The remaining part of the thesis follows this structure. chapter 2: a comprehensive review of fire detection which including point based methods, video based methods, camera hardware having sensitivity in various spectrum, sensor technology and machine learning techniques in regards to video fire detection; chapter 3: the property of NIR camera sensor whereby the different cameras sensors are tested and studied on fire detection as well as NIR imaging intuitions; chapter 4: expands upon the proposed multi-spectral fire detection system, where different stages of the processing pipeline is are presented and discussed; chapter 5: the extended use of NIR camera for gas and solid fuel combustion as a very low cost alternative to other IR cameras; chapter 6: thorough conclusion and discussion upon the whole of designing fire detection system, where future work is also suggested in several aspects.

Chapter 2

Literature Review

To define the scope of this research, previous researches have been reviewed in two broad areas, namely fire detection and passive imaging-based combustion diagnostics. For fire detection, the review covers traditional point-based detection, radiation detection, and existing video fire detection; the discussion will be based on hardware and algorithms in the pipeline. In particular, a thorough discussion on the use of infrared image sensor for fire detection will be provided, covering the aspects of hardware properties and corresponding algorithms. In addition, the application of machine learning algorithms in the field of fire detection will also be discussed. In respect to passive imaging-based combustion diagnostics, the application of near infrared spectrum in combustion will be discussed.

2.1 Traditional fire detection

2.1.1 Smoke detectors

The most common sensor used in fire alarm is the smoke detection. Because in many fire breaks, smoke arises first from smouldering before flaming. The basic structure of the ionization chamber smoke detector is shown in Figure 2.1. The detector contains a radioactive α particle source, which releases alpha particles in between the two electrodes to ionize the air. With the presence of smoke particles in

the chamber formed by the electrodes, the current is reduced, whereby alarm is rang based on the change of current. Later, new variation of such smoke detection without the use of radioactive material was proposed, which detects the positive charge of smoke particles from flame [3]. Fundamentally, these detectors detect the change in electromagnetic field in the chamber by charged particles, which is directly affected by its concentration and size. Problems arise because smoke particles vary greatly from different combustion in terms of structure and composition [4]. They can be insensitive to smouldering as the particle will lose the charge after cooling and can be sensitive to dust, steam or many other particles getting into the chamber.

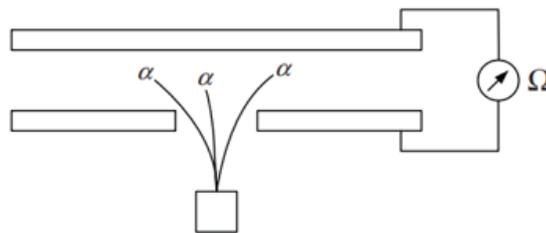


Figure 2.1: Schematic diagram of an ionization smoke detector

Optical smoke detection is an improved option with better sensitivity and robustness against ambient interference. Basically, the sensing relies on optical obscuration and scattering effects of particles. The suspended particles will scatter and attenuate the sensing light beam, which would be picked up by the system. The scattering of light is affected by many factors, for instance particle size, shape, refractive index, wavelength of the light source and angle of scattering [5]. These limited the detectable range of the particle size, where smaller particles from some flaming situation will not trigger the alarm.

The reliance of smoke detection solely will introduce many false alarms. Despite of the obvious drawbacks, it remains the most popular option in household and

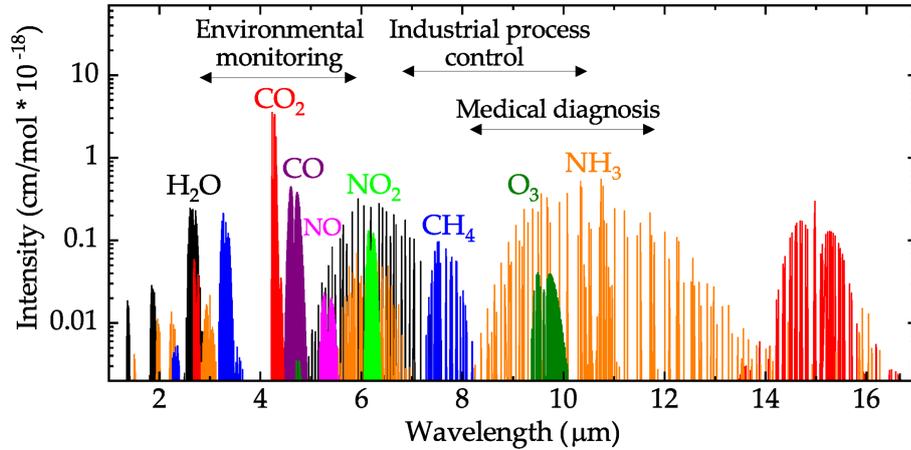


Figure 2.2: The emission and absorption spectral bands of common combustion gases [9]

buildings because of its low cost. It can be very effective in detecting smokes in small compartments, however, in large space its use is limited because of the transport time delay. The fire can develop significantly before the smoke particles accumulate and transported to the detector to trigger the alarm.

2.1.2 Optical gas detectors

Similar to smoke detectors that utilized the substances generated from fire, some fire detection systems are based on the detection of gases. Common gases such as CO_2 , H_2O , CO , H_2 and O_2 , from which the corresponding concentration is measured by gas sensors. CO was reported to be the most promising choice as early fires often started from smouldering from which the lack of oxygen give rise to increased carbon monoxide production [6]. More dedicated detectors were available for detecting HCl, which is usually generated from the pyrolysis of PVC materials [7]. Modern building construction extensively use PVC, for example the cable insulator in household circuits and appliances. The mid IR absorption spectrum is particularly useful in detecting the combustion products, which usually covers 2.5 to 25 μm [8].

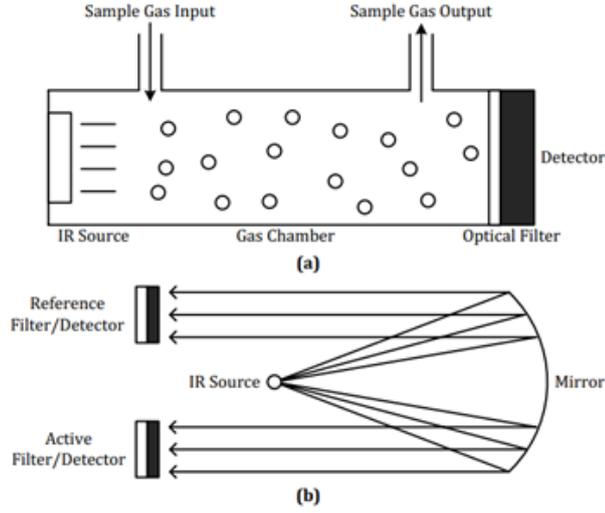


Figure 2.3: Direct and referenced gas sensing configurations [10]

Figure 2.2 shows the emission and absorption bands of some gases in the mid IR range. These spectrum signatures can be used to for gas detection. The prevailing type of optical gas sensors works based on absorption spectroscopy. The basic construction of the optical gas detector is schematized in Figure 2.3. The chamber includes an IR light source, which give a normal input to the detector. In the presence of a detectable gas, depending on its IR absorption, the attenuation is to be measured by the detector. Figure 2.3 (a) and (b) give the two configurations for direct sensing and referenced sensing. In the configuration of Figure 2.3 (a), the IR light is filtered at the desirable absorption bands for the gas of interest, whereby it is usually called non-dispersive IR sensors (NDIR) [11]. The measurement of the attenuation is based on Beer-Lamberts Law as in 2.1: $I(\lambda)$ where I_0 and I are the light source intensity and detected intensity (W/m^2), respectively; $\alpha(\lambda)$ is the gas absorption coefficient (L/gm); c is the gas concentration (g/L) and l is the optical path length (m).

$$\log(I_0/I) = c l \alpha(\lambda) \quad (2.1)$$

The system design can take various forms with different choices of light source and detector sensor types. Some implementations have been listed in Table 2.1. It should be noted that these choices can vary the cost significantly, therefore affecting their popularity in practical uses. The cost of the detectors may be reduce by CMOS fabrication process through mass production driven by the high demand. However, the whole system integration including downsized optical components, dedicated light source and detection mechanisms and other peripherals would deter the cost reduction. In addition, the mid IR sensor technology still faces the most prominent challenge towards miniaturization, low-cost, and low power consumption, despite of tremendous efforts previously [12]. Especially for fire detection based on wireless sensor network, which is made more practical than ever during the era of IoT, low cost and miniature is essential for mass distributed use.

Table 2.1: A collection of gas detectors

Emitter	Detector type	Wavelength (μm)	Detected gas	Reference
LED	photo-diode	4.26	CO_2	[13]
MEMS heater	Bolometer	4.26	CO_2	[14]
MEMS heater	Pyroelectric	4.65/4.26/3.31	$CO/CO_2/CH_4$	[15]
MEMS heater	Pyroelectric	4.26	CO_2	[16]/[17]/[18]
Lamp	Pyroelectric	4.66/4.26/3.33	$CO/CO_2/CH_4$	[19]
QCL	HgCdTe (photoconductive)	4.23	CO_2	[20]
HeNe	PbSe (Photoconductive)	3.4	CH_4	[21]
LED	photo-diode (InGaAs)	1.66	CH_4	[22]
Laser	PbTe (Photoconductive)	3.31	CH_4	[23]
DFB	photo-diode (InGaAs)	1.65	CH_4	[24]

In summary, gas detectors are passive as the gases have to be transported to

the detector, which will always give rise to a time delay. These may not be an issue for buildings with small compartments but will likely be less effective for large space indoor or outdoor. According to the development of the state-of-the-art optical gas sensors, the initial installation cost and the maintenance are still less economical. Research efforts have been made towards the direction of low cost and high reliability.

2.2 Fire detection based on radiation emission

The more conventional fire detection relies on detecting the substances involved during the combustion process, for example smokes, various gases as mentioned in the last section. Another approach is radiation-based, whereby the flame, as a radiation source, the selected radiation spectrum is sensed and then processed. The most common types of radiation detector are UV and IR detectors. This approach offers fast response and detection effectiveness in large open space.

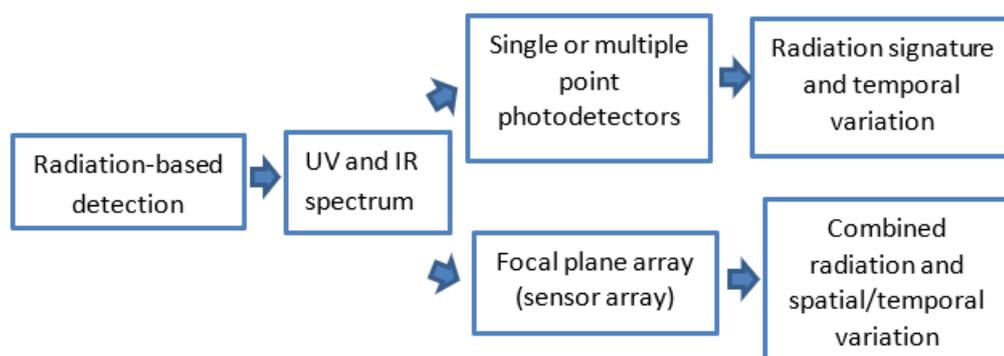


Figure 2.4: Detectors layouts for fire detection

As summarized in Figure 2.4, the traditional detection usually uses single or multiple photon detectors to gather fire signature radiations. Then based on simple logic, the alarm is triggered. Despite its improved sensitivity and reliability, it is still subjected to false alarms in certain scenarios. Over the years, the advancement of sensor fabrication technology, especially the success of monolithic CMOS, has promoted the use of focal plane arrays for fire detection. This enabled a wide range of options to exploit flame characteristics, including spectral properties, temporal and spatial properties, shape, size, etc. Combining with the blooming in hardware computational ability and machine learning algorithms, more feature fusion and abstraction can therefore be realized. These new emerging technologies have the

potential to bring robust fire detection to more practical scenarios.

2.2.1 Single or multiple photosensitive diodes

The combination of UV and IR detectors existed in literature a few decades ago but usually with single or multiple photosensitive diodes. It was an improvement in eliminating false alarm rates over the UV detector only system. The working principle was the mutual exclusiveness: for non-fire sources emitting or reflecting UV does not produce IR radiation; while for non-fire hot objects emit IR radiation but do not give off UV. However, the problem was associated with the significant UV absorption by aerosols such as smoke particles, water molecules, and other contaminants [25].

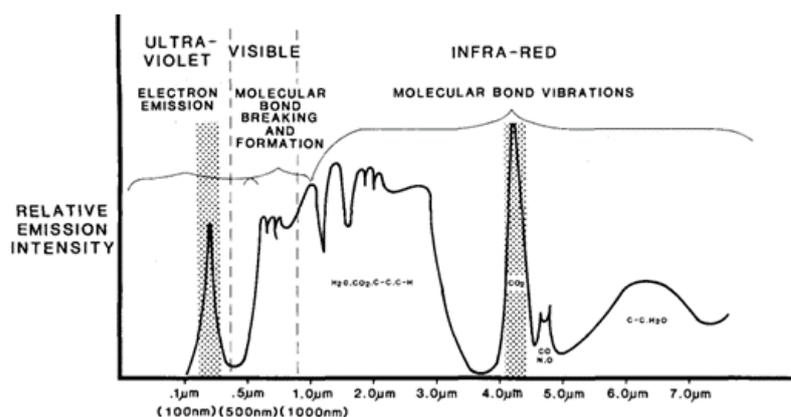


Figure 2.5: The complete spectral emission of hydrocarbon flames

The problems were circumvented by using dual IR detectors, monitoring the CO_2 peak from hydrocarbon flames at around $4.3 \mu\text{m}$ and the background at the valley near $3.8 \mu\text{m}$, as shown in Figure 2.5. At longer wavelength, those aerosol molecules vibrational state does not overlap with the wavelength frequency, therefore no absorption. The ratio of intensity from these two bands will be compared to infer the presence of a fire. The main disadvantage, however, is the drastic attenuation over distance, which give a ratio of unity at long distance. Other option also exists,

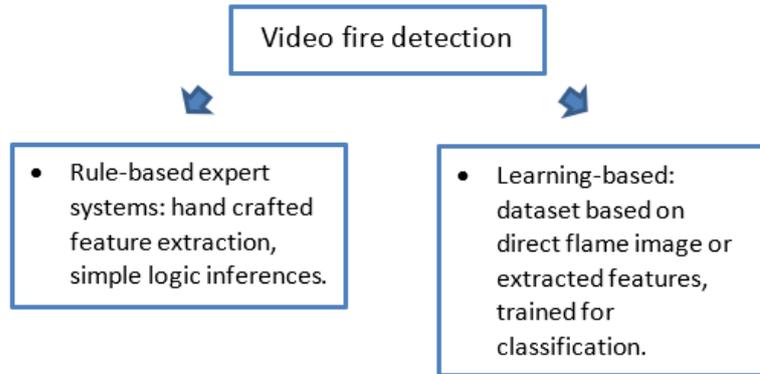


Figure 2.6: Video fire detection road map

such as triple IR detectors, which works in three selected narrow bands with high sensitivity, capable of detecting a range of gases. The inference to set off alarm is based on digital signal processing [26]. It has been reported to be very robust in the industrial environments, for instance offshore oil and gas, automotive industry, aircraft hangars, etc. Its high cost and bulky design prevented its domestic or public use.

2.2.2 Focal plane array: video fire detection

The focal plane array is just a technical nomenclature for camera image sensor, with the indication that optics are used to focus the light on to a 2-D plane where the image sensor is placed. Instead of just measuring the signal from a fire by single or multiple units of photodiode, the camera sensors offer 2D line-of-sight integration of the fire, from which a tremendous amount of data could possibly be extracted for fire detection. It is a natural choice to resort to pattern recognition and machine learning to deal with the huge amount of generated data. The video fire detection can be coarsely divided into two groups: the rule-based expert systems, where features are mostly engineered manually, based on which the decision is made, referring to

Figure 2.6. Generally, features are extracted by means of image processing in the aspects of image histogram, temporal behaviour and colour-based detection, etc. The learning-based detection algorithms are to train a mathematical model to find the presence of a fire in an image. Specifically, the model is just a function but can be complicated to map the inputs (direct image or extracted features from image) to the class label (fire or non-fire). In some learning-based algorithms, features are extracted automatically from the trained function, for instance the recent convolutional neural networks (CNN). Lines cannot be drawn distinctively for these two subdivisions as modern fire detection systems are usually a mix-and-match to form a data processing pipeline.

2.3 Video fire detection

Due to the development in sensor hardware, embedded system and machine learning algorithms, it is likely that video fire detection will substitute many other conventional fire detection systems with significantly improved robustness and reliability at very low cost. Previously mentioned flame detection can be very effective in specific environment and some are mostly applied in a more static environment such as the industry domain. Those deployments are more or less dedicated integration to the plant layout. The reliability of those systems will likely to be reduced dramatically for the use in common public area where many unexpected inferences exist. By contrast, video fire detection provides the possibility of estimate and predict the scene base on complex artificial intelligence algorithms. It should be reiterated that video fire detection usually involves different stages of data processing, where each stage can perform extraction, filtering or inferences. The final decision may be made according to a system of criterion. Intelligent is an overstatement because whether or not to use complicated and capable machine learning algorithms, it all boils down to the feature selection and fine-tuning of the model to achieve automated recognition. Generally, video detection algorithms can be categorized by the size of intake data, namely pixel-wise, blob-wise and patch-wise. The pixel-wise algorithms compute features or label individual pixels according to some measures. The following sections expands on the main components of VFD systems, although it is subjected to changes in specific cases.

2.3.1 ROI extraction and feature extraction

Modern cameras usually have large number of pixels, which generate a tremendous amount of information per image, and this number will only increase as sensor

technology evolves. However, in many detection cases, most of the information contained in an image is not relevant to the object of interest. More information can be a burden, especially for real-time applications. The throttle is at the processing time of the algorithms with the input data. Therefore, a task to reduce the input data while the relevant features are not compromised, becomes an essential stage for pre-processing.

To this end, the first stage is usually to extract the potential fire patches within the image, which is called region of interest extraction (ROI extraction). Common strategies are colour-based extraction, moving object detection, flickering for time series analysis, dynamic behaviours, texture pattern and other hand-crafted descriptors, etc. A comprehensive review of the pre-processing for feature extraction is as follows, whereby the advantages and disadvantages will be discussed.

2.3.2 Colour detection

Pixel-wise approach is among the most popular choices of the rule-based colour isolation, as presented in [27], [28], [29], in which fire-like colour pixels were selected by combined empirical relations in RGB and HSV colour space. The orange to red colour of flame is natively represented by RGB colour space as the RGB colour filters fitted to the camera can replicate any colour. However, more often than not, HSV is used in addition to RGB colour space due to the fact that colour can be isolated from illumination. By contrast, in RGB space, change in any channel values will alter both the colour and intensity. The outcome of the colour algorithm from those research is a subset of the image as potential fire candidates, which will be fed to the subsequent stages of an algorithm.

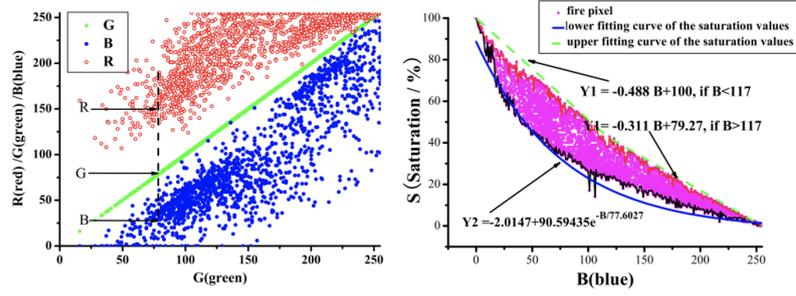


Figure 2.7: Rule-based colour model and curve fitting [30]

Moreover, [30] empirically determined three rules to select fire pixels: $R > R_T$, $R \geq G > B$, $Y1 \geq S \geq Y2$. S is the saturation in HSV space, as demonstrated in Figure 2.7. Through experimentation, the colour value correlations are formulated taking consideration of the illumination change of the scene. Similarly, [31] uses the colour representation in RGB with saturation and intensity to form a colour matrix of fire.

Other empirical rule-based filtering in colour space other than RGB and HSV exists in literature, namely YUV [32][33], YCbCr, and Lab [34] colour space. For example, [35] uses experimentally determined thresholds and rules to select fire pixels, which bears the relationship as: $Y > T_Y$, $|U - 128| < T_U$ and $|V - 128| < T_V$, where Y is the luminance component; U and V are the chrominance components.

$$R > R_T$$

$$R > G > B \quad (2.2)$$

$$S > (S_T(255 - R))/R_T$$

The research work established in [36] combines SURF feature extraction with rule-based colour model in 2.2 to generate features for fire detection. The first stage was to select potential fire pixels based on the colour rules, after which wavelet based local texture information were extracted using SURF. The feature vectors will then

be fed into a SVM for classification.

Probabilistic colour modelling is another approach to detect fire colours, where a set of training examples are used to extract the distribution in the corresponding colour space. For example, the research in [37] used colour model combined with temporal variation for each pixel to locate potential fire regions. Based on some fire training images, the colour histograms of fire pixels in RGB channels were fitted by Gaussian distribution with σ set to 2 in Eq(2.3). Then an empirical threshold was used to the Gaussian smoothed histogram to generate a probabilistic colour check when a new pixel is seen. Over a short period, an average of probability was generated for all the pixels. After that the pixel is labelled as fire if this probability exceeds a threshold.

$$hist(x)_{R/G/B} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_{R/G/B}-\mu}{\sigma}\right)^2} \quad (2.3)$$

The work from [38] also exploit the pixel-wise colour information as the first stage of generating potential fire candidates. The colour information in RGB channels is also modelled using Gaussian distribution using training examples. Based on the generated probabilities of the pixels of new observations, instead of directly using this probability, the authors normalized these over the maximum probability from the training examples to construct a colour metrics. Finally, a threshold is used to decide whether a pixel should be labelled as fire.

$$p(x_i|\mu, \Sigma) = \sum_{k=1}^K w_k \mathcal{N}(x_i | \mu_k, \Sigma_k) \quad (2.4)$$

Because the colour interpretation by RGB colour space is sensitive to scene illumination, a Gaussian mixture model using images from different illuminations was trained in the work of [39]. Basically, the mixture model is a combination of K

Gaussian distributions each with its own weight w_k , as formulated in Eq(2.4). With the help of Dirichlet Process and collapsed Gibbs sampling, the model parameters, w_k, μ_k, Σ_k , were obtained. For new observations, the model assigns a probability to the pixels which will be assigned as fire if exceeding a threshold.

In the research work of [40], a fire detection system was proposed based on common visible camera, which consists of six types of flame features: fire colour probability, spatial wavelet energy, spatial temporal energy, flickering energy, dynamic texture analysis and a spatio-temporal consistency energy. The system is rather complicated by multiple intermediate stages, including pixel-based filtering, patch-based processing. The stages can be generally categorized as ROI proposal stage, feature extraction and verification and final decision making by SVM. In the first stage, a manually constructed gaussian mixture model used to filter out fire-coloured pixels. Then wavelet-based edge detection is used to compute spatial energy because flame has more spatial variation than uniform flame-coloured objects. To further distinguish fire-like objects from real fire, pixel-based temporal flicker energy is calculated. The proposed ROI patch is then fed into a pre-trained linear dynamic system, which then produces nonlinear features. The histograms of the extracted features will finally be fed into a SVM for classification.

The ROI extraction used in [41] is solely based on rule-based colour model similar to the method used in [36]. The texture analysis and statistical moments of potential fire patches are then feed to a SVM for classification. Similarly, the work in [42] used a slight improved colour threshold in RGB colour space for initial ROI extraction. A neural network is trained using blue variation to segment foreground moving regions. The final decision is made based on the growth of the ROI from previous stages in three consecutive frames.

The research established in [43] applied a supervised k-nearest-neighbour algorithm to model the background in order to isolate moving targets. Then simple colour threshold in red and value channel of RGB and HSV colour space respectively is used for further potential flame filtering. The value channel ROI is used for histogram of oriented gradient (HOG) extraction, where local gradients are computed. Combining with local binary patten for flame texture feature, a SVM is used for classification.

In summary, colour model for fire-like pixel extraction is the most popular choice in literature. The colour of naturally occurred fire mostly a mix of yellow, orange and red. This distinctive feature is very intuitive and easy to be implemented. The main categories of colour detection are rule-based thresholding and probabilistic modelling. In terms of computation complexity, they are all capable of real-time scenarios. The only difference among these methods is that the probabilistic colour model requires training and fine-tuning before the actual deployment. Whereas the rule-based model is just off-the-shelf. It should be noted that many objects have similar colours to fire, which will also be selected as fire candidates. Therefore, the subsequent filtering is necessary to narrow down the ROI selection.

2.3.3 Moving region detection and dynamic features

Fire wobbles because of the buoyancy and the generated turbulence, making it move randomly and consistently. Popular methods for moving region detection are background subtraction, frame temporal difference, and optical flow. In addition, much research targeted the motion as dynamic features as one of feature selections for both fire and smoke. In comparison with pixel-level colour features, the dynamic features are extracted from patches, where certain geometric measures are used for

interpretation. In the background subtraction method, a reference image needed to be estimated. The reference image is strongly affected by the complexity of the scene, especially in cases where lots of movements are involved. Therefore, the result from background subtraction can be very noisy.

$$B_{n+1}(x, y) = \begin{cases} \alpha B_n(x, y) + (1 - \alpha)I_n(x, y), & (x, y) \text{ stationary} \\ B_n(x, y), & (x, y) \text{ moving} \end{cases} \quad (2.5)$$

The use of background subtraction as in the work by [44], constantly updating the estimated background image B_{n+1} , where α is a time constant determining how long to update the background image. The moving region is then obtained by subtracting current frame by B_{n+1} in Eq(2.5). Similar methods have been used in many other research [45], [46], [47], [48]. This approach requires the camera to be static. Optical flow based moving region detection was also reported in the study by [49], [50].

The research by [51] focuses on the detection of smoke by moving region segmentation extract in addition to dynamic features. The background of an image is estimated using a Gaussian mixture model. The dynamic feature is computed from the rate of change in frame difference.

The study in [52] used randomness of area size of the extract patch as a feature, where the difference in patch area size of two consecutive frame was normalized by the patch area from current frame. A threshold is then used for decision.

Wavelet based temporal behaviour of fire can also be useful features for detection. It has been reported that common flames usually exhibited a flicker frequency of 10 Hz [44], [48]. However, due to the erratic nature of flame flickering, the extracted frequencies disperse over the range of 0-10 Hz, as shown in Figure 2.8. Therefore, this is not viable to detect fire using a specific frequency. A more practical choice is to monitor high frequency band, for instance 5-10 Hz [30].

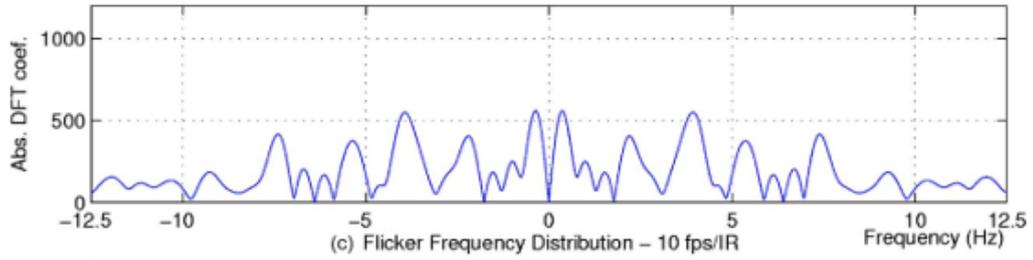


Figure 2.8: Flame flicker frequency spectrum, showing a flame oscillates at random frequency within 0-10 Hz [53]

For smoke detection, wavelet-based energy can be used to isolate smoke region. [46] extracted the variation of the contour of the ROI to the centroid as signals for separating the smoke from the background. Because smoke smears the edges in an image, which will show a drop of energy in frequency domain. The ratio of the patch energy and the background energy is used for identifying smoke region, as shown in Figure 2.9.

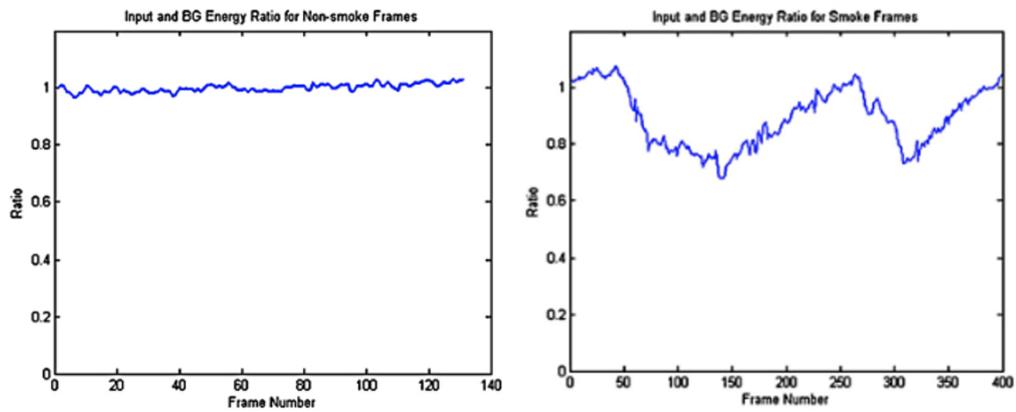


Figure 2.9: Energy ratio of foreground region and background; left: no smoke, right: presence of smoke [46]

2.3.4 Decision-making: machine learning methods

The detection of a fire in the scene is a task of recognition through algorithms based on some extracted features from the images. The types of features for recognition can vary greatly. Most common features to extract from an image are usually the colour information, object geometric properties and so on. For videos, variations over time can be captured to temporal analysis. In addition, the images of objects captured in spectrum other than visible light can be dramatically different from our intuition. This is because the material properties have different interactions with lights of different wavelengths.

The use of machine learning algorithms is to exploit signal features of the subject to help the recognition. Image is just one kind of signals but with a 2 dimensional layout. The research by [54] proposed a systems of three neural networks, which intakes the frequency spectrum in a time window of signals from four IR sensor; the core feature exploits on the flickering frequency (1-20Hz) of the flame, which is distinguishable from other objects. Each network corresponds to one identified false alarm sources. The work by [55] trained a CNN with a SVM for classification of fire and hot objects in the long wavelength IR spectrum. Optimal mass transport optical flow was used as a feature descriptor of flames in [56], which was combined with R,G,B values as input for subsequent neural networks for classification. The research by [57] trained a CNN based on normal RGB images for fire detection, leveraging the public surveillance system. In the work of [58], CNN models were used for fire detection with unmanned aerial vehicle, where several popular CNN architectures were adopted and compared. A Bayesian classifier was employed to distinguish fire from non-fire, which utilised colour and geometric features of the visible image of the fire [59].

In modern video fire detection, after feature selection and extraction, it always comes down to how to recognize the fire. In this case, it is a binary classification to distinguish fire from non-fires. As enough examples are collected, the extracted feature can take some distinctive distributions.

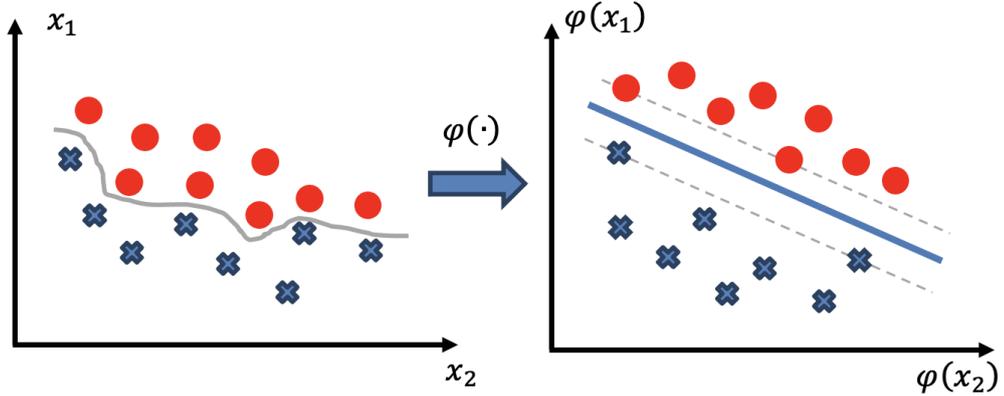


Figure 2.10: Nonlinear feature mapping using kernel function; left: feature space, right: transformed space

For a simple illustration with two dimensions in Figure 2.10 left panel, the fire and non-fire data has the distribution as shown. The classification task is to draw a line, linear or not, to separate each class. Lots of machine learning options available to perform this task. The performance of which is sometimes not dictated by the algorithm itself, but by the data available and other practical considerations.

2.3.4.1 Support vector machine

Support vector machine (SVM) was first introduced by Cortes and Vapnik in 1995 for the purpose of two-group classification problems [60]. Referring to Figure 2.10, the algorithm is to determine a function $f(x)$ to divide one group from the other, which take the form as in Eq(2.6):

$$f(x) = \mathbf{w}^T \varphi(\mathbf{x}) + b \quad (2.6)$$

From training, the weights w and bias b are obtained by convex optimization. The form of this equation indicates a linear hyperplane. With the kernel function $\varphi(\cdot)$, the original feature space is transformed into non-linear higher dimensional space where those transformed features are easily separable, as demonstrated in Figure 2.10 right panel. The popular choices of kernel functions $\varphi(\cdot)$ are linear, polynomial and RBF (radial basis function) as formulated in Eq(2.7):

$$\text{Linear kernel: } \varphi(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j \quad (2.7a)$$

$$\text{Polynomial kernel: } \varphi(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + 1)^p \quad (2.7b)$$

$$\text{RBF kernel: } \varphi(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2} \quad (2.7c)$$

The SVM for classification algorithm is essentially to find an optimal hyperplane with soft margins to separate the data classes, by means of numerical optimization. For a perfectly linearly separable case, the objective of training a SVM is to define the optimal separating hyperplane H_0 that has greatest margin from the closest data points, i.e. d in Figure 2.11. Those data points lying on the marginal planes are support vectors.

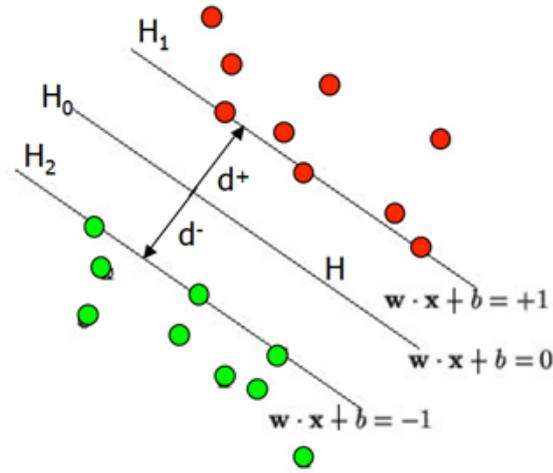


Figure 2.11: Finding the optimal hyperplane by maximizing the margin d , the data points lying on H_1 and H_2 are support vectors [61].

For points lying on the upper marginal plane or lower marginal plane, the expression of the hyperplane H_0 evaluates to positive or negative one, respectively. The expression of margin d is obtained by calculating the shortest distance from the point vector to hyperplane H_0 , taking the expression as in Eq(2.8):

$$d = \frac{\mathbf{w}}{\|\mathbf{w}\|} \cdot \mathbf{x} + \frac{b}{\|\mathbf{w}\|} \quad (2.8)$$

The margin d is dependent on the normal vector \mathbf{w} and the bias b of the hyperplane H_0 . d is then assigned with a direction by multiplying the class labels y_i to the first term. The algorithm boils down to an optimization problem that tries to solve for the optimal separating hyperplane, which is defined by normal vector \mathbf{w} and bias b , to give the largest margin d . To simplify the process of solving the above formulation, through mathematical manipulation the question is reformulated into a

convex quadratic optimization problem, taking the form as in Eq(2.9):

$$\underset{\mathbf{w}, b, \zeta}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \zeta_i \quad (2.9a)$$

$$\text{subject to } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \zeta_i \quad (2.9b)$$

$$\zeta_i \geq 0 \text{ for any } i = 1, \dots, m \quad (2.9c)$$

The addition of ζ is a compensation for data points that cannot meet the constraint. If ζ is exceedingly large, all examples will satisfy the constraint, rendering the classifier not classifying anything.

2.3.4.2 Neural networks in general

The idea of neural networks originated in the 1940s, inspired by the biological process of the brain [62]. However, it was made practical in the 70s and 80s by the introduction of backpropagation for training, from which the gradients of the objective function with respect to the weights are computed for update [63]. The typical structure of a neural network is shown in Figure 2.12. The input vector is passed to corresponding nodes by taking the inner product of the weight w , which is then transformed by a choice of nonlinear function before passing to the next layer. This function maps the input from the above layer to the required range, such as $[0, 1]$ or $[-1, 1]$. This nonlinear function is called activation function. The nonlinear mapping of the input data enables neural networks with deep enough layers can mimic any complex functions.

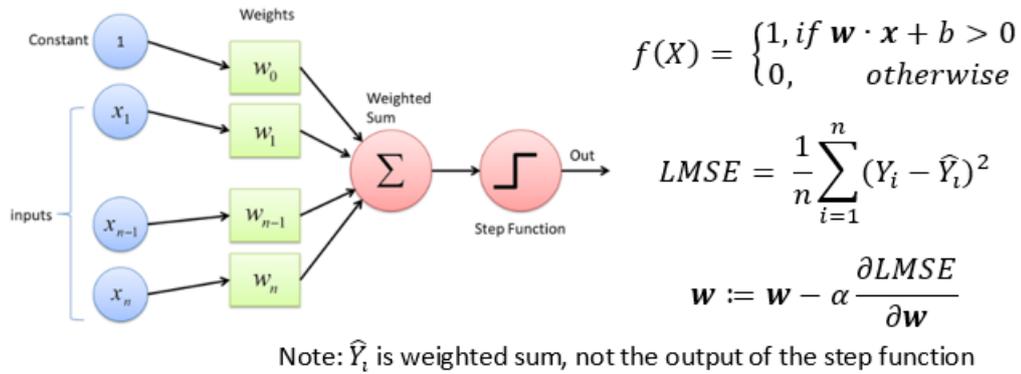


Figure 2.13: Schematic diagram of a perceptron with formulations: decision function, objective function, and weight update [65].

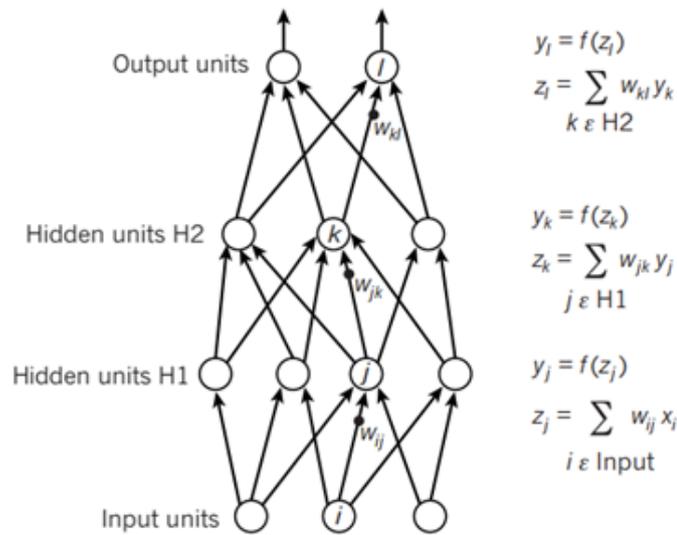


Figure 2.12: Feedforward neural network with two hidden layers [64].

The elemental unit of neural networks is perceptron, which is illustrated in Figure 2.13 with mathematical expression. A perceptron can be used as a logic function to represent the fundamental Boolean operator, such as AND and OR with appropriate weights and bias. The output of a perceptron is based on thresholding of the decision surface evaluation. Multiple of perceptrons can interpret any Boolean operations.

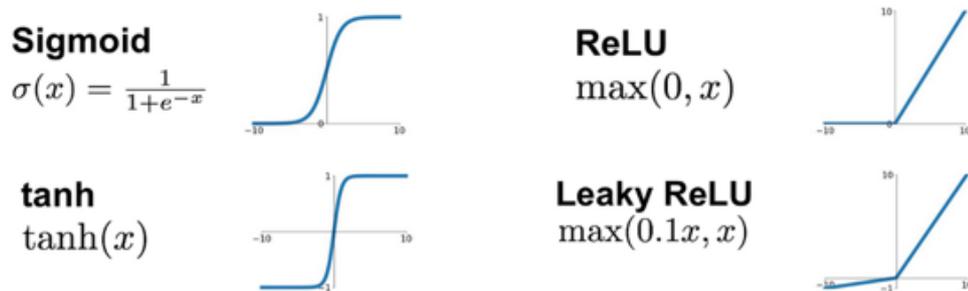
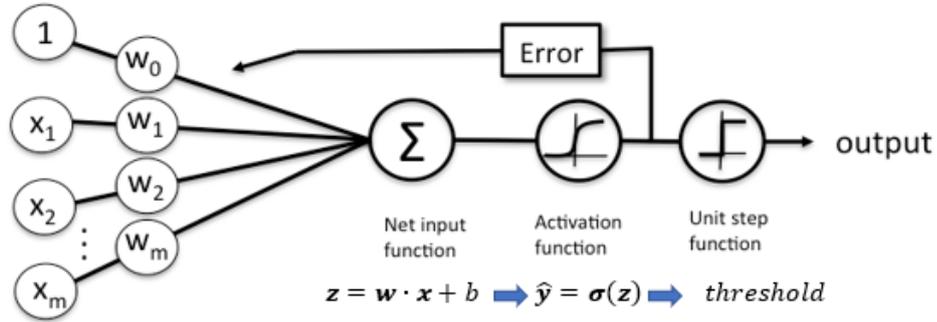


Figure 2.14: Common choices for nonlinear activation functions: *Sigmoid*, *tanh*, *ReLU*, *Leaky ReLU* [66], [67].

Even an interconnected network of perceptrons is still linear in nature, bounded by linearity of the expression of $f(x) = \mathbf{w} \cdot \mathbf{x} + b$. To model more complicated function mapping, the nonlinear capability is essential. An extra activation unit is added so that its output is nonlinear of its inputs. Meanwhile, such function should also be differentiable that the gradient-based update is operable. The earliest choices are the logistic *sigmoid* and hyperbolic tangent, as stated in Figure 2.14. However, they are depreciated over the years after deep learning becomes prevalent. This is because as the network gets deeper and denser the flat part of these functions gives very small gradients during the weight update, which will make the training process very slow for practical use [68].

With added nonlinear activation function to the dot product, it forces the weighted sum to be interpreted as conditional probability: $P(y = 1|x)$, meaning the probability of being in class y given the data x . This turns a perceptron into a logistic classifier, which is the fundamental building block of a neural network, which is shown in Figure 2.15. Noting that the classification is based on probability thresholding on the output of the activation function, it is a binary classifier.

To train the classifier, a cross-entropy loss function is used to measure the



Schematic of a logistic regression classifier.

$$f(X) = \begin{cases} 1, & \text{if } P(y = 1)|x > 0.5 \\ 0, & \text{otherwise} \end{cases}, \text{Loss}_{CE} = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

$$\mathbf{w} := \mathbf{w} - \alpha \frac{\partial \text{Loss}_{CE}}{\partial \mathbf{w}}$$

Figure 2.15: Schematic diagram of a logistic classifier with formulations: decision function, objective function, and weight update.

difference between predicted \hat{y} and true label y , using the objective function Loss_{CE} . This loss function comes from the fact that the likelihood of a binary classification is describe by Bernoulli distribution, which take the form of $P(y|x) = \hat{y}^y(1 - \hat{y})^{1-y}$. The gradient update of w and b are computed by taking derivatives of log likelihood of the loss function, which is called conditional maximum likelihood estimation. Loss_{CE} is essentially a convex function, where a global minimum can always be achieved.

Logistic regression is similar to linear SVM as both have $\mathbf{w} \cdot \mathbf{x} + b$ for a decision hyperplane separating the two classes. In nonlinear cases, they both employ a nonlinear function to transform the input vectors into higher dimensional space where the data is linearly separable. However, the way to determine the weights and bias is different. In SVM, the expression is found based on support vectors which is a subset of all training examples. Whereas the logistic classifier uses gradient descent by minimizing the log likelihood with respect to the weights to update the weights and bias. The gradient update is computed using all training examples. To use

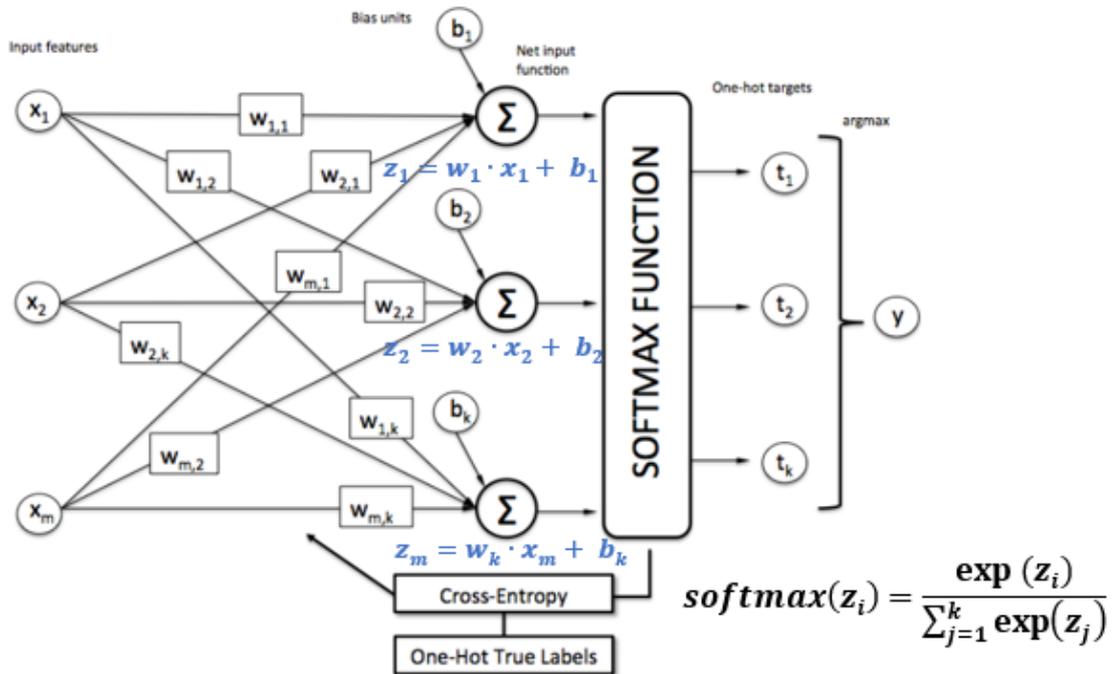


Figure 2.16: Schematic diagram of a multi-class logistic classifier with *softmax* function [69]

logistic regression for multi-class classification, the classifier will produce a vector of probabilities with the same dimension as the number of classes k . This is achieved using a nonlinear *softmax* function as shown in Figure 2.16. Its denominator is a normalization of the probability of each class to have a range of $[0,1]$. Similarly, the weights are updated by gradient descent using cross-entropy loss, but derivative are taken with respect to each class.

A neural network can be regarded as many logistic classifiers stacked together in an organized structure, without the step function for thresholding, as shown in Figure 2.12. The layers in between input and output layers are called hidden layers. It is also called feedforward network because the outputs from of a unit is passed to the next layer only.

It should be noted that the output layer in the figure is a vector of real numbers.

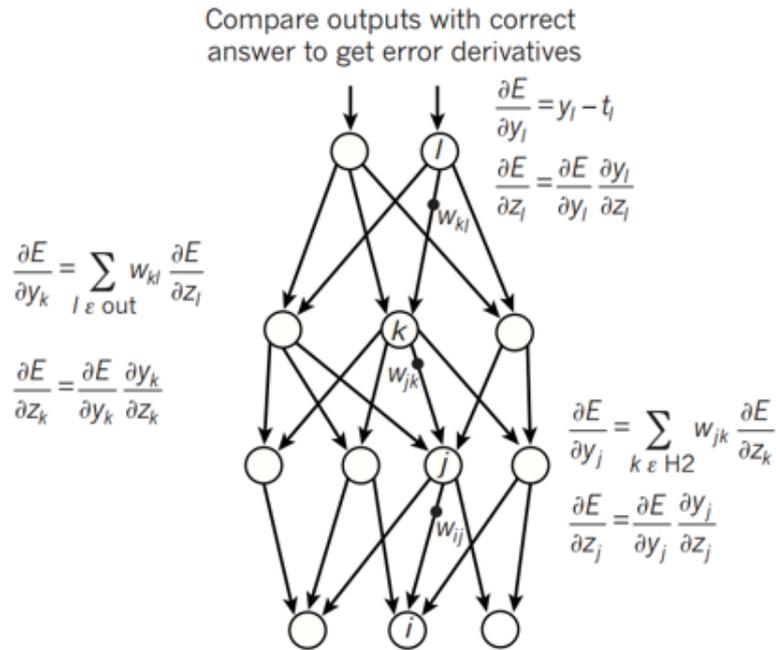


Figure 2.17: Schematic diagram of backpropagation to update neural network parameters [59].

Probabilities for each class for classification can be computed using the *softmax* function. Same as previously introduced, the learning is based on the updates of the w and b using gradients of the loss function. Since neural networks have multiple layers, the gradient is first computed from the output layer and propagate to reversely to the input layer. This is the key to make the practical use of neural networks in the late 80s. As illustrated in Figure 2.17, the partial derivatives are taken with respect to the intermediate parameters using chain rule. The weights are updated iteratively until the loss is minimized according to some criterion.

2.3.4.3 Convolutional neural network (CNN)

Convolutional neural network can be think of as an extension of the traditional neural networks with integrated feature extraction component. A typical architecture of a CNN is shown in Figure 2.18. It is more convenient to deal with 2D or 3D data

comparing to traditional neural networks, which takes 1-D vector as input. The main components are convolutional layer, pooling layer and fully connected layer. From the functional point of view, the network can be divided into two parts: feature extraction, and fully connected neural networks for classification. In the feature extraction layers, convolutional and pooling layers with a nonlinear activation layer in between forms a basic unit.

In the convolutional layer, a number of filters has corresponding patches of the previous layer, where the feature maps are formed by taking the convolution and passed to activation function. This is just the same as using filters to detect edges in image processing, but the difference is the parameters of filters in CNN is obtained by training through optimization. The number of filters and filter size are usually determined through trial and error. As shown in Figure 2.18, the filter sweep over the input from the previous layer, then produce a convoluted map, which is then applied with a nonlinear activation function element-wise to generate a feature map. A pooling layer is used to down-sample the output from the convolutional layer thereafter. These together forms a structure, which can have many repetitions depending on the architecture design.

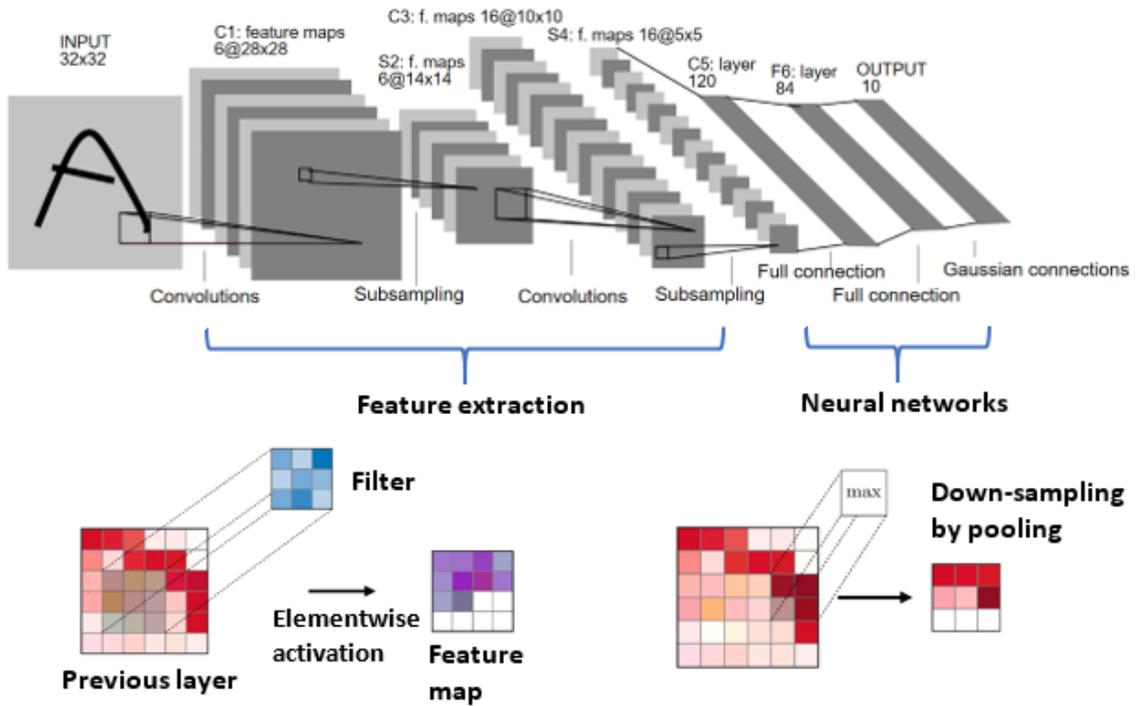


Figure 2.18: Schematic diagram of LeNet model for CNN illustration; lower left: convolution layer, lower right: pooling layer for down-sampling [70], [71].

The output of the feature extraction part will be flattened to a 1-D vector for the subsequent neural network for classification, which is usually called fully connected layer. The fully connected layer is the same as the traditional neural networks as elucidated in previous sections. The choice for classifier is not limited to *softmax* for a probabilistic interpretation at the end of the neural networks. Training of a CNN is just similar as a regular neural network by backpropagation, where the only difference is the addition of trainable weights in the filters. As deeper it gets in the feature extraction part, the size of the feature maps from each layer shrinks because of the down-sampling, whereby the features are more focused on the details.

In summary, the fundamentals of SVM and neural networks have been explained. As a variant of the traditional neural network, the CNN is also discussed. In terms of similarities for classification, they both use nonlinear functions to map the input

into higher dimensions to generate nonlinear decision boundaries. However, they are fundamentally different regarding how the decision boundary is obtained through optimization. For neural networks, backpropagation is used to update the weights iteratively by taking partial derivatives until the cost function is minimized. By contrast, SVM is more geometrically motivated [72], where an optimal separating hyperplane is found by maximizing the margin using a subset of the data called support vectors.

The training for these algorithms is essentially an optimization problem. For neural networks, gradient based optimization is used to minimize the cost function. As presented in previous sections, solving for SVM is a convex quadratic optimization problem subjected to inequality constraints. Solving the constrained optimization requires Lagrange multiplier, from which the weight w and bias b is obtained [73]. In comparison, the cost function of a logistic classifier is quadratic, and its optimization is convex, which means a global minimum is guaranteed. However, its interconnection to form a neural network render the cost function non-convex even though the cost function bears the same quadratic form. This is because of the nonlinearity of the function $\hat{y}(x)$ that maps the inputs to output class labels. Neural networks are always thought to be stuck in local minima as its not convex. However in practice, it is usually not a problem.

Another major difference is the integration of feature extraction functionality to the classifier for neural networks. Whereas SVM is by itself a classifier. As previously explained, features are generated from the data through the intermediate layers (hidden layers), which is fed into the final classifier (*softmax*). CNN is the same but with spatial feature extraction by 2-d filter bank. Neural networks are designed to take raw data directly, especially for the case of CNN (images as inputs).

Once it is trained, the features are automatically extracted. For SVM, features are usually extracted manually according to domain expertise before feeding to the SVM for classification. It is viable to replace a *softmax* classifier by a SVM for a neural network.

In practice, the choice of a classification algorithm depends on many factors. The most significant one is the amount of available data. For example, the CNN model in [70] has 60000 trainable parameters. The network is trained on 60000 examples of handwritten digits images having size of 32x32. The large number of parameters needs to be trained to achieve desirable performance. To device a neural network model for specific task, one crucial question before undertaking is whether enough data from the task domain is provided. This gives an edge of SVM over neural networks because generally SVM will perform very well with small dataset, which will usually intake manually crafted features from the dataset. However, if the dataset is abundant, training a neural network is actually faster than training a SVM. This is because the optimization process requires the computation of the Gram matrix that measures the similarity of every data point pairs by taking inner product. The memory requirement of this matrix scales quadratically with respect to the number of training examples.

Moreover, SVM is more deployable than neural network models. The structure and combination of layers is determined through iterative experimentation. Not only fine-tuning the hyperparameters is workload-intensive, but the training requires many practical tricks to achieve a good balance of the variance and bias trade-off to avoid over-fitting. For example, adding batch normalization to re-centre the output of each layer can accelerate the training speed [74]. Another practical trick is dropout, which randomly shut down the connections at nodes to avoid over-fitting [75].

2.4 Video fire detection in IR spectrum

The use of IR cameras for fire detection is significantly different because objects appear differently in terms of image texture. Especially in the long wavelength IR range, some objects in the image are bright because of their own radiation coming from the vibration and rotation bands of the molecules. Figure 2.19 demonstrated some sample images captured by the longer wavelength IR detectors. The overall image is more artificial than intuitive.

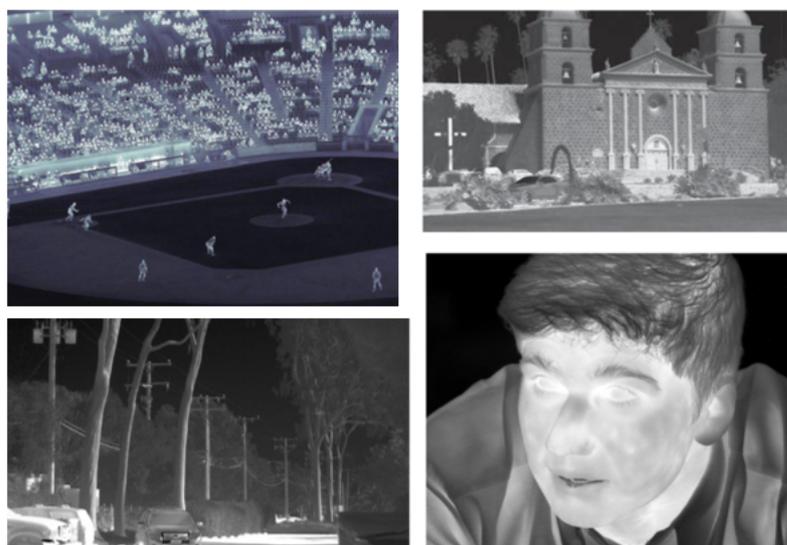


Figure 2.19: Long wavelength IR images; top left: MWIR, nBn InAsSb FPA [76]; top right: uncooled VO_x microbolometers LW IR [77]; bottom left: MWIR, nBn InAsSb FPA [78]; bottom right: MCT, mercury cadmium telluride (HgCdTe), liquid argon cooled [79].

The use of IR detectors for video fire detection is scarce. The main reason could be the high cost comparing to common silicon-based CMOS image sensors. Despite of significant advancement in image sensor fabrication technologies, IR camera sensors are still significantly more expensive. Almost all the existing fire detection research used microbolometer based IR detectors. Even the most cheapest such IR cameras are more expensive than its visible counterpart, not to mention they have much lower



Figure 2.20: Fire images captured by long wavelength microbolometer IR detector (7-14 μm) [80], [81], [53].(The red highlights were post-processed in the reference, not the original captured image.)

resolutions. In addition, from image processing point of view, the properties of a flame in different IR spectrum are not well understood. Figure 2.20 shows the images of flames by microbolometer IR detectors.

2.4.1 Long Wavelength IR fire detection

The work by [82] used an IR camera operating in the long wavelength band of 8-12 μm . For detection, the ROI extraction stage combines adaptive thresholding and centre of gravity analysis. Two experimentally determined thresholding were proposed to generate heat source and hot source region as in Eq(2.10) and Eq(2.11):

$$TH_{heat} = I_{avg} + 2I_{std} \quad (2.10)$$

$$TH_{hot} = \begin{cases} 0.5(I_{avg} + 3I_{std}) + 0.5I_{Ymax} & : \text{if } I_{max} < I_{lmax} \\ 0.5(I_{avg} + 3I_{std}) + 0.5I_{max} & : \text{otherwise} \end{cases} \quad (2.11)$$

Then the distance of the centres of gravity of each regions were calculated and thresholded empirically. Regions went through this ROI stage will then be fed to temporal texture variation analysis. All of these features were finally used for decision making.

A long wavelength camera was used for fire detection in the research by [83]. The type of camera was microbolometer with sensitivity of 7.5 to 1.3 μm , and a resolution of 160x120. The foreground is separated by Otsus method, which will produce ROI containing human, fire and other higher temperature objects. The discrimination among those extracted objects is by tracking the centroid of the ROI contours.

The combination of multiple spectrum can promote feature extraction. For example, [81] combines a long wavelength camera with visible camera. Because of the large difference of intensities between human and fire, a histogram roughness is used as a feature. The histogram for fire has a wider span while that of human has a spike at lower intensity range. In addition, the temporal variation of the intensity range of the ROI is also used because human exhibits an almost uniform and small range of intensity.

A neural network was used in the research by [80], where a CNN model was trained for feature extraction. Then the features are fed to a subsequent SVM for further classification. Data augmentation by horizontal flipping and adding salt and pepper noise, whereby the training dataset was enriched. A very high performance of the trained model has been reported. The IR camera used is FLIR-A310 320x240 VO_x microbolometer based detector.

Another research used similar long wavelength IR camera where a FLIR camera

with 160x120 resolution was employed [53] . The first stage was an experimentally determined threshold to isolate local maxima as potential fire candidates. The reported extracted ROI were fire and people in the tested cases, which were shown in Figure 2.20. The discrimination was made by comparing wavelet domain energy ratio of ROI contours, because of the random movement of a flame than relative steady movement of a human. By incorporating wavelet based flicker frequency, a hidden Markov model was trained.

2.4.2 Choice of IR detectors for fire detection

Infrared sensors can be particularly useful for fire detection because the flame emits more energy in the infrared spectrum as illustrated in Figure 2.21. According to Plancks blackbody radiation curve, for a typical fire temperature, the curve should lie in between 1000K and 2000K. By integration of the curve, the amount of energy emitted in the visible spectrum is much more than that of visible light (shadowed region). As the temperature of the blackbody decreases, the curve peak shifts towards longer wavelengths.

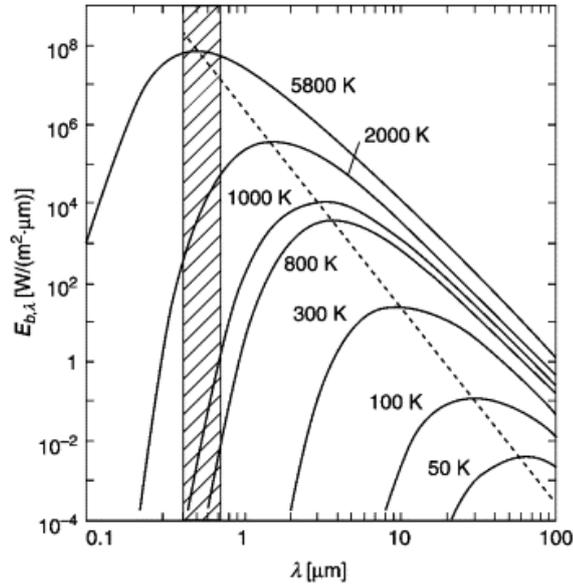


Figure 2.21: Blackbody radiation curves at low temperatures [84].

The infrared spectrum has a wide span, within which one may have to resort to specific sensors for IR sub-band imaging. The whole of infrared spectrum is conventionally divided into several sub-bands, namely Near-IR (700-1400 nm), short-IR (1.4-3 μm), mid-IR (3-7 μm) and long-IR (7-14 μm). For the sake of fire detection, considerations of the choice of sensors are important because of the atmospheric absorption caused by CO_2 and water molecules. From Figure 2.22, the atmosphere has good transmission for spectral windows in NIR band, mid-IR, and long-IR.

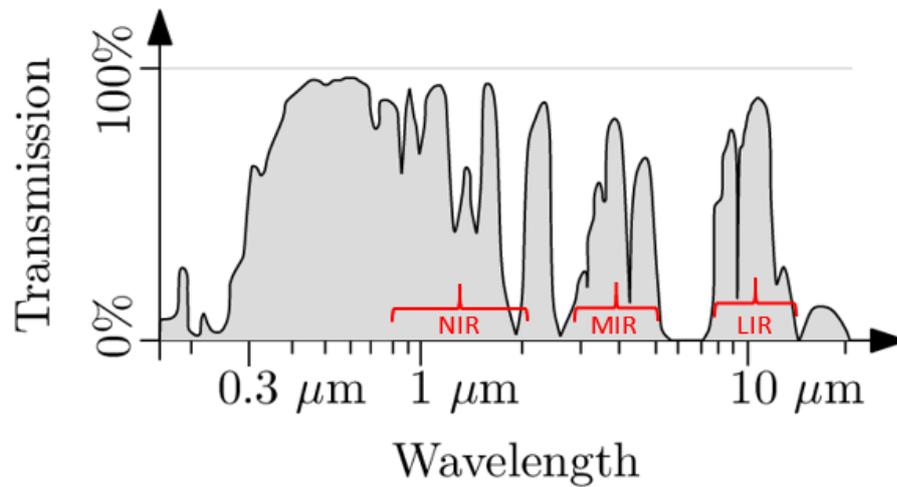


Figure 2.22: The atmospheric transmission with respect to wavelength.

The type of sensors used in literature to detect long-infrared radiations is, without exception, silicon-based microbolometers, simply because of its relative lower cost compared to short and mid uncooled IR detectors. The advance in CMOS monolithic fabrication with silicon makes it possible to produce high quality microbolometers in batch, therefore at much lower cost [85]. However, the cost is still high enough to hinder its prevalence. In addition, through this fabrication process the produced sensors achieved higher dynamic range and better linear response, which is suitable for a variety of applications [86]. The most common materials in the fabrication of microbolometers are VO_x , αSi and silicon diodes.

According to the best of the authors knowledge, there is not a single case in literature using the mid-IR spectrum for fire detection. The main reason would be the tremendous cost of those detectors; not only do they have higher price in fabrication but the required cooling adds more cost. For safety surveillance, the amount of deployment of the thermal cameras can be huge, which in turn the cost-effectiveness is a necessity. Common types of mid-IR detectors are: InSb (indium antimonide) and HgCdTe (mercury cadmium telluride). The InSb photodiode is sensitive to the

spectrum of 1-5 μm , which has to be cooled to a temperature range of 77 to 160K with optimized quantum efficiency [87]. The main application of this type of detectors is for low-background astronomy applications [88]. On the other hand, the HgCdTe detectors have a much wider range of sensitivity compared to InSb, covering from 1 to 20 μm from SWIR to LWIR. Depending on the choice of doping and substrate to fabricate the HgCdTe sensor array, its sensitivity can be tuned to specific sub-bands [86]. Then for each dedicated sensitivity range, its corresponding cooling strategies should be applied. The main application of such a detector is also for astronomy observations.

Apart from the advantages offered by using infrared detectors, there are some drawbacks worth considering for the system design. Using different sub-band detectors in the IR spectrum could give very distinctive appearance of the environment, where disturbances come from different sources. For example, background thermal radiation can produce false alarm in the spectrum above 5 μm [89]. This is because objects at lower temperature in the surroundings will emit considerable thermal radiation, rendering many objects looking saturated. This is very obvious from the provided thermal image examples. In contrast, in shorter IR spectrum from 1.0-3.0 μm , most of the disturbance comes from sunlight reflections.

The use of Long wavelength can be very effective in ROI extraction because flame will always have high intensity comparing to the background. The resultant image is readily bi-modal to make Otsus method very effective. However, more processing and decision making will be intended to discriminate fire from other bright objects.

2.5 Sensor hardware: focal plane array

Focal plane array is a general terminology of an imaging system, which means an array of detectors located at the focal plane of an optical imaging system. The electrical signals generated by the detectors, whether from photoelectric or thermal effects, have to be digitized through electronic readouts. The combination of detectors with electronic readouts forms the fundamental architecture. The advances in fabrication have enabled the focal plane array to be made in a monolithic manner. Specifically, in monolithic fabrication, the signal readout circuits are integrated into the detection material. The arrays can be produced in bulky pieces with less production stages, increased yield and lower costs. The two well-known imagers from silicon fabrication technologies, namely charge-coupled devices (CCD) and complementary metal-oxide-semiconductor (CMOS) imagers, are realized in this monolithic approach. Therefore, CMOS and CCD are different approaches to transform the generated charges from photosensors to digital signals. This transformation is realized by the readout integrated circuits (ROIC) [90]. The basic element for CCD is the MOS capacitors while MOSFET is for CMOS.

2.5.1 CMOS vs CCD

After the electrons generated from photoelectric effect, they need to be stored and transported for subsequent digitization. In CCD detectors, the pixels provide potential wells from the applied voltage clocking to achieve electron charge storage and transportation. The clocking performs the timing of the applied voltage to shift the charges to an output amplifier. As shown in the figure below, the elements for clocking are the phases, which can be regarded as charge manipulator. The silicon layer located below the phase layer confines the generated free charges. As the clocking

calls upon each phase, the corresponding charges will be shifted. This charge shifting process is illustrated in Figure 2.23. Then the output amplifier transforms the charge into voltage. Finally, the chain of voltage output is converted into 2D images through external electronics. The common types are 2, 3, or 4-phase CCDs, where each phase is connected to a separate external voltage sequence.

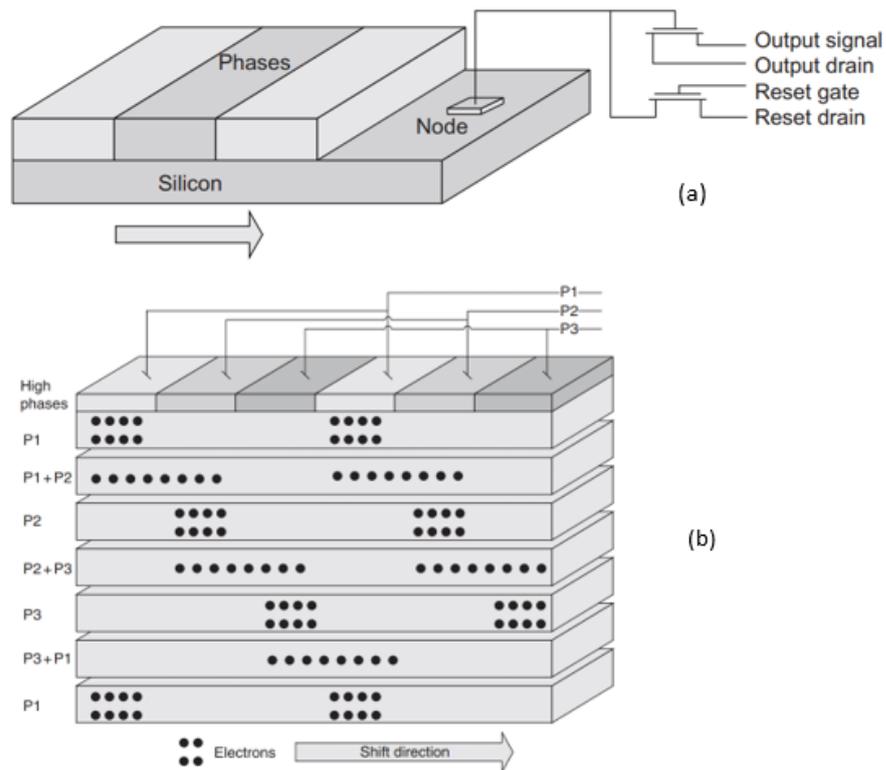


Figure 2.23: CCD charge integration and shifting architectures: (a) charge readout process; (b) 3-phase charge shifting process

While CCD sensors require external readout and controls, CMOS has the control transistors integrated. The whole control circuits are realized in several metalization layers. The integration of control circuits provides the on-chip processing ability for CMOS sensors. Functionalities, such as setting the exposure time, converting charges to voltage and readout timing, are all done within the sensor array.

Since the release of back-illuminated CMOS image sensors from SONY, it soon

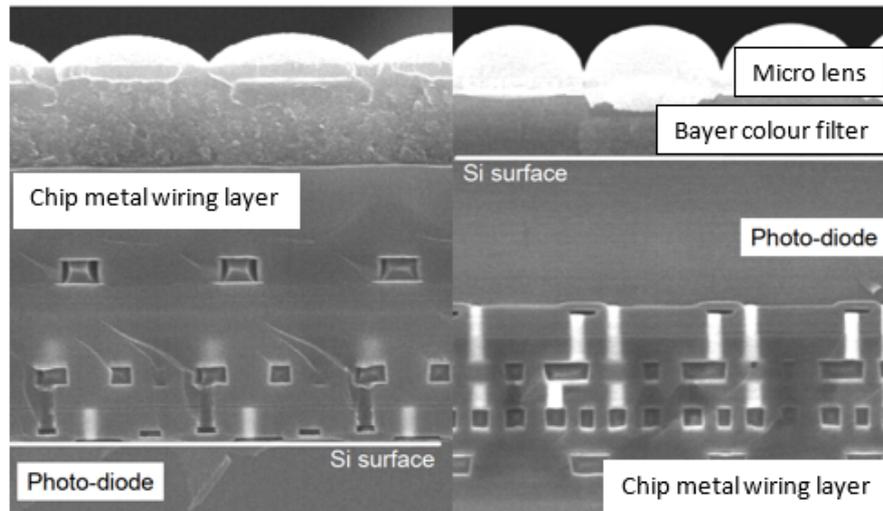


Figure 2.24: Magnified cross-sectional view of CMOS image sensor pixel [91]. left: front-illuminated CMOS pixel; right: back-illuminated CMOS pixel.

became dominant in the portable device market [91]. Comparing to the traditional front-illuminated CMOS sensors, they have superior quantum efficiency compared to frontside illuminated sensors. The Figure 2.24 shows the magnified comparison of frontside and backside illuminated CMOS sensors. The active photodiodes are embedded in the silicon substrate layer. The metallization wiring layer, containing all the control circuitry, is placed below the diode layer, giving a closer contact with the micro-lens layers. This is a natural choice to improve the sensors quantum efficiency because there are less obstacles to reflect, block or scatter the incident lights.

Silicon detectors are innately sensitive from UV to NIR. With the back illumination configuration, it can even be used to detect soft X-rays. However, towards longer wavelength, a sharp drop of quantum efficiency usually exists at around 1100 nm. The reason is low photon absorption due to its indirect band gap. Different materials have different photon absorption depth as a function of wavelength. For silicon, the absorption length of photons near 1000 nm is significantly larger than that of shorter wavelengths, as depicted in Figure 2.25. To provide efficient detection

in the NIR region, the thickness of the silicon layer needs to be at least 100 microns. However, most silicon sensors are produced with thickness of 10-20 μm .

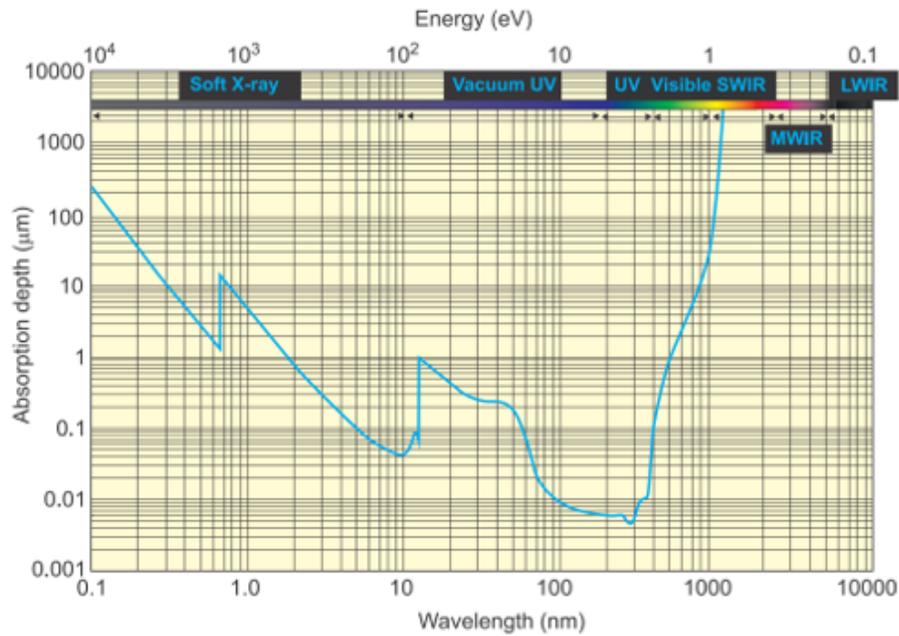


Figure 2.25: Photon absorption depth and energy in silicon [86].

In summary, CMOS image sensors became more popular than CCDs. CMOS sensors have more complex architecture than the CCDs. Despite the superior sensitivity and quantum efficiency of CCDs, the demand for low-power with on-board processing ability is one of the reasons for the increase of CMOS sensors.

2.5.2 Infrared detectors

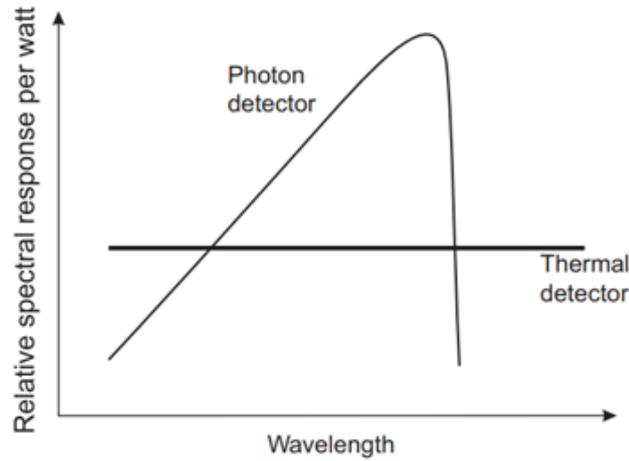


Figure 2.26: General spectral response: photon detectors vs thermal detectors.

In general, IR radiation detectors have two categories according to the sensing mechanisms: photon detectors and thermal detectors. Photosensitive detectors usually have wavelength dependent spectral response for different materials, while it is uniform for thermal detectors, as illustrated in Figure 2.26

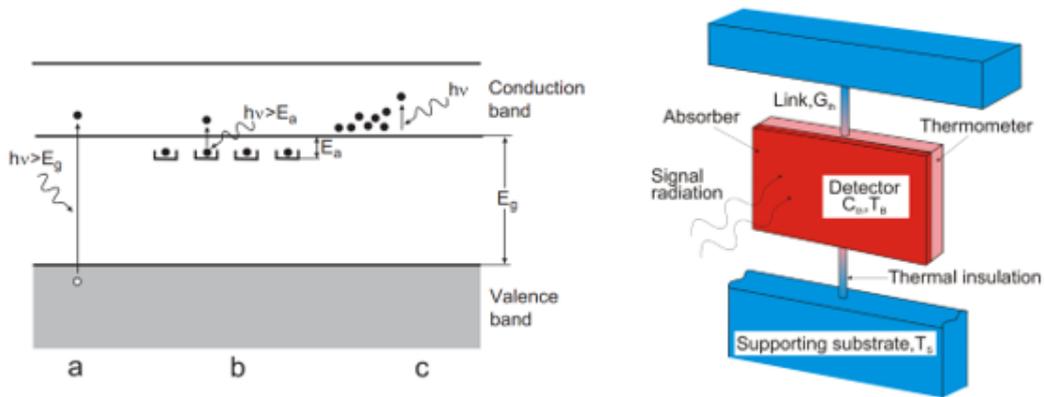


Figure 2.27: Sensing mechanisms: photon detectors vs thermal detectors [86].

2.5.2.1 Photon detectors

For photon detectors, the material absorbs the incident radiation through the interaction with electrons in three forms: electrons bounded to the lattice atoms,

bounded to impurity atoms or free carriers. When the energy of the incident radiation exceeds the band gap energy of the material, electrons will be ejected from the material. The strength of this interaction give rise to the difference in the generated photo-current, which is fundamentally wavelength dependent, as shown in Figure 2.25. Based on the different interaction mechanisms, photon detectors can be categorized mainly as: extrinsic, intrinsic and photo-emissive detectors [88]. The mechanisms of generating electrons are shown in Figure 2.27. The emitted electrons are the primal constituents to form an image.

2.5.2.2 Thermal detectors

Unlike photon detectors, which rely on photoelectric effect to generate electrical signals, thermal detectors absorb the incident radiation to change the physical properties of the material. Usually, the temperature, resistance or electrical polarization of the sensing material is changed, which is then used to generate electrical signals. A schematic diagram of the structure of a typical thermal detector is shown in Figure 2.27 right panel.

In general, photodetector type of IR sensors at longer wavelength range usually needs cooling systems in order to be functional, as shown in Figure 2.28. By contrast, thermal detectors do not need cooling. Most of the IR detectors that requires cooling is used for in the field of defence and astronomy. The materials used for IR detectors are mainly semiconductor and alloy based. Based on their properties, different doping and use of substrates and architecture design vary. For example, HgCdTe alloy based detector has a adjustable band gap, which provides great flexibility for the design with respect to different infrared spectrum from short IR to very far IR [92]. Many of the IR detectors have dedicated market targets, which requires bespoke design and fabrication process.

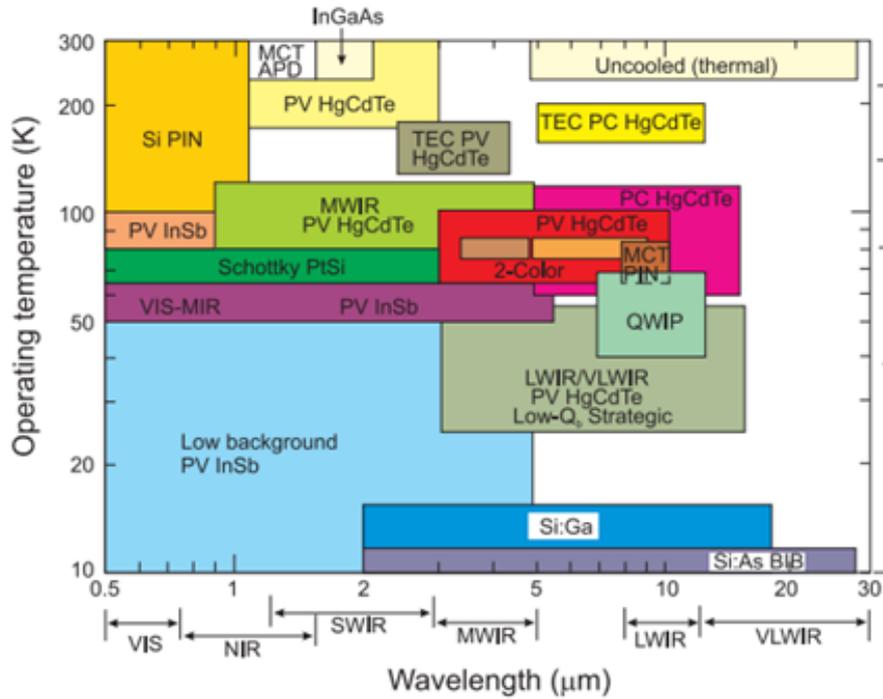


Figure 2.28: IR detector materials and operating temperature [86].

Silicone based uncooled microbolometer is the most popular IR detector comparing to others. It was a breakthrough for long wavelength IR detectors to operate at room temperature. In addition, they can be produced in large volume in monolithic structure based on CMOS fabrication technology, therefore relatively low cost. A hand-hold microbolometer camera for maintenance and inspection can be purchased nowadays with under 1000 dollars from FLIR [93].

The core component of the sensing unit in a microbolometer is the absorber, as illustrated in Figure 2.29. The absorber has a very small heat capacity and a large thermal coefficient, whereby the incident infrared radiation raises the temperature of the absorber. This then generates a large change in resistance, which is detected as a change of current. The commercial microbolometers are mostly made from VO_x , amorphous silicon or silicon diodes, among which the VO_x bolometer is the dominant. The trend for microbolometers is to reduce the pixel size, which is the main limiting

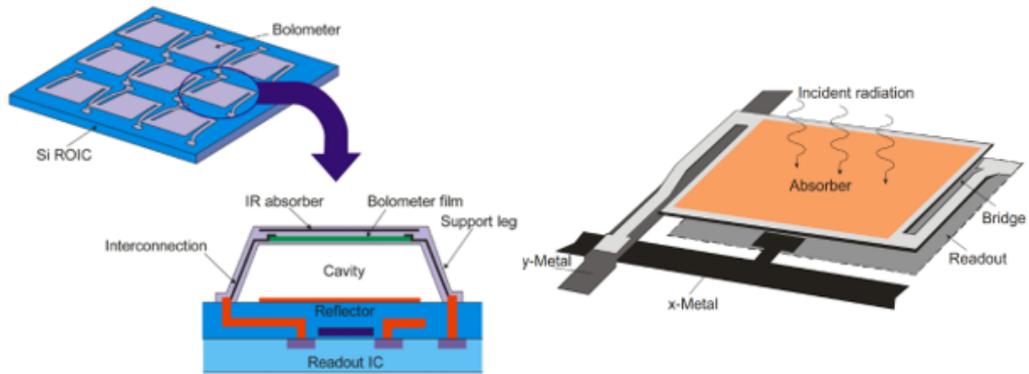


Figure 2.29: Illustration of microbolometer pixel structure [93].

factor than the sensitivity. However, as microbolometer is thermal detector in nature, the response time of the material becomes a compromise with sensitivity [94].

2.6 Fire detection in different scenarios

Based on the type of information gathered from the event of a fire, different machine learning algorithms may be in favour. The detection of fire can rely on algorithms to process a fusion of information gathered through ad-hoc wireless sensor networks (WSN). The problem associated with this technique is the redundancy of data flow gives a heavy network working load. The research by [95] tested a hybrid method combining k-means clustering with feed-forward neural network, Naïve Bayes and decision tree respectively. This approach achieved higher fire detection accuracy with reduced data complexity. The sensorial data was collect by four types of sensor, namely temperature, ionization, photoelectric and CO gas sensor, having the ability to distinguish smouldering from flaming.

$$J = \sum_{j=1}^k \sum_{X \in S_j} \|X - \mu_j\|^2 \quad (2.12)$$

The algorithm computes the mean and radius of the clusters by minimizing square of the Euclidean distance as in 2.12 (J in the objective function, also called distortion measure) between the data points and assigned centre of the clusters, iteratively. The k-means cluster algorithm is used as a pre-processing tool to reduce data redundancy. Only the data within the predefined clusters will then fed to the classifier.

In the case of heavy data workload scenario, machine learning can be an effective way to process the data and make inference. It can be handy comparing with manually engineered features when the data is overwhelming, because the process involving feature extraction and selection may require extensive trial and error. In addition, the choice for the decision-making process should also have a good match to the manually selected features, which entails further experimentation. As previously

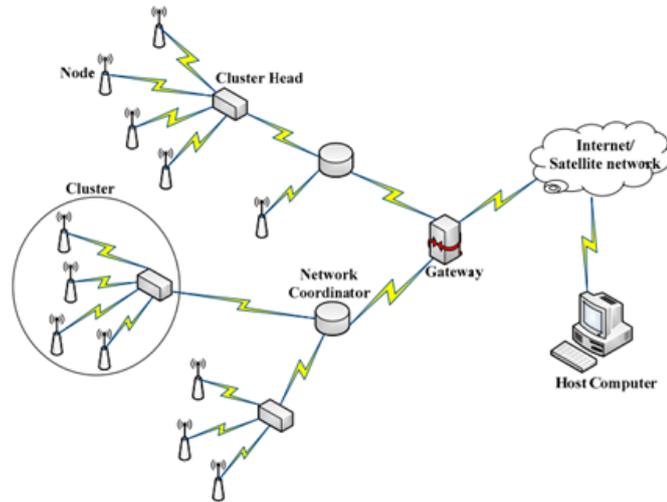


Figure 2.30: Illustration of wireless sensor network (WSN) [96].

mentioned, neural networks have the ability for the integration of feature extraction and decision making, the only thing needed to be done is the selection of structure and fine-tuning the network. This is especially the case for wireless sensor network-based fire detection, where various types of sensors are integrated to generate data. [94] proposed using neural network for forest fire detection by combining smoke sensor, CO and CO_2 sensors, air temperature and humidity sensors. The system was able to detect smouldering and flaming conditions according to the different gas concentrations. For flaming condition, CO_2 , H_2O , NO_x and SO_2 are primarily produced while CO and CH_4 are dominant from smouldering.

The sensor network schematic diagram is shown in Figure 2.30. The unit in each node is an integration of processing chips, solar panel, GPS, and all the previously mentioned sensors. The trade-off between nodes distribution and communication efficiency was reported. Similar research was also reported by [97], where an FPGA unit is used to integrate sensors and neural networks.

2.7 Passive NIR instrumentation in literature

Even though the spectral properties of CO_2 and H_2O consists a major part of the emission spectrum of the combustion products, especially for the case of non-luminous flames, the use of such information for combustion diagnostics is still limited. These two products have banded emission and absorption spectrum in the infrared domain. For the sake of temperature measurement of flames, several popular techniques exist, which can be divided into two categories: measurement of luminous flame (sooty flame) and non-luminous. The most conventional approach is to used thermal couples, which is known for its slow response and intrusiveness. In addition, the measured temperature by this approach may differ from the actual flame temperature because of the catalytic activity of the coating on the thermocouple [98]. The thin-filament technique and two-colour pyrometry are also available to determine flame temperature [99][100][101]. The thin-filament material is ceramic made of silicon carbide (SiC), which is regarded as grey body with known emissivity. This is especially useful for measuring non-luminous flames. The measured temperature can be derived from the measured intensity of the filament in the flame after calibration.

The measurable range of such techniques predominately depends on the sensor spectral sensitivity and the choice of the bands. For example, if a silicon-based CMOS sensor is used, the sensitivity range is up to 1100 nm, will yield a measurable temperature of 375 Celsius with two-colour pyrometry. In terms of practicality, the sensors response at those two narrow bands can be too weak in lower temperature scenarios, which will require longer exposure to obtain usable response. Therefore, its use for more dynamic combustion is limited. The better choice for such case, would be to use a sensor that is sensitive to longer wavelengths, such as InGaAs or silicone microbolometers. To circumvent the limited sensitivity of silicon-based CMOS sensors

with two-colour pyrometry, the banded thermometry can be used. In this approach, a short band of spectral response is collected. However, the uncertainties come from the integration of emissivity for the temperature inference. For precise measurement, the emissivity as function of wavelength and temperature is required, which is usually a difficult task.

For a non-intrusive approach, two-colour pyrometry offers a simple and effective choice for luminous sooty flames. The temperature is inferred by the intensity ratio of the chosen two narrow bands. The selection of the two narrow bands depends on a few factors, such as response change w.r.t temperature, sensor SNR, etc [102]. For example, if the infrared narrow bands are selected, the choice should avoid the strong absorption band of H_2O and CO_2 . In some research, the common digital cameras were employed for such task, which utilized the in-built banded colour filter [100]. This requires additional calibration to account for the emissivity variation, which can introduce more uncertainties. However, the emissivity variation can be ignored in the case of using narrow bands. The measured temperature distribution depends on the soot particles distribution in the flames [103].

In contrast to the aforementioned methods, the non-intrusive temperature measurement of non-luminous flames is not as practical and simple. Most of the existing research relied on techniques based on laser diagnostics. For instance, Rayleigh scattering technique [104], laser induced fluorescence [105] and coherent anti-Stokes Raman scattering [106], have been used to measure flame temperature. In general, the laser diagnostics offer thorough information of combustion physics with high precision and accuracy. Nevertheless, the experimental setup complexity and tremendously high cost rendering its practical use limited [107][108][109]. Other techniques also explored the possibility of temperature measurement using the near-

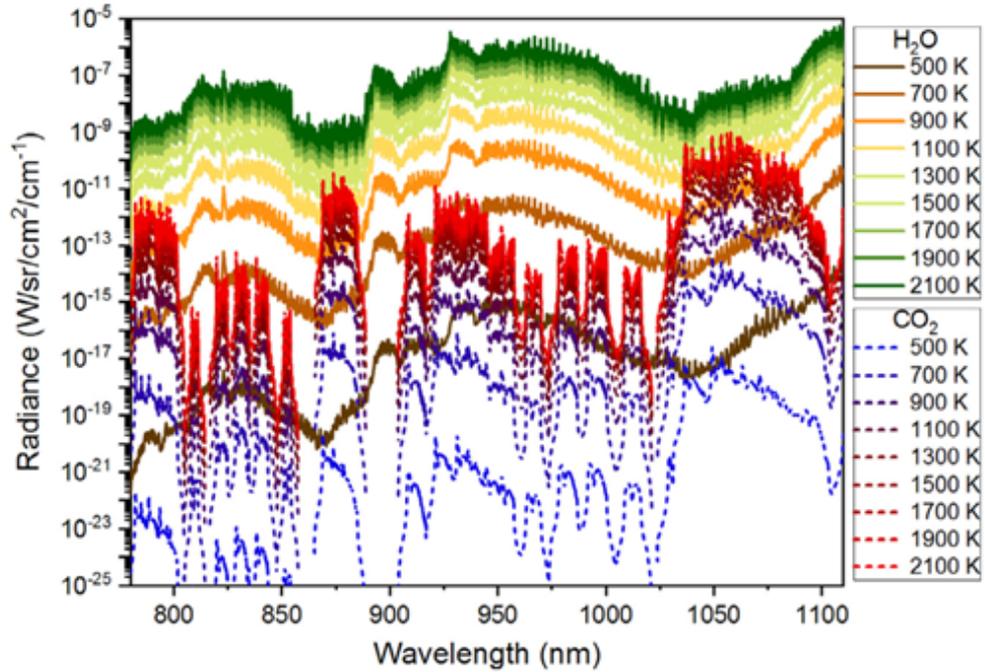


Figure 2.31: Radiance of CO_2 and H_2O at various temperature from HITRAN database [113].

infrared absorption bands of water molecule [107][110]. These techniques exploit the absorption bands of the combustion product, from which the temperature can be inferred. The usage of emission band of the H_2O and CO_2 for such purpose is very scarce. [111] used mid-IR camera to gather emission from water molecule in hydrogen flames at $2.77 \mu\text{m}$, which corresponds to the transient line of H_2O . The temperature and water molecule fraction were then inferred. [112] extended this approach for hydrocarbon flames by using the emission spectrum of water molecules in the near infrared range (less than $1.1 \mu\text{m}$), within which the CO_2 emission is negligible. Such technique is made more approachable as it is the sensitivity range of the common CCD and CMOS sensors, which is an economical choice comparing to other scientific sensors. Figure 2.31 shows the spectral intensity of H_2O and CO_2 of hydrogen flames compared with HITRAN database, indicating that CO_2 emission spectrum is orders of magnitude weaker than that of H_2O .

Other than the use for temperature measurement, CO_2 and H_2O spectrum is other fields of combustion diagnostics, such as premixed flame instability, flame propagation speed, equivalence inference etc. The thermal radiation heat loss of CO_2 and H_2O was reported to have strong effect on lean flammability limit as well as flame speed of near-limit laminar premixed flames [114]. In addition to the emission in IR spectrum of these two species, they also have strong absorption IR absorption bands, from which their re-absorption played a role in fundamental combustion physics. [115] has demonstrated that the re-absorption can increase the flame velocity and expand the flammability limits of premixed flames. Research has been established to show that radiation heat loss and flame stretch can lead to the extinction of near-limit flame, which is further coupled with spectrum re-absorption. Theoretical modelling of the radiation re-absorption was reported that the radiation re-absorption significantly enhance the flammability limits [116].

Those fundamental research into combustion physics in the domain of infrared spectrum, either experimental or theoretical modelling, shed lights on the use for such information in the field of practical combustion diagnostics. [117] deployed high speed infrared imaging in combination with acoustic properties, to study the radiation heat loss and local extinctions in a turbulent premixed combustor as shown in Figure 2.32. A narrow-band filter at $4.38 \mu\text{m}$ was applied to gather hot carbon dioxide emission spectrum, whereby the localized extinction region can be identified.

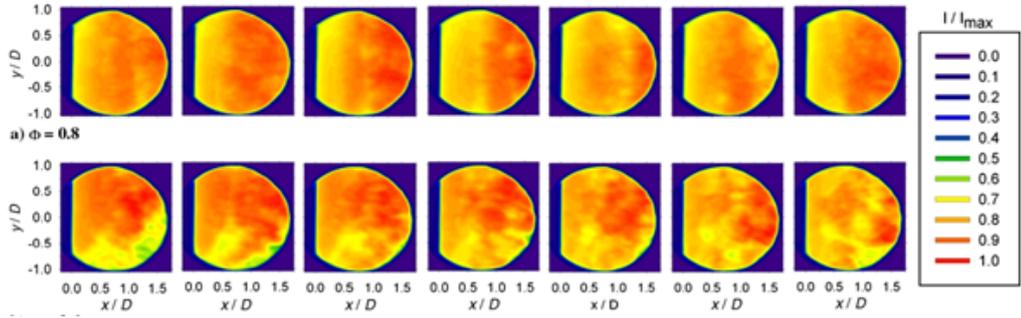


Figure 2.32: Visualization of radiation heat loss and local extinction at $4.38 \mu\text{m}$ [117]

Similar application was conducted by [113], where near-IR camera in spectral range of 780-1100 nm was used to visualize the entropy wave of the downstream of acoustically excited premixed swirling flame. As within this spectral range, CO_2 emission is orders of magnitude less than that of H_2O , and therefore neglected. The variation of distribution of H_2O intensity provides a valuable tool to study the bearings between thermal induced instability and the flow field.

2.8 Literature review summary

Traditional fire detection relies on point-based sensors, which mainly detects the concentration of species or the rate of change of temperature. The mechanism of how these detector works also poses limitations. The species detection is subjected to ambient air flow and transport delay, which limited its effectiveness. This can be improved by point-based photosensitive diode which constantly monitors fire signals in UV or IR. However, besides the high false alarm rate, they do not offer any information in an event of fire. A manual check is always performed without any guidance.

Video fire detection extended with multi-spectrum capability offers the potential for very robust fire detection. The high resolution NIR image sensors is more suitable for machine vision applications, in terms of performance and cost-effective. To

recognize fire, certain image features as fire signature needed to be extracted. The features can be divided into three categories: temporal features, spatial features and colour information. All of which requires image processing pipeline which usually combines multiple stages of computations. These methods differ greatly depending on the choice of camera sensors, because the image properties varies accordingly. The difference in image property is proportionate to the difference in sensitive spectral wavelength.

The development of video fire detection is correlated to sensor hardware advancement, algorithm improvement and knowledge of flame image properties. From the hardware point of view, low-cost is crucial that not only does it motivate research but also can make the actual deployment very economical. Despite the reduced cost of long wavelength IR image sensors, its cost is still much higher than common cameras with much less comparable resolution. As a niche product, its cost is unlikely to be comparable with common CMOS image sensors.

In terms of decision making algorithms, many classification algorithms can be employed, such as SVM, Adaboost, neural networks, Bayesian probabilistic classifiers, Markov models and rule-based models, etc. There is no better choice of classifier than the suitable one. All depends on the how to deal with the image data or the extracted data. In addition, practical considerations have to be made to design the system as a whole. Moreover, the available resources to ensure the functionalities of the algorithms sometimes dictates the choice of algorithms. Fire in near infrared spectrum is rarely study in the field of fire detection. The research established here is to expand the knowledge on fire property and fire detection in the near infrared spectrum, showing a low-cost and robust video fire detection system can be made possible.

Flame spectroscopic analysis is an indispensable way to study the nature of combustion. In addition to the macroscopic physical properties of the flame, such as burning velocities, temperature distributions, flammability etc, spectrum analysis offers ways to study combustion in microscopic domain, which is more closely related with chemical processes. More specifically, the spectroscopic analysis is particularly suitable to study energy release in the process of combustion. Combustion is a complex process, which involves many forms of energy, for instance, kinetic energy, internal vibration and rotation of molecules, ionization and other chemical and electronic energy. The cheap modern CMOS camera sensor offers a very cost-effective way of spectral instrumentation within the range of 300-1100nm.

Chapter 3

Raspberry Pi camera system for NIR fire detection

This chapter focuses on explaining why the near infrared spectrum is a better option for fire detection. A background of NIR as to its engineering application and how it interacts with the world will be provided. In addition, a method of removing the Bayer colour filter will be introduced. The sensor will then be tested to show its improved spectral response. Finally, NIR fire image will be tested to show its greater contrast helps image segmentation.

3.1 Background

3.1.1 The near infrared spectrum: active imaging

The part of infrared spectrum just next to visible spectrum is vaguely defined. In literature, the spectrum from 700 nm to 3 μm is named as short-IR spectrum sometimes. Sub-divisions of this part of spectrum also exists. One of the Divisions can be made to have the near-IR and short-IR, for range 700-1400 nm and 1.4-3 μm respectively. The same convention will be followed herein. The most common detector in this spectrum is InGaAs (Indium Gallium Arsenide) detectors, which usually have a sensitivity range of 950-1700 nm [118].

The active near IR imaging has various applications in the fields of product inspection, anti-fraud, agriculture, biomedical, food and pharmaceutical process control. The NIR absorption characteristics by substances is the fundamental principle for NIR spectroscopy. More specifically, the absorption bands comes from the molecular vibrational overtones and combinations of the chemical bonds in the materials, such as $-CH$, $-NH$, $-OH$, etc [119]. The transition between various vibrational states of the molecules give rise to NIR overtone bands, which usually have multiple peaks. The combination bands originated from the vibrational interactions, whereby polyatomic molecule with their corresponding fundamental frequencies interacts to create new frequencies or shift the existing ones.

Based on the absorption and scattering of NIR in substances, chemical and physical properties can be inferred. For example, in pharmaceutical raw material qualification, NIR is used to infer the water content, because the active ingredients are significantly affected in terms of product consistency and storage effects [120]. The rationale of this is the intensive absorption of $O - H$ bonds at five peaks in the NIR spectrum, namely 760, 970, 1190, 1450, 1940 nm. Moreover, NIR spectroscopy can also be used for measuring particle size distribution of raw materials. This utilized the scattering effect of the powders by measuring the diffusive reflectance, from which the correlation can be established [121]. Furthermore, other uses of NIR spectroscopy were also available in literature, including quantification of crystal structure, visualization of tablet coating and etc.

It was reported that NIR spectroscopy gained its favour rather late comparing to the use of Raman and Mid-IR in pharmaceutical industry, because of its versatile FOV imaging capability and excellent tolerance to sample topology [100]. Similar application is also extensively practiced in the biomedical and food industry due

to the fact that organic substances have distinctive NIR absorption signature from the molecular overtones and combination bands. Usually chemical composition and molecular structure can be inferred using the spectral absorption curves, which requires data mining techniques [122].

3.1.2 The near infrared spectrum: passive imaging

The two common types of cameras readily available commercially for NIR imaging are silicon-based CMOS/CCD cameras and InGaAs cameras. Almost all cameras on mobile devices nowadays are CMOS Silicon sensors, mainly due to its compact size and lower power consumption. While the latter is significantly more expensive for dedicated scientific and industrial uses. Their typical spectral response is shown in Figure 3.1.

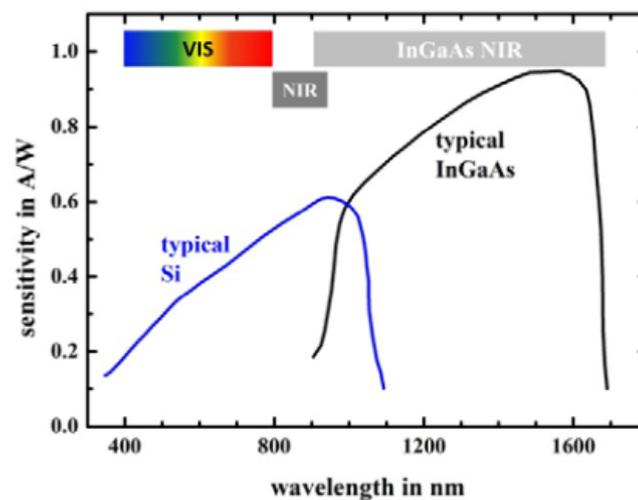


Figure 3.1: Spectral response of Si-CMOS and InGaAs cameras [123].

Unlike Mid-IR and Long-IR images, NIR images capture the reflected lights of the scene, which will require NIR illumination during the night. The images in NIR are noticeably different from normal visible images, yet still appear less artificial comparing to Mid/Long-IR images. The main difference comes from the

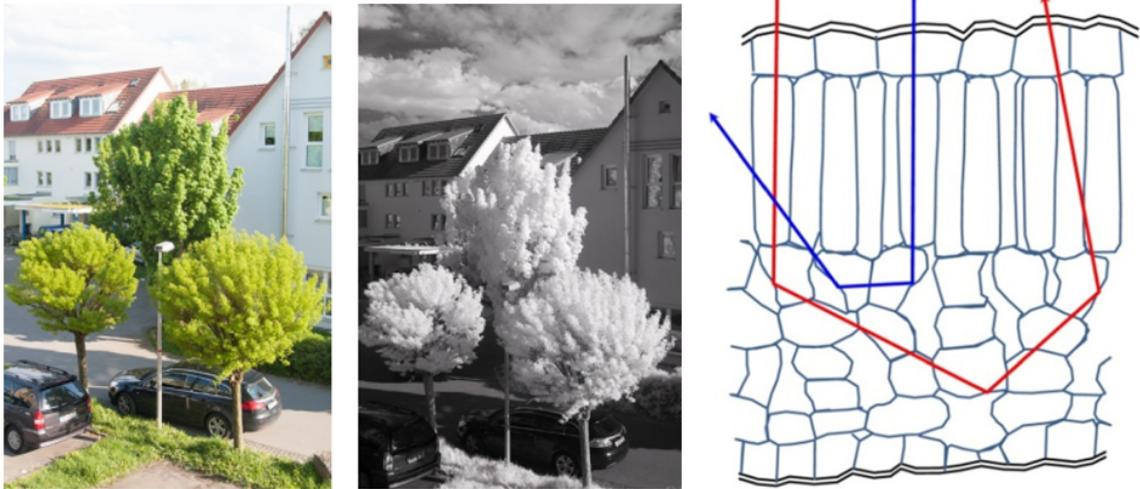


Figure 3.2: NIR and visible image comparison. left: visible image; middle: NIR image; right: NIR scatters more frequently in the deeper layer cells [125].

absorption and scattering of NIR from ambient subjects. The most obvious change is the brightness of green leaves of vegetation, as illustrated in Figure 3.2. This was called the chlorophyll effect by [124]. Due to the better ability to penetrate through substances, NIR rays reach cells in deeper layers where multiple scattering happens, leading to the increased brightness.

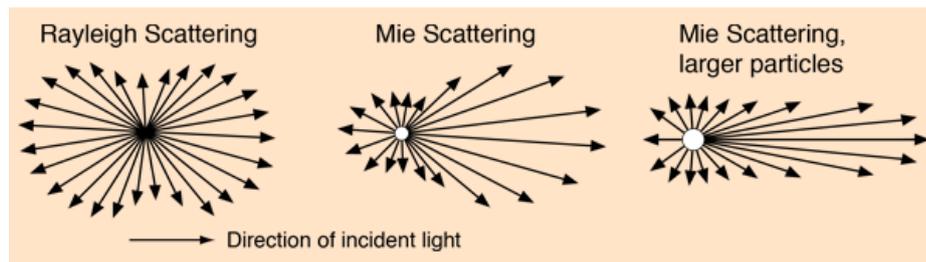


Figure 3.3: Mie and Rayleigh scattering [126].

The next noticeable difference is the large contrast of the cloud and sky because of the Rayleigh scattering and Mie scattering as shown in Figure 3.3. Rayleigh scattering describes the light scattering effect of molecules, which has a wavelength dependent relationship as: $I \propto 1/\lambda^4$. The expression explains the blue colour of the sky as I

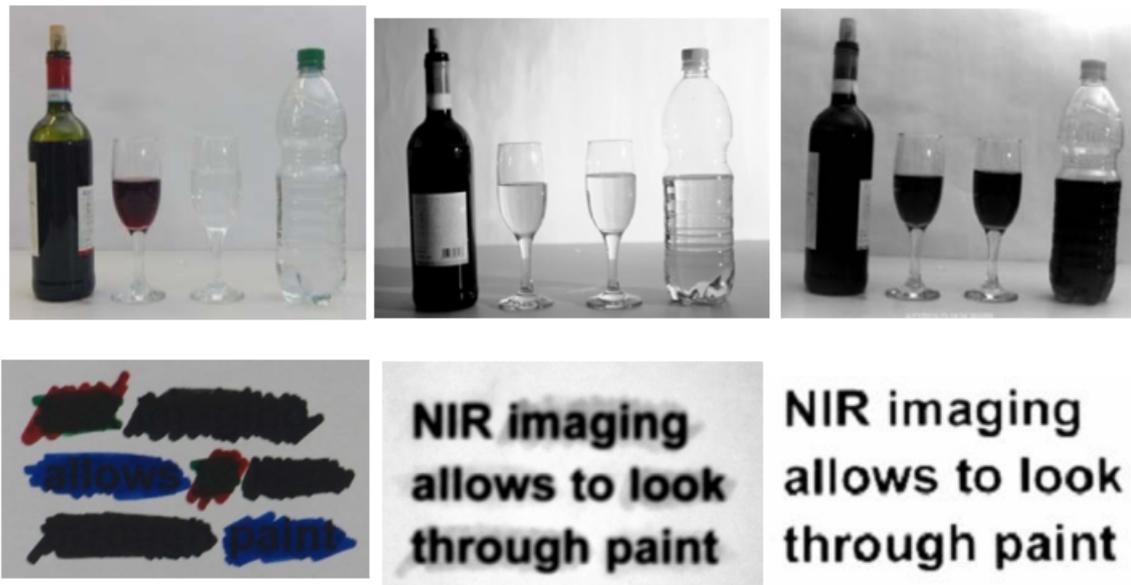


Figure 3.4: Wine and water in visible and NIR spectrum. Top left: visible image; top middle: 830-1100 nm; top right: 950-1700 nm by InGaAs camera. NIR penetrates through pigments. Bottom left: visible image; bottom middle: 830-1100 nm; top right: 950-1700 nm by InGaAs camera [127].

favours shorter wavelengths. For NIR spectrum, the intensity decrease dramatically, resulting darker appearance of the sky. In comparison, Mie scattering corresponds to the scattering effect of much larger particles compared with molecules. The reduction in intensity is much less intense. The combined effect gives the NIR image a contrast look.

Some interesting NIR phenomenon are shown in Figure 3.4. Comparing the images of water and wine captured by 830-1100 nm and 950-1700nm, both wine and water have strong absorption in the higher frequency part of the short-IR spectrum (1100-1700nm). This is mostly likely due to the IR absorption of water molecules within such spectrum range, which coincides with Figure 2.22. In addition, longer wavelength leads to less scattering and more penetration, which explains the clearer view of the paint covered text from InGaAs camera than that from NIR camera.

3.2 Back-illuminated CMOS sensor: Raspberry Pi cameras

Raspberry Pi foundation offers two versions of *Si*-based images sensors: V1, Omnivision 5647 with 500 megapixels from Texas Instruments; V2, Sony IMX219 with 8 megapixels, referring to Figure 3.5. Both versions have the NoIR options that the IR blocker is removed for low light sensing. As previous shown in Figure 2.24, the backside illuminated CMOS sensors have a Bayer colour filter array (BCFA) deposited onto the pixel layer, from which colour mosaic is formed. The BCFA allows red, green and blue banded spectral radiations reach the pixel. Depending on their respective spectral response, certain amount of photons will be collected by the pixels. The mechanism is illustrated by Figure 3.6 top panel. The mosaic image will processed by pixel colour interpolation to generate normal looking images, as shown in Figure 3.6 bottom left. If the BCFA is removed, the sensor will be turned into a monochrome camera collecting photons with respect to the complete sensor spectral sensitivity range in all pixels, as illustrated in Figure 3.6 bottom right.

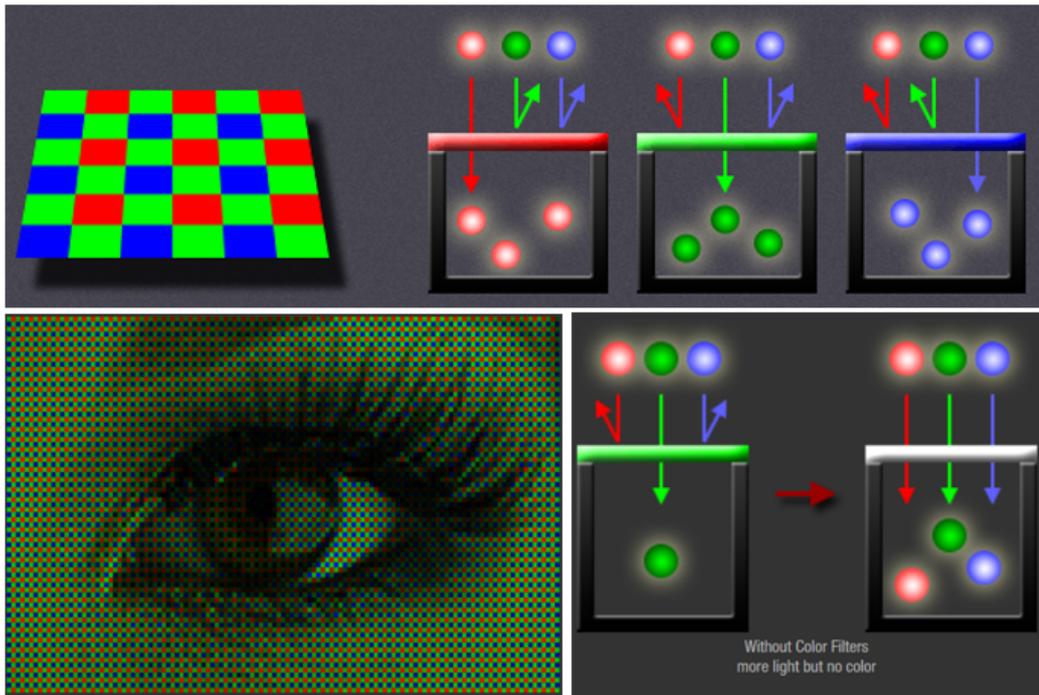


Figure 3.6: How Bayer colour filter array works. Top: Bayer colour filter selective collect photons; bottom left: raw unprocessed colour mosaic image; bottom right: photons collection without the colour filter [128].

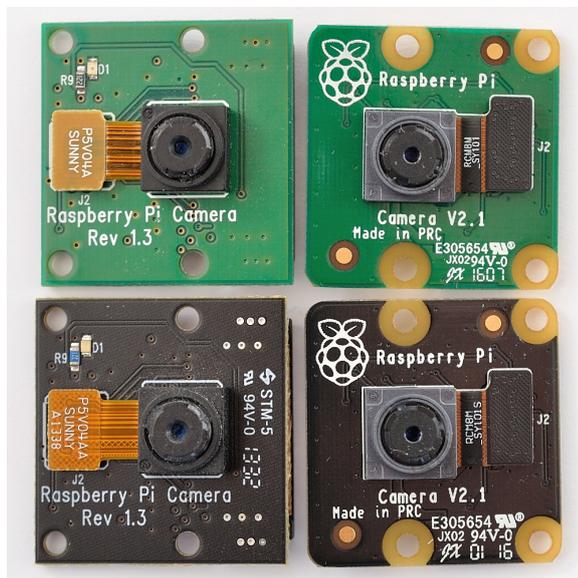


Figure 3.5: Pi cameras V1 and V2. Black ones are NoIR version. Sensor: V1, Omnivision5647, 5 mega pixels; V2, Sony IMX219 8 mega pixels

The NoIR Pi cameras have the IR blocker removed but the BCFA is still in place.

This will produce false colours because the BCFA has certain amount of NIR quantum efficiency for all three colours. The extra NIR photons blend into the pixels of all three coloured pixels. The amount of leaked NIR is determined by their respective quantum efficiencies for all three colour filters. The false colouration due to the NIR leakage is illustrated in Figure 3.7. As previously explained, the strong NIR scattering from green leaves of vegetation introduce more false colour due to the NIR leakage comparing with other objects. This effect will be further explained in subsequent sections.



Figure 3.7: From top to bottom: V2, visible; V1, visible, V1, visible+NIR; V2, visible+NIR [129].

3.3 Turning Pi camera into a monochrome sensor

The Pi camera V1 OmniVision 5647 from Texas Instruments can be modified to be a monochrome sensor by chemically removing the BCFA. As shown in Figure 3.8, the modification enables the use of Pi camera sensor for spectral instrumentation as the any desirable optical filters within its sensitivity range can be applied.

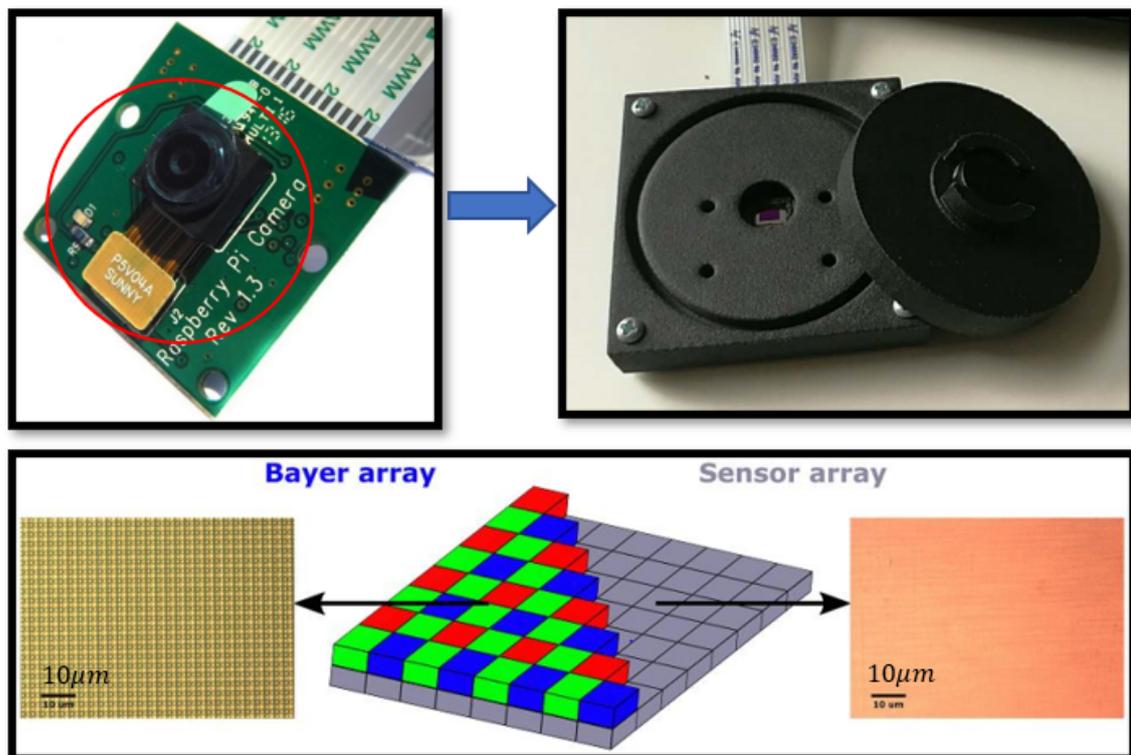


Figure 3.8: Removal of BCFA of Pi camera V1. Top left: the sensor module with integrated lens and IR blocker is highted; top right: de-Bayered sensor with bespoke lens system; bottom: illustration of BCFA removal.

The process of removing the BCFA takes six steps: firstly, using a knife to remove the fitted lens to expose the sensor; then put the module into a bath of photoresist remover (*EKC 830TM*) at heated temperature of about 100 °C until the bulk of the colour filter falls off; Next, a second bath is needed to further remove the residue; in final three stages, the module will go through consecutive baths using butyl ethanoate,

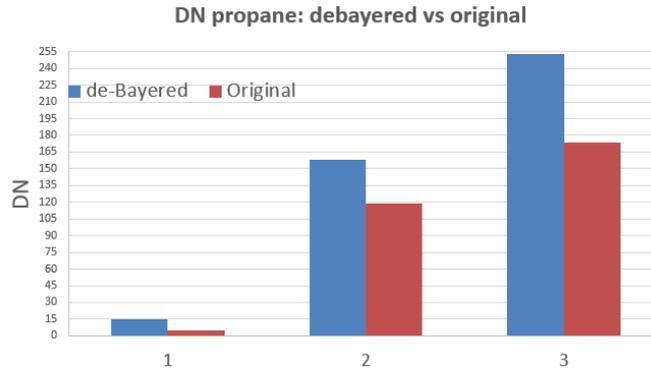


Figure 3.9: Pixel value (digital count) comparison of original and de-Bayered V1 Pi camera. (1): 310 nm narrowband filter; (2): 430 nm narrowband filter; (3): 515 nm narrowband filter.

acetone and isopropyl alcohol [130], [131]. The case was then designed and 3D-printed to mount the camera module with 9-mm focal length focusing lens with high spectral transmission in 200-2000 nm. The lens cap can be screwed to adjust the focus.

Without the BCFA, the sensor have more quantum efficiency as the extra layer of blockage is removed. Narrowband optical filters were applied for comparison by capturing premixed methane flames, as shown in Figure 3.9. In addition, the linearity tests have been performed to check the sensor performance. As illustrated in Figure 3.9, the sensor response is very linear with respect to shutter speed and ISO sensitivity, which ensures the reliability of spectral instrumentation.

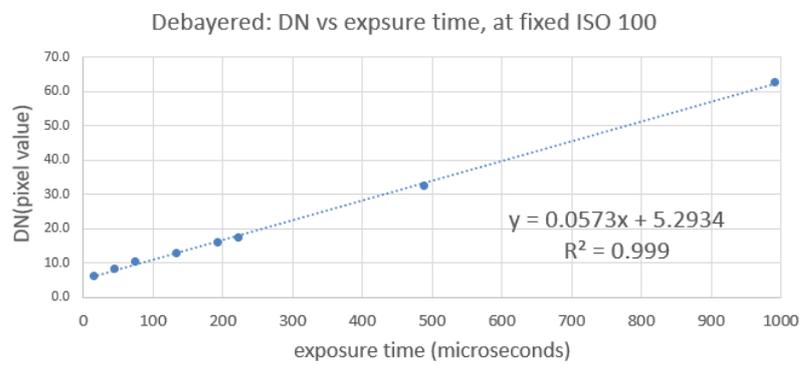
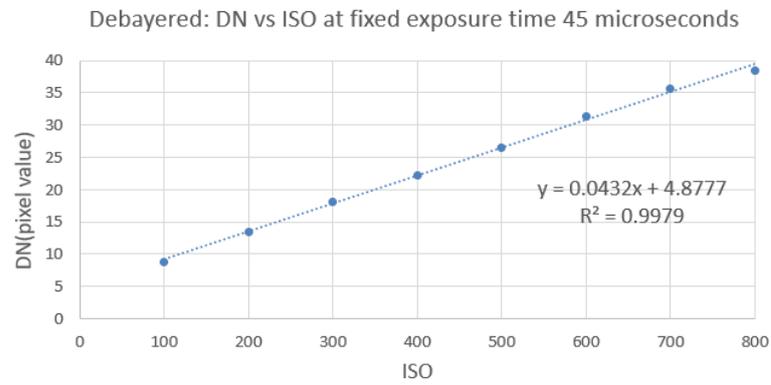


Figure 3.10: de-Bayered Pi camera linearity tests. Top: w.r.t sensitivity; bottom: w.r.t shutter speed.

3.4 Advantages of NIR for fire detection

NIR spectrum seemed to have been much less favoured comparing to other parts of the IR spectrum. In pharmaceutical, food and biomedical industries NIR spectral absorption was the latest adopted technique comparing to that of Mid-IR and Long-IR. It is somehow more overlooked in the field of video fire detection.

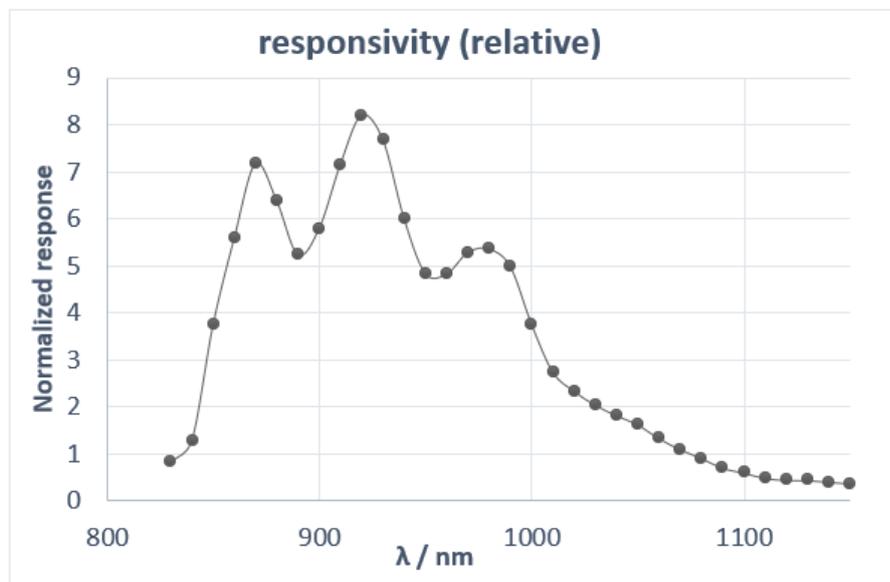


Figure 3.11: De-Bayered Pi camera spectral response, tested using a monochromator.

The NIR fire images have much higher contrast than those from of visible images. This is a combined outcome from three sources: the nature of Si-based CMOS image sensors, the NIR spectral properties in atmosphere and the spectral emission of typical flames. Referring to Figure 3.1, Si-based sensors are more sensitive to NIR radiation than to visible lights. In addition, the spectral response of the de-Bayered Pi camera module was measured using a monochromator, revealing its higher sensitivity in NIR spectrum up to 1100 nm, as provided in Figure 3.11. Because NIR sensor is not able to detect thermal radiations from objects at ambient temperature, the captured background comes from reflected NIR radiations. In an outdoor situation, the main

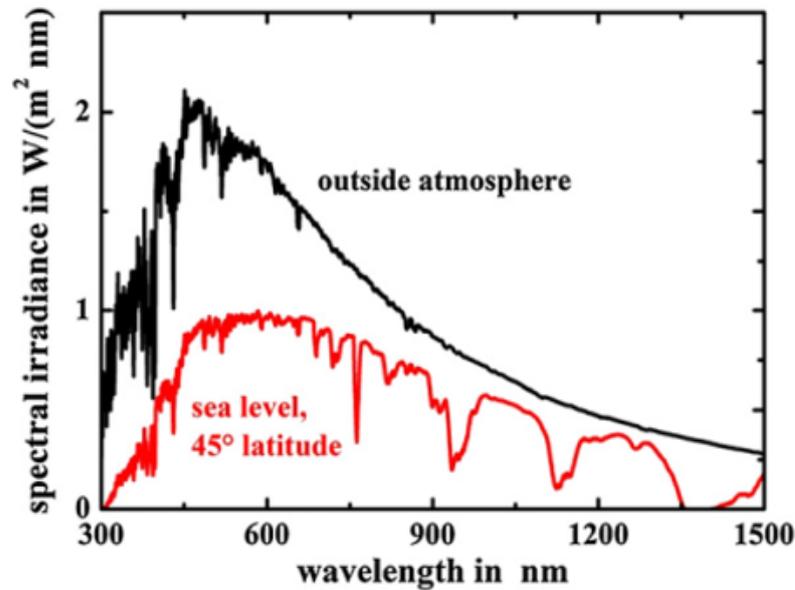


Figure 3.12: Solar spectral irradiance in $W/(m^2nm)$, The curve shows the NIR radiation is weaker than visible radiation from the sun [132].

source is sunlight. While in an indoor environment, the NIR radiation from lighting is much weaker comparing to sunlight, especially if LED and energy efficient lighting are used. Meanwhile, typical flames have stronger radiation in NIR spectrum than in visible, as shown in Figure 2.5.

If the exposure is properly controlled, the background can be suppressed while the flame is still very clearly imaged with enhanced texture comparing to visible images. This property helps to improve the region of interest detection as one of the detection algorithm pipeline. Not only does the NIR camera produce more contrast image, it is able to detect thermal radiation of an object at temperature as low as $\approx 375^\circ C$. The sample images of smouldering charcoals and a flame in Figure 3.13 illustrates the effect.

The flame image in Figure 3.13 was captured on a cloudy day which provides very weak ambient NIR reflection. To further illustrate the high contrast image with strong NIR reflections as disturbance in the surroundings, the test was established



Figure 3.13: NIR and visible image comparisons: smouldering charcoals and a flame. Left: NIR image; Right: visible light image. The flame images were captured on a cloudy day. A cut-on filter at 950 nm was applied to the NIR sensor.

on a sunny afternoon at 3PM during the summer when the sunlight is intense. It can be observed that the fire region is much brighter, even with most intense of sunlight and NIR scattering from the vegetation. From the histograms, NIR scattering from vegetation gives intensity ranging approximately from 50 to 100. The flame region pixels intensities have a range of 100 to 250 with almost no saturated pixels, therefore preserves intensity spatial variation that corresponds to image texture. Moreover, Figure 3.16 shows how simple threshold technique can be very effective in ROI extraction. Furthermore, as shown in Table 3.15, the NIR has better atmosphere spectral transmission comparing to longer wavelengths where CO_2 and H_2O have strong absorption bands. Better atmosphere transmission make it a promising choice for longer distance fire detection. Besides, less attenuation could potentially increase the robustness of the algorithms as the data remains less deviated from its usual distribution.

There are many applications of NIR spectrum for fire detection in remote sensing through satellite imaging, where NIR is part of the hyper-spectral data used for analysis. However, the use of NIR for video fire detection in large indoor space or public open space is very scarce in literature. The low-cost of CMOS sensors with its certain thermal capability comparing to mid-IR and long-IR make it more appealing, notwithstanding it is very underdeveloped. One of the early uses of NIR for indoor scenarios was to detect fire in shipment compartments [2]. The sensitivity of CCD cameras in spectral range 700-900nm was exploited by applying a long pass filter. The triggering criterion is simply based on comparing the sum of pixel intensities of current frame to a reference value. Despite of its lack of sophistication in the detection algorithm, fairly good results were achieved because there were very few disturbances in such confined space. It also reported that the NIR flame images have much better

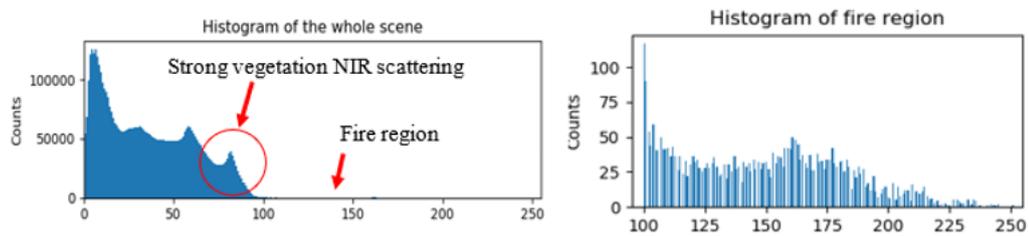


Figure 3.14: NIR image scene test with 950 nm cut-on filter. Image was taken at 3PM on a summer day with intense sunlight. Bottom left: histogram of the whole scene; bottom right: histogram of the fire region. The exposure time was $551 \mu s$ at ISO 100.

Figure 3.15: Integrated atmosphere IR spectral transmission [133].

Spectral range ($\Delta\lambda$)	0.75–1 μm (NIR)	1–3 μm	3–5 μm	5–8 μm	8–12 μm
Integrated τ_{atm} ($\Delta\lambda$)	> 0.95	0.75	0.78	0.19	0.93

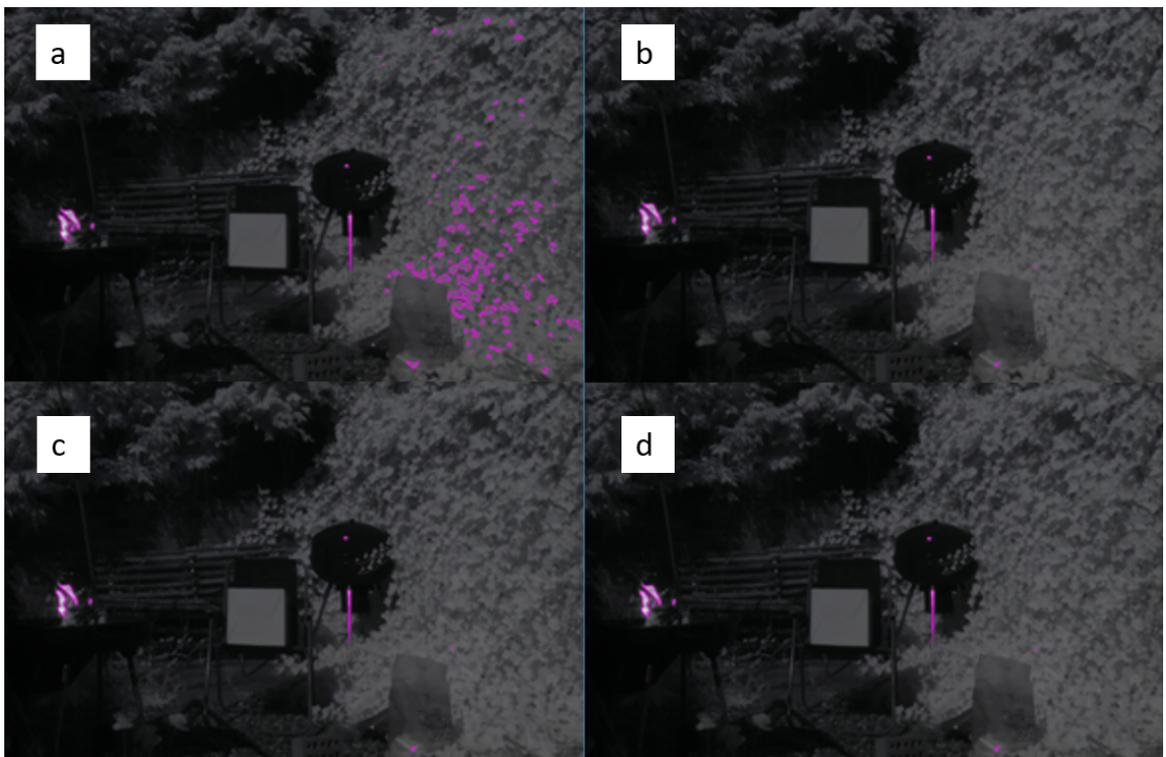


Figure 3.16: Illustration of thresholding for ROI extraction of fire in NIR images, where the highlighted regions are the thresholded regions. (a) $T = 90$; (b) $T = 100$; (c) $T = 110$; (d) $T = 120$. Threshold at 120 is the empirical choice for ROI extraction.

contrast.

[133] tested some features of NIR on hot spots and flames with various materials. A hot spot can be distinguishable at a minimum temperature of 350 Celsius. In addition, a few light sources were compared in spectral range around 950nm, namely magnesium light, yellow and red lamp. Moreover, energy fluctuation in terms of pixel intensity, flame height fluctuation (FFT used for analysis), movement of flame centroid as well as equivalent blackbody temperature were all combined to reach a final decision.

3.4.1 NIR simplifies ROI extraction

As illustrated previously, the use of NIR cameras really simplify the ROI extraction stage. However, considerations and experimentation are required for specific situations. Otsu's method for segmentation can be effective in certain situations where a bi-modal image is usually captured, for example UAV scanning down to the forest [134]. But the method can easily fail in situations like urban environment where reflections and light source disturbances are frequent.

Similar to many other fire detection, the first stage in this work is to extract ROI (region of interest) as potential fire candidates. For Otsu method to work properly, the pixel intensities of the image are assumed to be bi-modal. This imposes limitations to the effectiveness of its application on some scene images containing multiple histogram peaks. In [134], Otsu method seems robust because the UAV hovers over the forest or vegetation which sees a uniform texture in the background. The intensity variation of the leaves would be within a small range, and fire has higher contrast from the background. However, disturbances from the scene can cause the failure of Otsu method, for instance, rock reflection, sunlight reflection from water

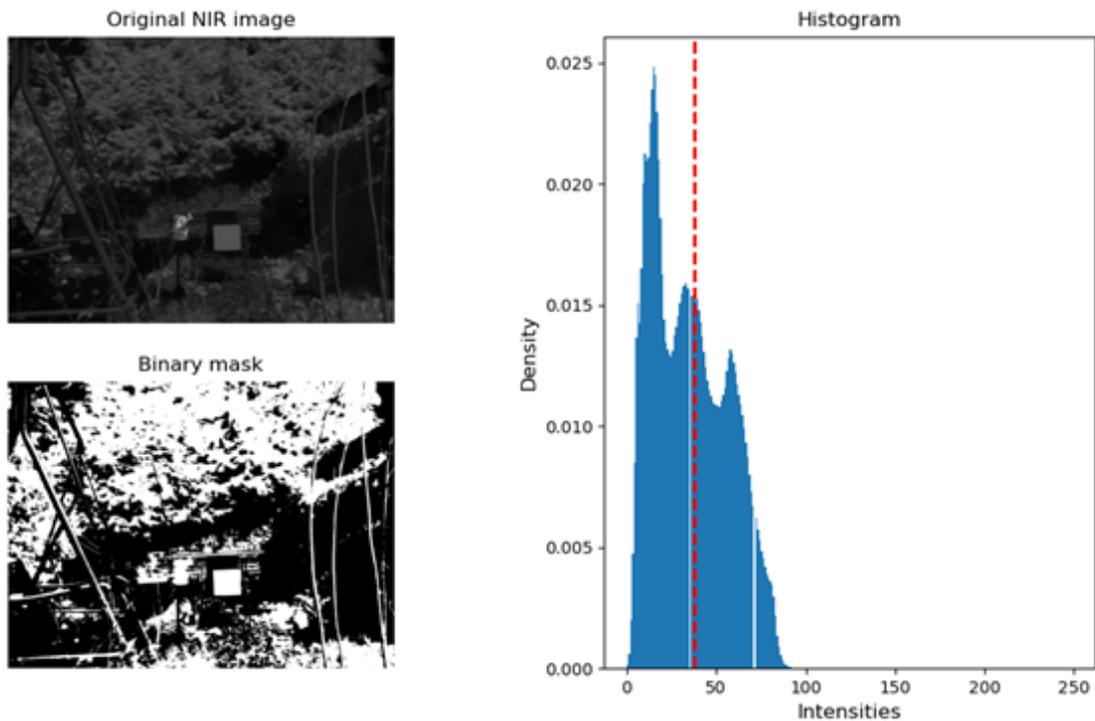


Figure 3.17: Scene test for Otsu's method of thresholding: fail to segment for images having multiple histogram peaks: red line indicates the threshold obtained from Otsu's method.

surface or any other bright objects.

Otsus method for ROI extraction would most likely fail under two conditions: 1. other bright objects or light sources appear in the scene; 2. the flame in the scene occupies a small region. In the first situation, bright non-fire objects or light sources will be extracted together with the flame, which then needs further filtering algorithms. Whereas in the second situation, Otsus method gives a threshold far away from the part corresponding to a flame that the resulting segmentation contains many other objects or just fragments all over the scene. This is illustrated by Figure 3.17, where the red dashed line indicates the threshold generated by Otsus method. In addition, the flame is too small to be manifested in the histogram plot, which locates around intensity of 200.

Clearly, when the NIR images contain multiple peaks, Ostus method is not robust.

Especially for the case of urban open space, there can be reflections from windows, car lights, etc as disturbance. Besides, the diverse texture of the scene in such a situation highly likely gives rise to multiple peaks, which will hinder the performance of Ostus method. For early fire detection, the flame is supposed to be small and may be at distance to the camera. Thus it is very probable that the early fire would only occupy a fraction of the image.

The high contrast NIR flame images were also reported in [135]. The higher contrast between the flame and background in NIR images can be attributed to the two factors: 1. more blackbody radiation in NIR at typical flame temperature than that in visible spectrum; 2. according to Figure 3.18, the modelled spectral irradiance from the sun by SMARTS radiative transfer model reveals that sunlight is more intense in the visible spectrum than in NIR. As a result, the flames have higher pixel values whereas the reflected NIR radiation from the background is much weaker. The contrast of the resultant scene image is greatly improved. This is also confirmed in the quantitative analysis, where the contrast ratio of NIR images is between 1.28 and 2.40 comparing with visible images between 0.86 and 1.58 [135].

According to the literature, three choices for thresholding are, Ostus method, fixed threshold and adaptive threshold. It is worth noting that these methods are greatly affected by the IR camera choice and the camera setting. As previously mentioned, the most popular IR cameras for fire detection in literature are operating in long wavelength band of 7-14 μm . The images from both LWIR and NIR cameras are significantly different. Because LWIR sensors capture the thermal radiation emitted by the objects themselves, it has a typical measurable temperature range of -25 to 300 Celsius. By contrast, the NIR spectrum is close to visible light that the captured image remains relatively intuitive. Same as the visible light, the NIR of the scene

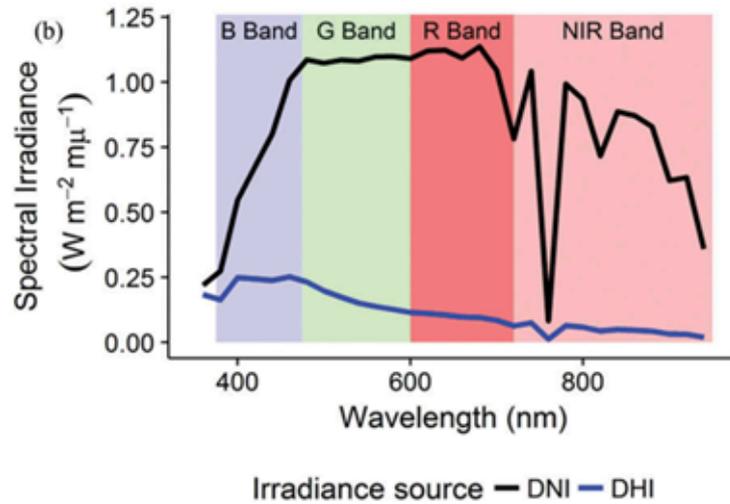


Figure 3.18: Modelled spectral irradiance from the sun by SMARTS radiant transfer model [135].

comes from diffused sunlight. The most use of Ostus method in literature couples with the deployment of LWIR cameras. It may be robust but requires extensive scenario tests for verification. Adaptive thresholding for ROI extraction is also an option but with empirical formulation. [82] uses the mean and standard deviation of the pixel values of the whole image, based on which a set of equations was formulated to generate a threshold value. These equations contain empirically determined constants and coefficients.

Some issues were reported in literature in terms of LWIR camera features: the fire imaged by LWIR sensors are always saturated because of its intense radiation within the sensitive spectrum; moreover, if the sensor gain is not adjustable, human body temperature can even saturate the pixels. Therefore, the spatial intensity distribution of fire pixels is lost, which means less information to be exploited as flame features [53].

3.4.2 Using NIR false colour for fire detection

A purplish pinkish colour was introduced if the IR blocker is removed because NIR radiation blends into all three colour pixels. This false colour is particularly advantageous in fire detection. For example, this can be used to distinguish fake fire (fire advertisement on display, posters) from really fires. The fire from displays are emulated by screen pixels emitting visible lights, which will remain its normal reddish colour. Similarly, the pigments in advertisement posters reflect certain visible wavelengths to mimic the look of a fire. The false colour of a flame is demonstrated in Figure 3.19.

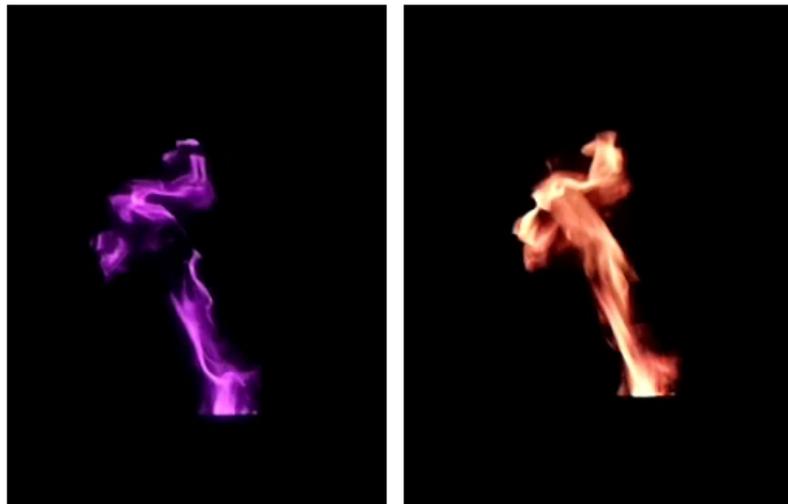


Figure 3.19: NIR false colour vs common visible colour of a fire. Left: A 760 nm cut-on filter was applied to Pi camera; right: normal visible fire

3.5 Summary

Despite the InGaAs (Indium Gallium Arsenide) detectors' high-speed and low noise quality, its cost prevents its broad use for general safety surveillance. In contrast, silicon-based CMOS image sensor, as used in this research, is also sensitive to NIR. This type of sensor is dominant in the photography market that the low cost is

propelled by mass production. To avoid NIR leakage to the pixels, manufacturers integrate IR blockers to the lens system allowing only visible lights to reach the image sensor. By removing the IR blocker and Bayer filter, it becomes a low-cost NIR image sensor with relative high speed. Moreover, comparing to the mid and long infrared sensors, the silicon-based sensors have much higher spatial resolution and framerate. In terms of intelligent fire detection, high resolution is crucial for stereo localization of the fire and mapping and planning for robot firefighters. It is also capable of detecting non-luminous flame, such as alcohol and hydrogen flame, as CO_2 and H_2O are the main combustion products having spectral emission in NIR. Furthermore, the false colour can be used in combination with a Pi camera limited to NIR spectrum only for fire detection to increase system robustness. The camera system used for fire detection in this research is Pi-based stereo NIR camera system, where one camera without IR blocker for colour detection (Visible + NIR) and the other operates in NIR spectrum only for texture feature extraction.

Chapter 4

Multi-spectral fire detection system

This chapter explores the potential of using low-cost CMOS image sensors for fire detection in combined multiple spectrum. The scene image is captured by a stereo system, with one camera working with visible and NIR spectrum of (400-1100 nm) and another camera with NIR spectrum only (950-1100 nm). The captured stereo image will then be the input of the fire detection system. The system overall structure is demonstrated in Figure 4.1, which contains three stages: ROI extraction stage, texture-based feature extraction and decision making stage.

In the first stage, a threshold was used to roughly locate the flame regions. A two-stage crop was applied to ensure the ROI crops into the object. The crops will then be fed into the second stage, where all the texture based features will be extracted to form the feature vectors. The final stage incorporates a SVM for classification with a GMM colour model for false positive elimination.

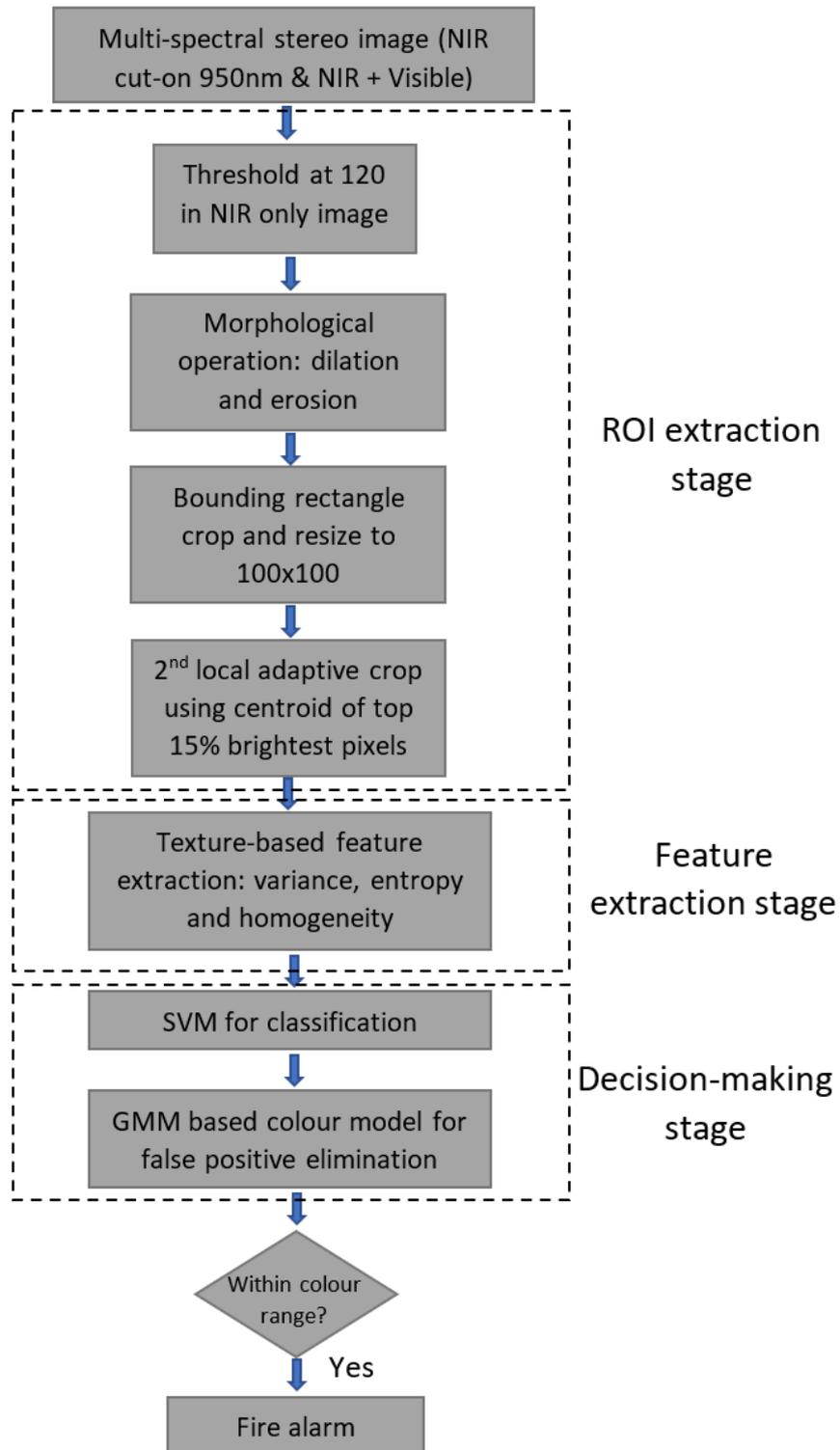


Figure 4.1: Multi-spectral fire detection system: ROI extraction stage, feature extraction stage and decision making stage.

4.1 ROI extraction stage

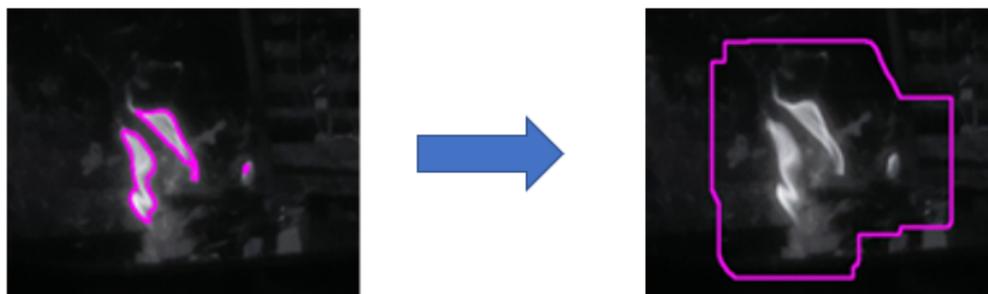


Figure 4.2: Morphological operation to combine blobs into patch.

In the first stage, the main goal is to extract patches of potential fire candidates. The NIR fire image will be threshold at 120. The threshold is empirically determined at fixed camera setting to ensure higher contrast: exposure 551 μs and ISO 100, showed previously in 3.16. After that, a morphological transformation by erosion and then dilation was used to remove the noisy spots and combine the clustered bright regions into one patch. For example, if there is a fire in the scene, the fire region after thresholding contains clusters of blobs as shown in Figure 4.2. The morphological transformation combines the cluster into one patch. A rectangle crop of the patch is used to extract the pixel values as a subset of the image. Dilation and erosion are particular useful in emerging regions and remove very small noisy regions. These two techniques applies a Boolean filter of a chosen size to binary images. The operation can be formulated as in Eq(4.1):

$$\text{Dilation: } g(n) = OR[\mathbf{W}f(\mathbf{m})] \quad (4.1a)$$

$$\text{Erosion: } g(n) = AND[\mathbf{W}f(\mathbf{m})] \quad (4.1b)$$

The $f(\cdot)$ represents an image; \mathbf{W} is the subset size and m is the stride size. In dilation operation, the filter sweeps through the binary image with the defined stride size, where each stride is a convolution. If any pixel within each convolution is 1, all of the pixels will be set 1. The process is the same for erosion with reversed logic. As shown in Figure 4.3, the increase of filter size resulted in increased blob size.

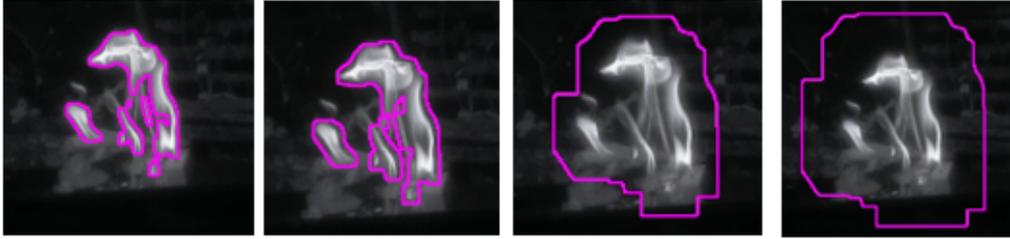


Figure 4.3: Dilation with different filter size, from left to right: 5, 10, 50, 100.

Next, rectangular patch will be cropped using the boundaries of the patch generated in Figure 4.3. According all the scene tests, the generated ROI patch is larger than needed for local texture extraction. In addition, in some cases, the crop would have background around the frame. Therefore, an adaptive crop will be performed again based on the centroid of the pixels having values in the range of $[I_{max} - 10\%, I_{max}]$, as in Eq(4.2). This ensures the second crop is always covering the centre which is usually the brightest region. In addition, the second crop reduced the patch size from 100x100 to 60x60. The cropped region will be used in the next stage for feature extraction.

$$\bar{x} = \frac{\sum_{n=1}^m x_i}{m} \quad \bar{y} = \frac{\sum_{n=1}^m y_i}{m} \quad (4.2)$$

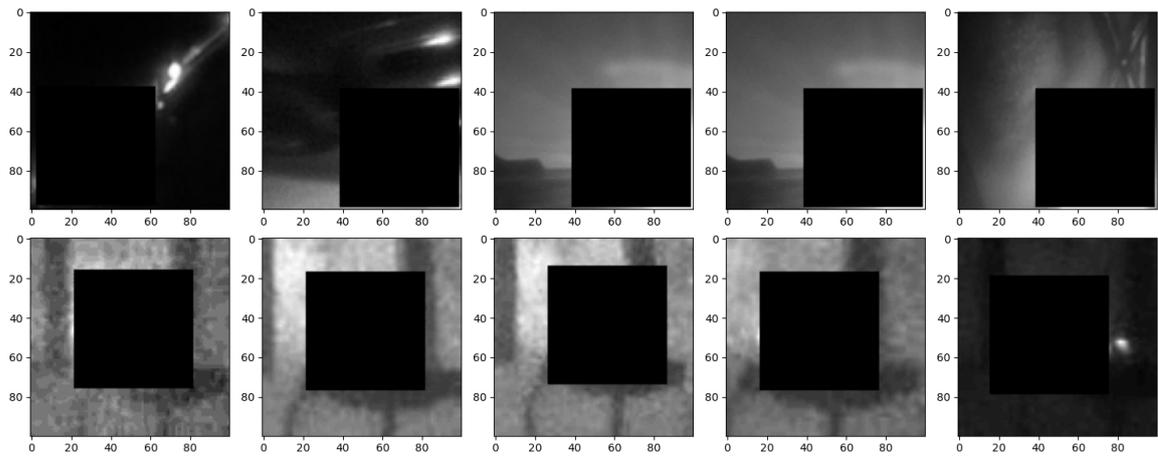


Figure 4.4: Second stage adaptive crop of 60x60. The black patch indicates the cropped area. Top: car light, non-fire reflections from car front grill, windows; bottom: fire on BBQ stand.

4.2 Feature extraction stage

The ability of human perception of an image can be categorized by three fundamental aspects, namely perception of colour (spectral information), texture (tone gradient and spatial distribution) and semantics (meaning or representation of an image). As human are only able to see visible light, many of the image sensors outside such spectrum presents image as gray scale without any mosaic layer to form a colour space. The simplest texture of an image are the edges, where the spatial tone gradient is high. Then to understand the meaning of the image, a comprehensive evaluation of the surrounding, actions, pose, etc, is needed.

In the case of fire detection, texture feature can be beneficial because the intensity of a fire has its unique tone variation and spatial distribution. The tone is simply the variation of shades of intensity, which bears no directional information. Whilst the texture of an image is the spatial distribution of tones [136]. Different objects have distinct image texture, ascribed to their ability of diffuse lights based on the surface micro-structures and materials properties. Some texture can be form by a swarm of the same objects, for instance a grass land. Grass clustered together to have a whole background with distinct texture properties. Moreover, same object appears distinctively imaged in different spectrum in terms of image texture. [137] showed some statistical measures between visible and LWIR image, indicating the lack of texture in LWIR images can help simplify computer vision problems. It is a trivial task for human to recognize the texture in an image and promptly shout out a word to describe it. However, the precise definition of texture by computer still requires thorough mathematical interpretation. Three textural features will be used in this study for fire and non-fire classification, namely homogeneity from GLCM, entropy and variance.

The gray-level co-occurrence matrix (GLCM) method is an effective way of extracting texture features [136]. The GLCM algorithm assumes the textural information of an image can be expressed as a spatial relation. The base of the algorithm is to compute a spatial probability distribution matrices of a gray image. These matrices can then be used to calculate 14 textural features. The generated probability matrix is proportional to the input image. Practically, a tiny subset of an image may suffice. Basically, the core of this algorithm is to generate a relative-frequency matrix $P(i, j; d, a)$, where i and j are the intensities of two neighbouring; d specifies the distance between the pixel pair; a is the angle of orientation along which the occurrence of pixel pairs are counted. The detailed computation can be illustrated in Figure 4.5.

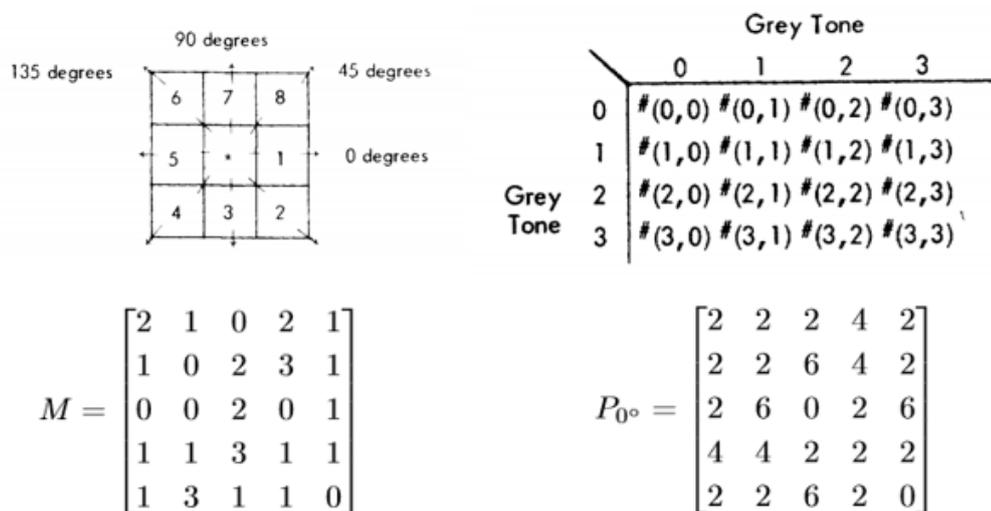


Figure 4.5: Computation of co-occurrence frequency matrices: (a) occurrence counting orientation; (b) obtained frequency matrix; (c) example image M ; (d) frequency matrix obtained by counting horizontally.

The next step is to normalize the frequency matrix by the total number of occurrence of pixels pairs, which then gives a pseudo-probability interpretation matrix. In the example above, every element of P_{0° will be divided by the sum of all

its elements, which is 68. The resultant matrix will thereafter be used to compute texture features. During this computation, each texture measure has its own weight matrix. The final measure is a single number by taking dot product of GLCM and the weight matrix.

The choice of texture used for infrared fire texture description is homogeneity, which will give a high value if the image has less texture or less contrast. It can be observed that in its weight matrix, the weights away from the diagonal decays exponentially. The weight matrix biases the diagonal since the diagonal in GLCM matrix is the frequency of pixel pairs with the same intensity value. Because the camera is set to take advantage of the high contrast of fire to the background, most of the background will be very dim, except for solar reflections from glossy surfaces, car lights, traffic lights, etc. These non-fire objects will form the non-fire class.

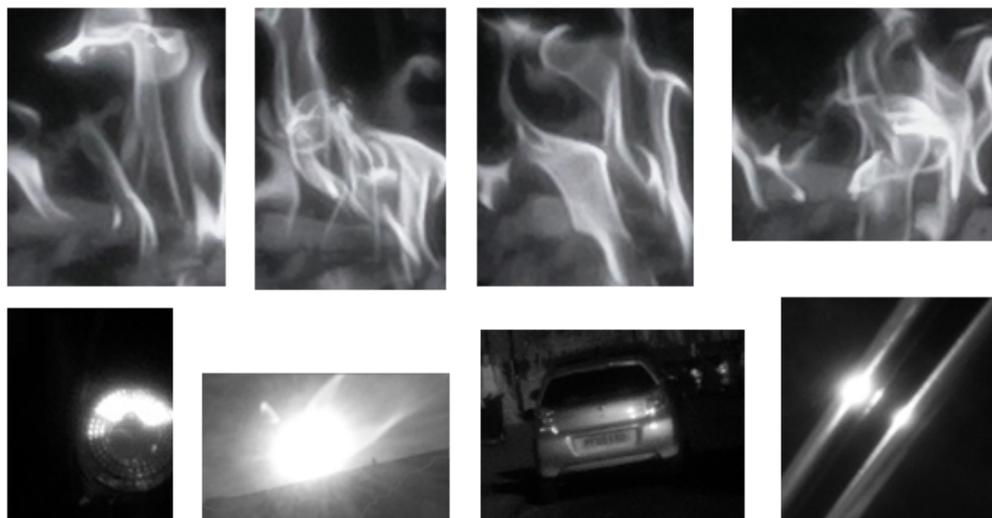


Figure 4.6: Illustration of texture difference between fire and non-fire objects. Top: fire; bottom (left to right): car light, sun reflection, reflection from car body, reflection from glossy bars.

Almost all uses of LWIR (7-14 μm) for fire detection through literature reported the saturation of fire images, in which case the texture information is no longer

available. Therefore, the shape of outline of the fire patch was used to extract some geometrical measures to generate features. Figure 4.6 is a comparison between fire and non-fire cropped bright objects at this fixed camera setting. It is obvious that the flame images have more texture, such as more edges and steaks. While the non-fire examples mostly come from light sources and reflections that are smooth and even. Many attempts have been carried out to find the optimal parameters for GLCM textures. As shown in Figure 4.7, the orientation to get co-occurrence matrix does not affect the homogeneity feature. However, the distance d between two pixel pairs greatly affects the difference of homogeneity between fire and non-fire objects, as illustrated in Figure 4.8. The homogeneity decreases as the distance increase for non-fire objects, while the fire is not much affected.

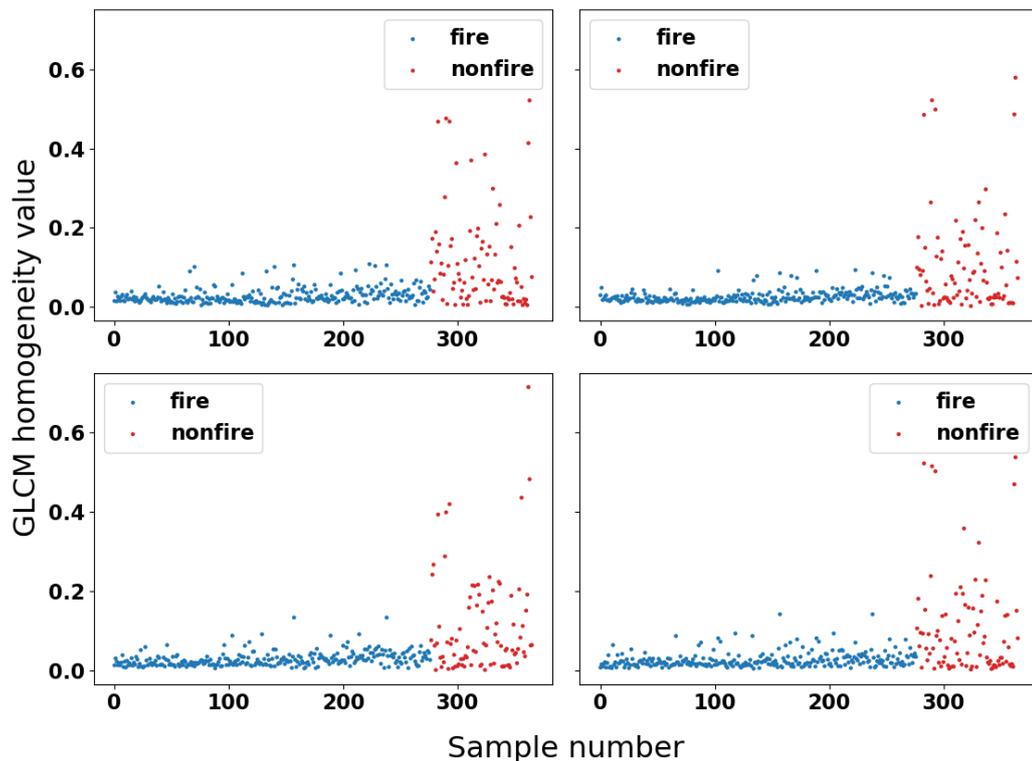


Figure 4.7: GLCM texture homogeneity vs sample number at different orientations for fire (blue) and non-fire objects (red). Top: 0° , 45° ; bottom: 90° , 135° .

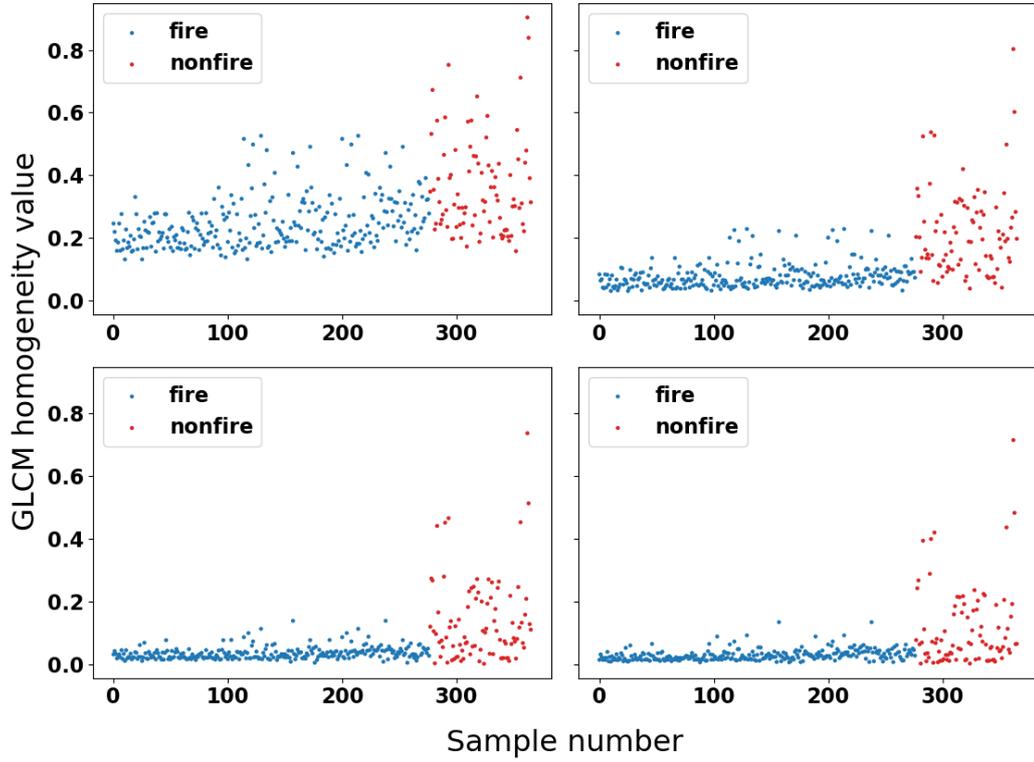


Figure 4.8: GLCM texture homogeneity vs sample number at different pixel pair distances for fire (blue) and non-fire objects (red). Top: $d=1$, $d=5$; bottom: $d=10$, $d=20$.

The concept of entropy is commonly known from thermodynamics. From image processing point of view, it can be regarded as the average amount of information required to describe the state of a random variable [138]. The entropy is calculated by Eq(4.3), where $p(x)$ is the probability of the random variable x . The summation is over the image patch of n pixels. For an image, the pixel intensity can be thought as the random variable $x \in [0, 255]$.

$$H[x] = - \sum_{i=1}^n p(x) \log_2 p(x) \quad (4.3)$$

Figure 4.9 shows the entropy image of a fire and a non-fire object. It can be observed that the regions having higher entropy values hold more variations in intensity gradient and spatial distribution. In the non-fire case here, the image is a

sun reflection from a window. The bright region has saturated pixels which makes the texture flat. More specifically, the entropy of pixels is computed within the available 8x8 patches within the image. For example, in an 8 by 8 matrix, the appearance of unique pixel values will be counted. Then a probability $p(x)$ for each value can be obtained by normalizing with respect to the total number of pixels in the patch, 64 in this case. After that, the entropy of each pixel can be calculated by applying Eq(4.3). In the case of flat texture within the patch, such as the bright region, the probability $p(x)$ of the bright pixel value is close to one because there are many pixels with the same intensity. Then the value of $\log_2 p(x)$ will come close to 0 and its entropy will have a small positive value. In contrast, in the fire image, more texture means a wide range of pixel values. For each pixel value, its $p(x)$ will be close to zero. Then the logarithm term will be close to $-\infty$. As a result, its entropy will be large. From probability point of view, entropy can be through as a measure of surprise: if one was told that a highly unlikely event has just occurred, the information conveyed is more than that one was told an ubiquitous event just happened [138].

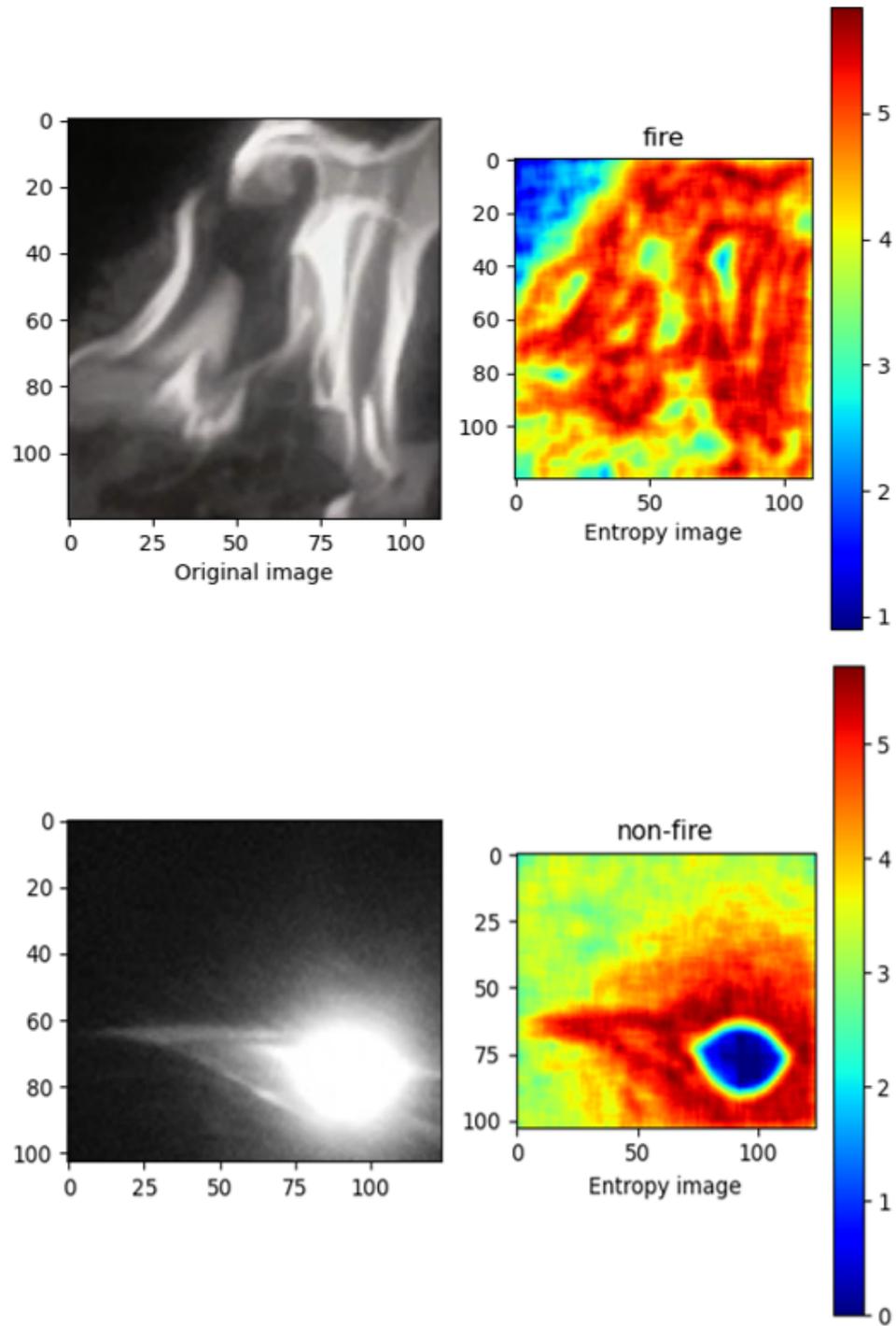


Figure 4.9: Entropy image of fire and non-fire objects obtained by down-sampling with 8x8 convolution. Entropy is proportionate to spatial variation.

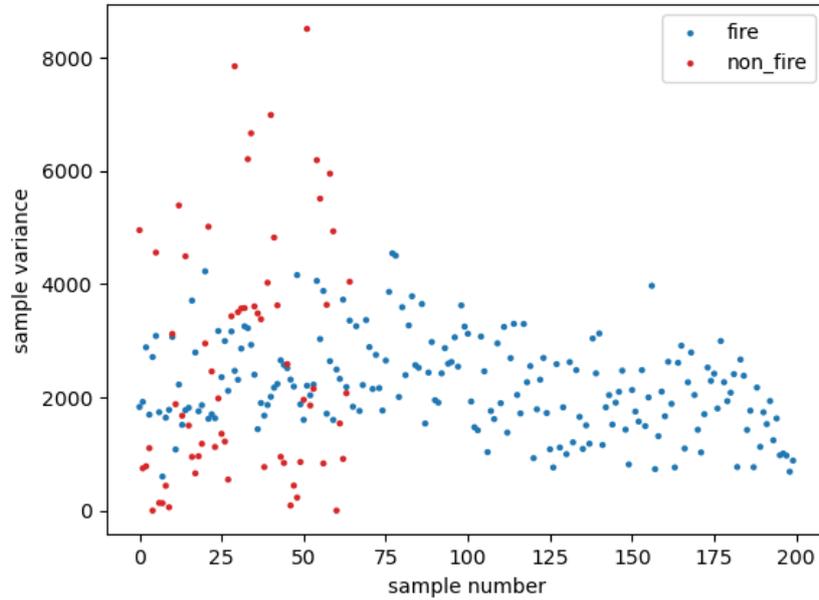


Figure 4.10: Variance plot of fire and non-fire objects: fire (blue); non-fire (red).

The variance is also an effective measure of the difference between fire and non-fire objects. Fire as a dynamic combustion phenomenon constantly interacts with the surrounding cooler airflow, gives a relative steady distribution of pixel intensities. However, for non-fire objects, the variance is bipolar depending on the type of bright objects. For example, if it is sun reflection or car light, the uniform intensity across the crop will give very small variance; whereas if some background got cropped, the variance can be very large. Figure 4.10 illustrates the variance of some fire and non-fire examples.

4.3 Decision-making stage

4.3.1 Classification using SVM based on extracted features

Through previous stages, the potential fire candidates in the scene are produced and filtered. Features were extracted from the remaining fire candidates. The 3 features, namely pixel value variance, GLCM homogeneity and entropy, forms a 3D feature space. A support vector machine is trained for classification. The training dataset consists of 282 fire examples and 89 non-fire examples. Because of the scarcity of NIR spectrum for fire detection in literature, there were no available datasets to the best of the authors knowledge. The whole dataset used for the classifier is obtained in-house. The fire examples were collected by burning common kindling wood. The fire examples contain flames with different dispersion because of the ambient air flow and combustion intensity. The non-fire examples were collected by shooting at random scenes in urban areas. To cover a wider range of non-fire candidates that could pass through the data processing pipeline before feeding into the classifier, some bright objects were deliberately captured to test the classification robustness. Due to the design of this fire detection system, only very bright objects in NIR spectrum can be extracted for subsequent processing. Therefore, through the pipeline, most non-fire objects were any kind of light source having strong emission or reflection in NIR, for instance car light or sun reflections from glossy surfaces, such as from windows, ground, car, metal bars, etc.

The support vector machine is trained with radial basis kernel function. Different choices of kernel functions have been attempted to compare the classification performance. Using radial basis kernel yields slightly better classification comparing to linear and polynomial kernels. Referring to Figure 4.11, the visualization of the training data space is roughly linearly separable. It is probable that if more data

can be generated, the feature distribution in the feature space may become more intertwined. Therefore, the decision boundary will become more nonlinear.

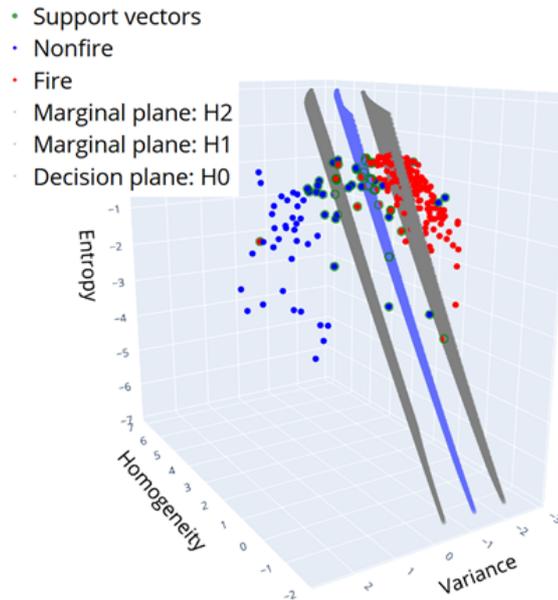


Figure 4.11: The trained SVM in feature space. Green: support vectors; blue dots: non-fire examples; red dots: fire examples; gray planes: marginal planes H_1 , H_2 ; blue planes: decision plane H_0

Referring to Equation 2.9, the regularization term contains C and ζ . ζ is a slacking variable to tolerate some examples that does not strictly satisfy $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$, whereas C controls the degree of tolerance. In addition, γ in RBF controls the degree of non-linearity. More specifically, the parameter C controls the trade-off between classification rate and the margin length. At higher values of C , the marginal planes will stay close to the decision hyperplane H_0 , while they will be far part at lower values. This is illustrated in Figure 4.12. The parameter C shifts the decision plane as its value changes. It is apparent that the classification accuracy in the latter two cases is higher than the first case. In addition, smaller C will lead to more support vectors to construct the decision plane.

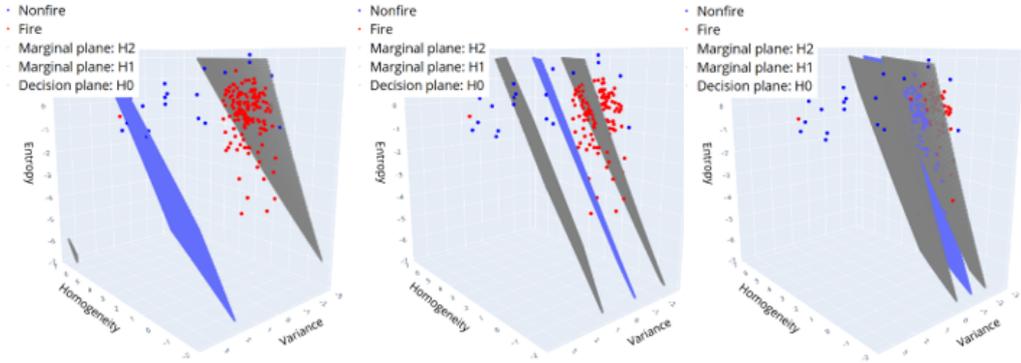


Figure 4.12: Effect of C on SVM decision plane (left to right): $C = 1$, $C = 10$, $C = 100$

The decision hyperplane H_0 is very sensitive to parameter γ in RBF. In 3D space, γ controls the radius of the volume of similarity of the support vectors. Referring to the SVM optimization formulation, the dot product is just replaced by a radial basis function. The kernel function is a similarity measure based on the Euclidean distance between a support vector and another datapoint in the set. The value of γ is disproportionate to the volume of influence. This is illustrated in Figure 4.12. Large γ allow the model to try to capture the volume of the data cluster. As seen in the figure, larger γ gives smaller radius of curvature of the decision boundary, and vice versa. At smaller values of γ , the decision function is similar to that with linear kernel.

The dataset used for training and testing were gathered at four distances from the camera, namely 5, 10, 15, 20 meters with the fire roughly the same size. Based on the system design, after the first stage of ROI extraction, the greater the distance, the smaller the fire proportion in the crop. Then the second ROI always crops within the first around the brightest region. Hence the final extracted ROI fed into the classifier contains similar proportion at different distances. The main difference, however, is the smear effect as the fire becomes smaller in the scene. Less pixel is used to represent

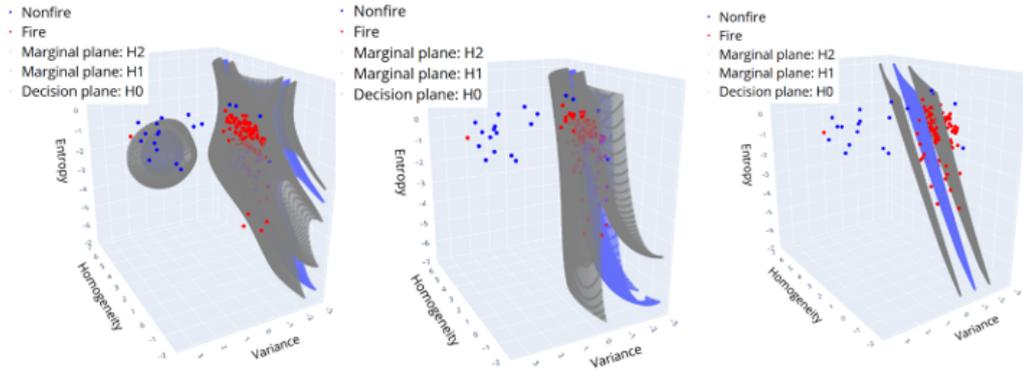


Figure 4.13: Effect of γ in RBF on SVM decision plane (left to right): $\gamma = 0.05$, $\gamma = 0.01$, $\gamma = 0.001$

the flames; therefore, the texture information may be reduced.

Figure 4.14 illustrates the misclassified examples from training. According to the confusion matrix, totally 14 examples are misclassified, where 13 cases are false positive and only one case for false negative. Even judged by human perception, it is difficult to tell that it is a fire. In a real scenario, fire is always irregular because of buoyancy and ambient air flow. Especially in a windy situation, the flame can be frequently diluted rendering them to be dimmer. Fortunately, this is not a problem given the system processing rate is high. Given a short period, the flame can regain its temperature and brightness, which will be detected again by the system. The emphasis on reducing the false positive alarms while maintaining an effective detection, is the key for a robust fire detection system. It should be noted that the confusion matrix measures the performance of the SVM, not the whole system. The subsequent colour filtering algorithm will check if those predicted fire examples are real fires.

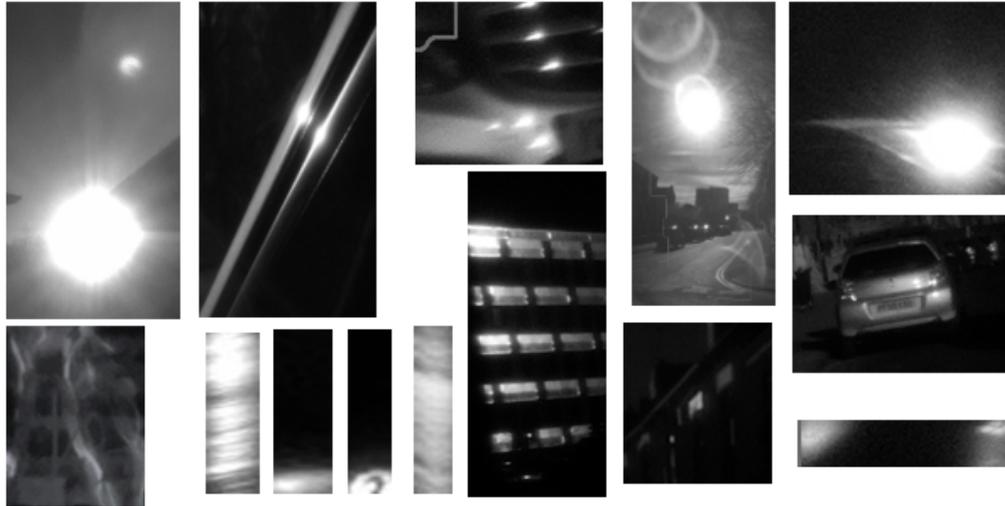


Figure 4.14: Misclassified examples in training: only one false negative at bottom left.

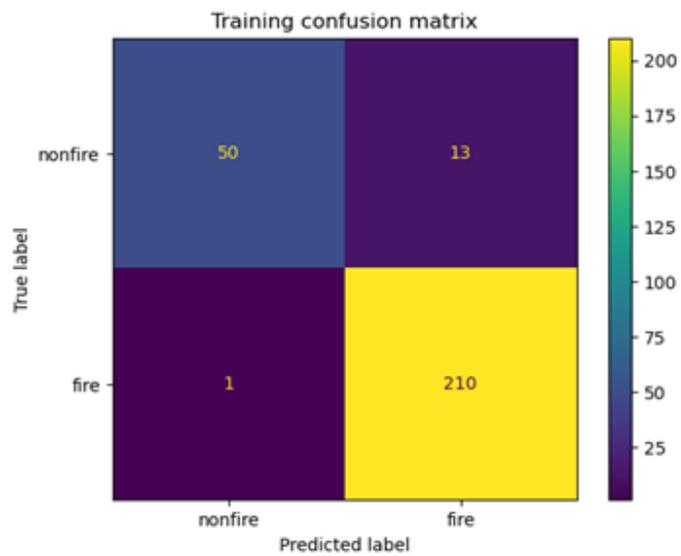


Figure 4.15: Training confusion matrix.

Table 4.1: Test set detection rates at each distance.

Case	Detection rate (precision)	False negative rate
5m	98.5%	20%
10m	87.9%	15%
15m	88.1%	15%
20m	93.8%	21%

Overall, the test shows good result. The detection rates in the test set are shown in Table 4.1. Generally, the detection rate does drop slightly. The smaller proportion of occupation does not seem to affect the performance because of the two-stage cropping algorithm. It is observed that those mis-detected fire examples all appear dimmer and with lower contrast than the correctly detected examples. This makes sense because when the fire is disturbed by the wind, more mixing and cooling happened in the flow field, giving reduced texture. The test false positive rate for each test case is around 20% at this stage. This is probably due to the lack of dataset. If more test examples are made available, the rate is expected to drop. Secondly, the colour model in the next stage will filter out false positives to give a much better overall false positive rate. The results from all test cases will be shown in Figure 4.16, Figure 4.17, Figure 4.18 and Figure 4.19, where the confusion matrix and misclassified examples will be demonstrated.

4.3.2 GMM colour model to eliminate false positives

In real-time systems, the data throughput can reach tens of images per second. Even the algorithm does not have a very high detection rate, the system may still detect the subject in time with such large amount of data. Therefore, the robustness of a fire detection system depends on the rejection of false positives. In this section, a Gaussian mixture model with Dirichlet process is used to model the unique colour

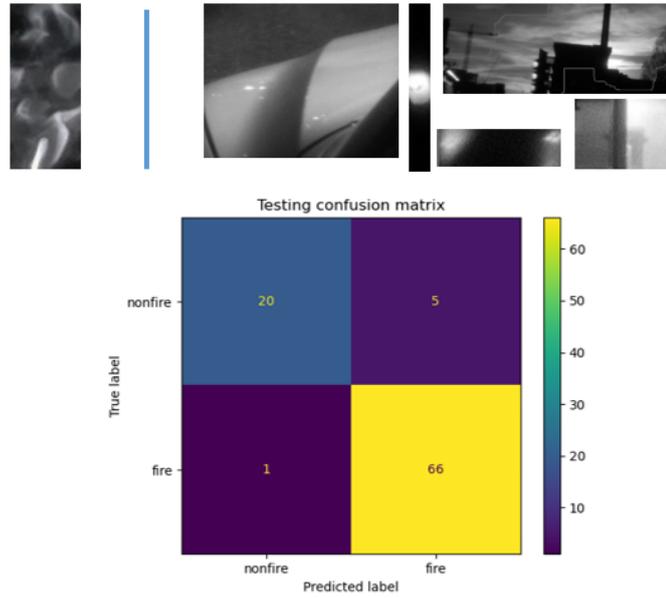


Figure 4.16: 5m test case. top left: false negative; top right: false positive; bottom: confusion matrix

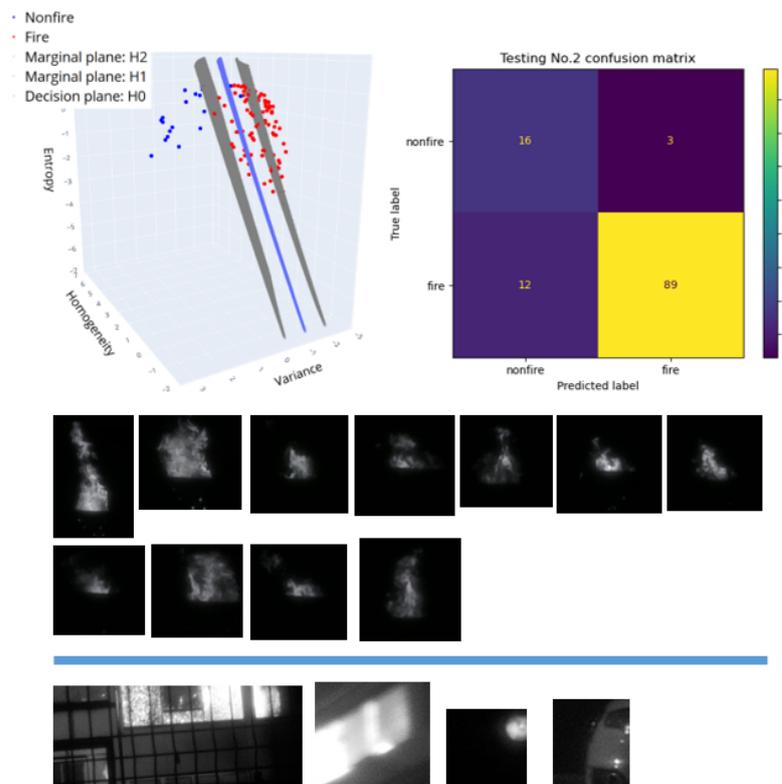


Figure 4.17: 10m test case. Top left: trained SVM on test examples; top right: confusion matrix; bottom: misclassified examples.

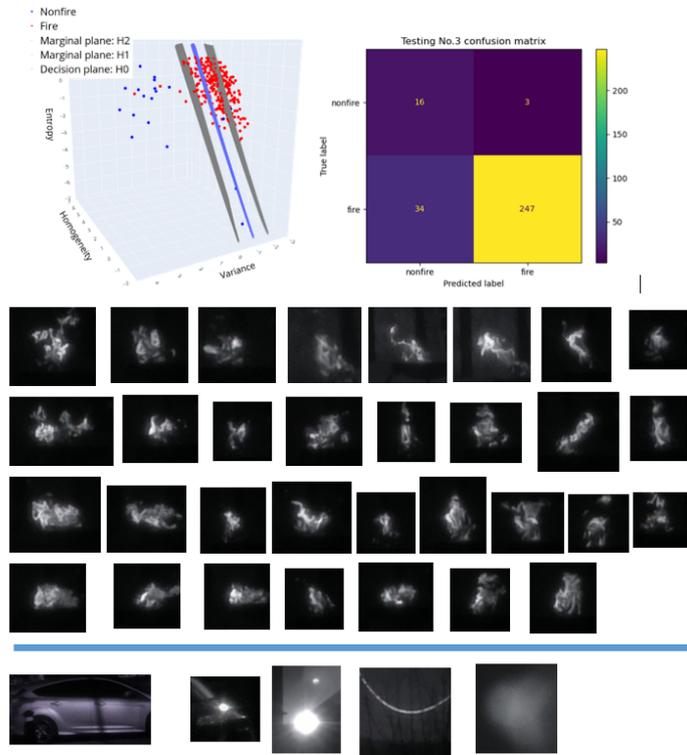


Figure 4.18: 15m test case. Top left: trained SVM on test examples; top right: confusion matrix; bottom: misclassified examples.

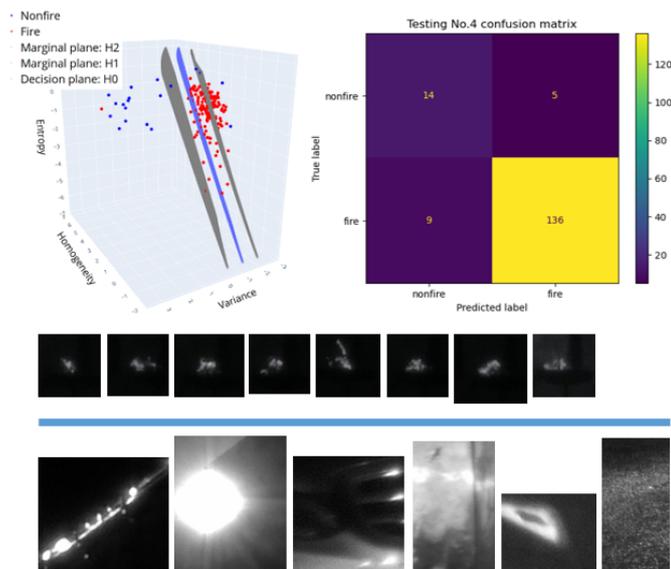


Figure 4.19: 20m test case. Top left: trained SVM on test examples; top right: confusion matrix; bottom: misclassified examples.

of flame captured by Pi NoIR camera, which combines visible and NIR spectrum sensitivity with Bayer CFA. To reach this choice, a series of study was established to investigate the effect of different NIR cut-on filters and exposure on flame colour distribution.

The colour of a flame mostly comes from the radiation of soot formed in the flame. The soot is usually regarded as blackbody, which has a radiation curve similar to Planck's blackbody radiation curve. The visible part of the soot radiation perceived by human eyes are orange to reddish colour. Common flames usually appear to be such colour because of the presence of soot in the flame. This type of flame falls in the category of diffusion flames, where the fuel is separated from the oxidizer before entering the reaction region [139]. More specifically, the substances are heated to a temperature for the material to decompose into combustibles. Most of the fuels contains carbon, hydrogen and oxygen. Others, such as plastics, may be composed of nitrogen, chlorine and fluorine [1]. In real scenarios, the spontaneous flames can happen in two situations: direct flaming combustion and smouldering combustion which can develop into flaming combustion [1]. The decomposed combustible gases will then be transported into the reaction zone through diffusion. The phenomenon of diffusion is driven by the gas concentration difference, which can be described as Fick's Law [139]. As shown in Figure 4.20, the concentration of the gas and oxygen is very low comparing to where they were transported from. Then the gas and oxygen will be transported in the direction of high concentration to low concentration, so that to feed the reaction. Whereas the combustion product has higher concentration at the reaction zone, they will be transported away from the it.

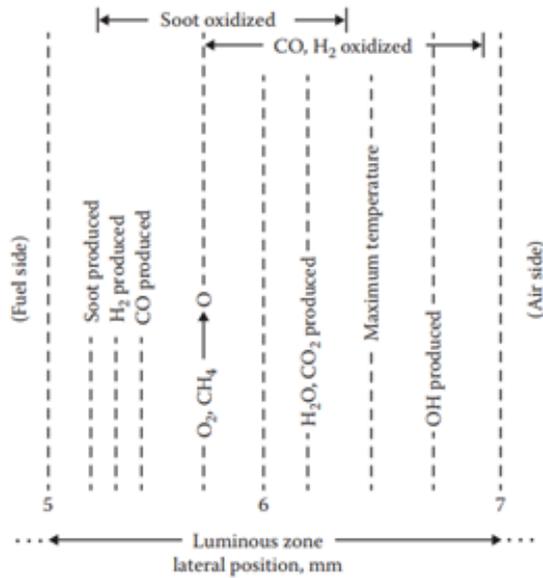


Figure 4.20: Substance concentration in lateral position of a diffusion flame.

4.3.2.1 The cause of false colour in NIR image

The objectively true colour of the flames comes from the physical phenomenon of reaction. However, reproduction of this colour information by means of imaging, can be affected profoundly in terms of spectral response of the hardware, such as the Bayer CFA, the sensor, and other optical components. To be precise, the camera needs to be colour-calibrated with colour compensation algorithms. Because the RGB micro filters are wide bands, which has its respective peak quantum efficiencies, each of the obtained raw image pixels is an amalgamation of photons of a range of wavelength. Therefore, the interpretation of colours is dominated by the properties of Bayer colour filter. Manufacturers integrates an IR blocker to let the sensors only to the visible spectrum (400-700 nm). If the blocker is removed, there can be false colours because of the spectral transmission of the Bayer CFA as illustrated in previous sections.

Figure 4.21 shows the typical spectral response of a Sony ICX285HQ CCD with Bayer CFA [140]. The top left corner shows the Bayer pattern, where one colour

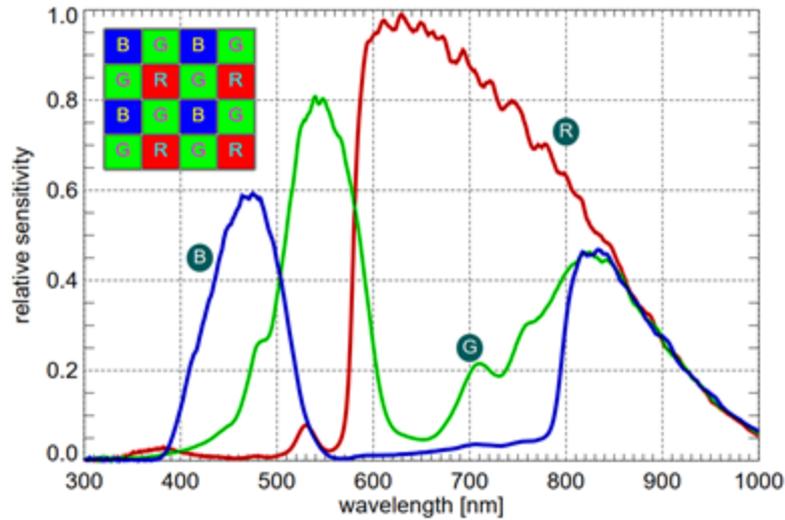


Figure 4.21: Typical spectral response of a RGB camera and Bayer CFA pattern.

filter corresponds to one pixel. Referring to the spectral sensitivity curve for each channel, the pixel value is an integration of photon counts over the sensitivity range of its corresponding colour filter. However, the Bayer CFA is also transmissive of IR in the range of 800 to 1000 nm. If the IR blocker is removed, NIR photons will be collected by the pixels. Notwithstanding the decrease in sensitivity in the NIR spectrum, the colour filter still allows a considerable amount of NIR. The additional NIR photons give rise to the overall pinkish or purple false colour. This artificial colour can be problematic for many other machine vision tasks. But for fire detection, the NIR sensitivity can be used to reduce false alarm. For example, the fake fires from advertisement posts or display will still appear as normal. Only in the events of a real fire emit NIR to render a purple fire image. Therefore, the final stage of the fire detection system takes advantage of this property for false positive rejection.

Figure 4.22 shows a comparison of image of diffusion methane flame with 3D visualization of the pixel colours in RGB colour space. The left flame image was captured without the IR blocker. In this case, the whole sensitivity spectrum with the Bayer CFA is used. The flame image on the right captured with fitted IR blocker,

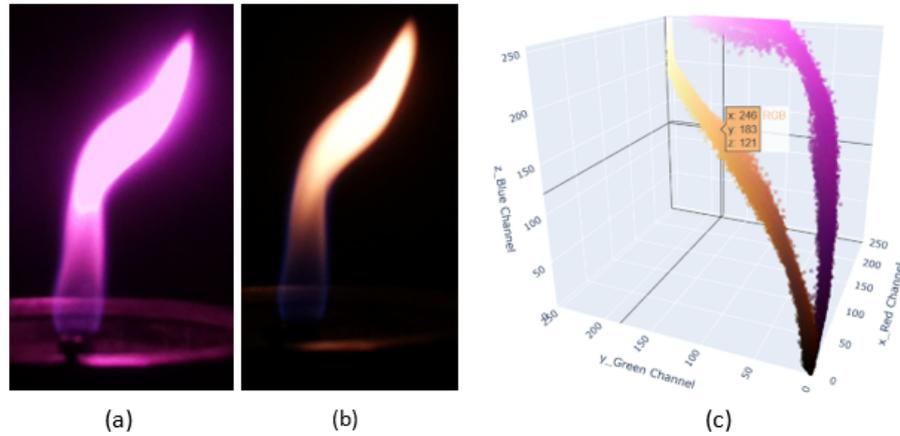


Figure 4.22: Colour space visualization of a diffusion methane flame : (a) NIR + Visible; (b) visible only; (c) 3D distribution in RGB colour space

which limits the sensitivity range to below 700nm. Because more light is allowed to reach the sensor, the flame image on the left is considerably brighter. The colour distinction is clearly presented in the RGB colour space plot.

To investigate the cause of the purplish false colour, NIR cut-on filter was applied to the camera. Figure 4.23 shows colour histogram of flame image obtained with 760 nm cut-on filter. The histogram peaks towards the right side. There are more blue pixels having larger values than the red pixel. This is counter-intuitive because at typical flame temperature flame emits more radiation in the infrared spectrum than in the visible spectrum. In addition, the silicon-based CMOS sensor is more sensitive to NIR than to visible lights. These two intuitions suggest that there should be more red pixels than blue pixels having higher values. However, the reason to cause the false colour is more near infrared photon leakage into blue pixels than red pixels. Referring to Figure 4.24, the quantum efficiency of the blue Bayer filter is higher than that of the red filter. As a result, the blue channel of an image will have a huge increase in pixel intensity because the response of the sensor to NIR photons is much higher. In addition, the red channel pixel values also increase. The compound effect

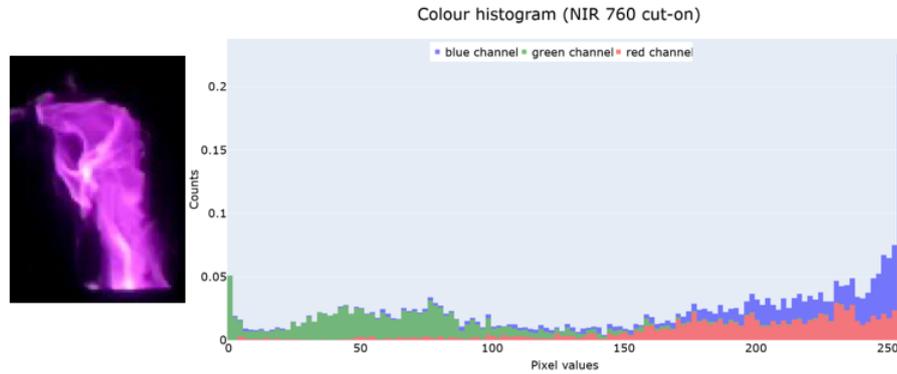


Figure 4.23: Colour histogram of diffusion flame captured with 760nm cut-on filter.

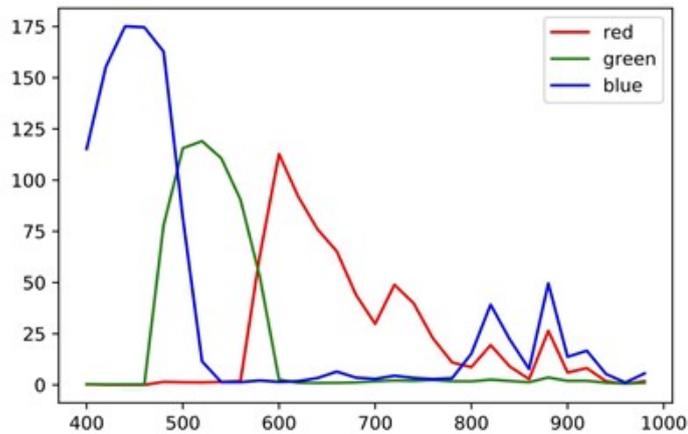


Figure 4.24: Bayer CFA spectral response up to 1000 nm [141].

is the pinkish colour of the flame images.

In the following, the colour distribution comparison will be made for two cases:

1. NIR+visible versus NIR cut-on at 760nm; 2. NIR cut-on at 950 nm at two exposures. In Figure 4.25, the pixel cluster of NIR case is more planar, but it spreads widely on the plane it forms. From the side view, the correlation of blue and red is proportionate, which corresponds to red and blue colour filter quantum efficiency curve. In comparison, the cluster for the whole spectrum case is tighter but thicker. From Figure 4.26, it is obvious that the distribution of pixels for NIR 760 nm case is wider than that of the NIR 950 nm case. In case 2, two exposure setting were tested: exposure 1 was 29991 μs , and exposure 2 was 9991 μs . The clusters of two exposure

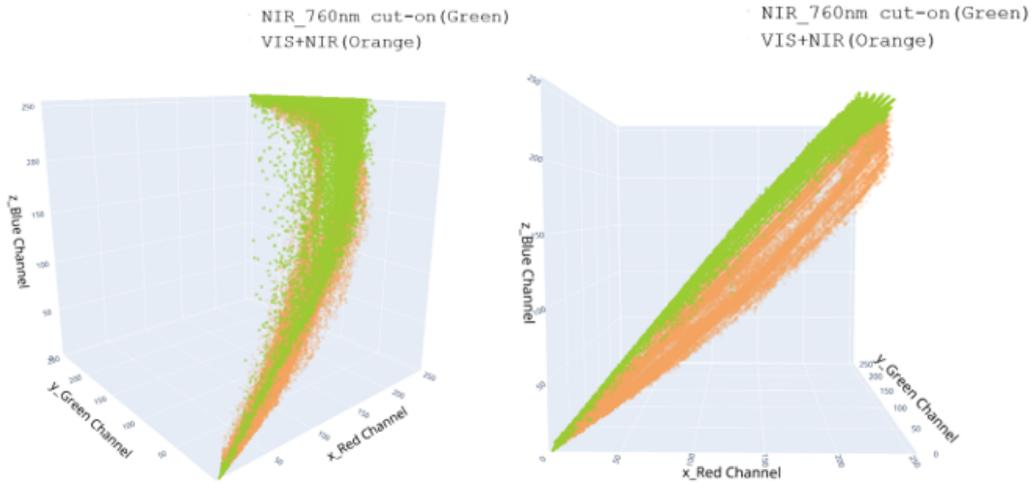


Figure 4.25: Colour distribution: NIR+Visible (orange) vs NIR cut-on at 760 nm (green).

cases overlap indicating that the exposure does not seem to affect the overall colour pixel distribution.

4.3.2.2 Colour distribution modelling

To make inference, the colour distribution is modelled by a Gaussian mixture model with Dirichlet process as prior. The conventional probabilistic models have fixed and finite parameters, which is likely to over-fit or under-fit because the number of parameters of the model does not match the complexity of the available data. Therefore, the determination of the parameters poses difficulties for model selections. On the other hand, there is the Bayesian non-parametric models to mitigate the situation, where the underlying distribution is inferred from the data. The Pi image sensor can be fixed at the experimentally determined setting, regardless of the scene illumination as well as the auto-white balance to avoid colour temperature change. The flame colour distribution showed good consistency. Therefore, it is possible to fit a function to the flame colour space, from which the pixel colours can be checked

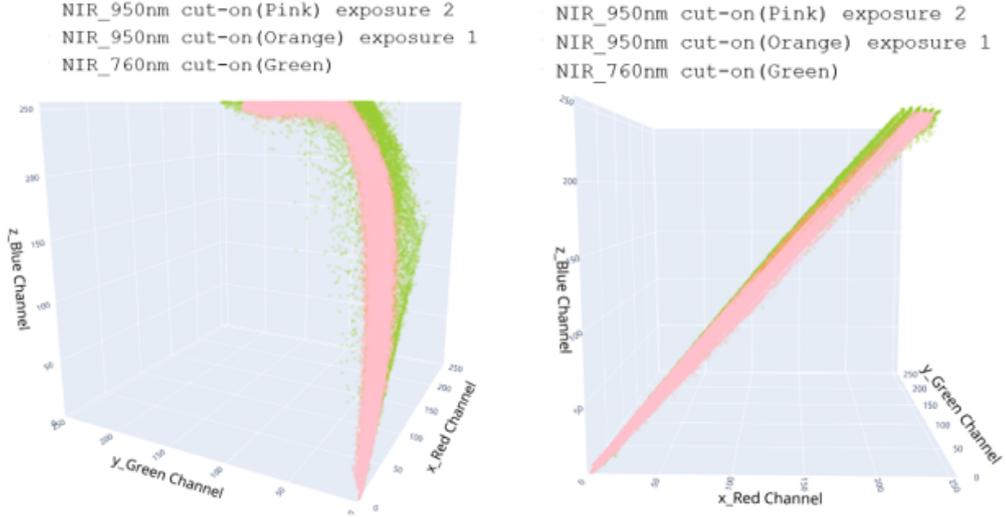


Figure 4.26: Colour distribution: NIR cut-on at 950 nm at exposure: 1. 29991 μs (orange); 2. 9991 μs (pink)

whether they belong to the flame category.

$$Dir(\mu|\alpha) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\dots\Gamma(\alpha_K)} \prod_{k=1}^K \mu_k^{\alpha_k-1} \quad (4.4a)$$

$$\text{where } \alpha_0 = \sum_{k=1}^K \alpha_k \quad (4.4b)$$

Dirichlet distribution is a generalization of multivariate beta distribution, as in Eq(4.4). The Dirichlet distribution is a distribution of distributions. This means if one samples from a Dirichlet distribution, a probability value sampled in the distribution determined by alpha will be returned. From Figure 4.27, Dirichlet distribution with a vector of three random variables (as in RGB space) are used for the sake of visualization. The concentration factor α affects the location of higher probability regions. The sampled distribution is multinomial implying that the sum of thetas is one. If all alphas are large, the distribution of the vector will concentrate towards the centre of the simplex. Whereas the probability will concentrate to the edges. Since

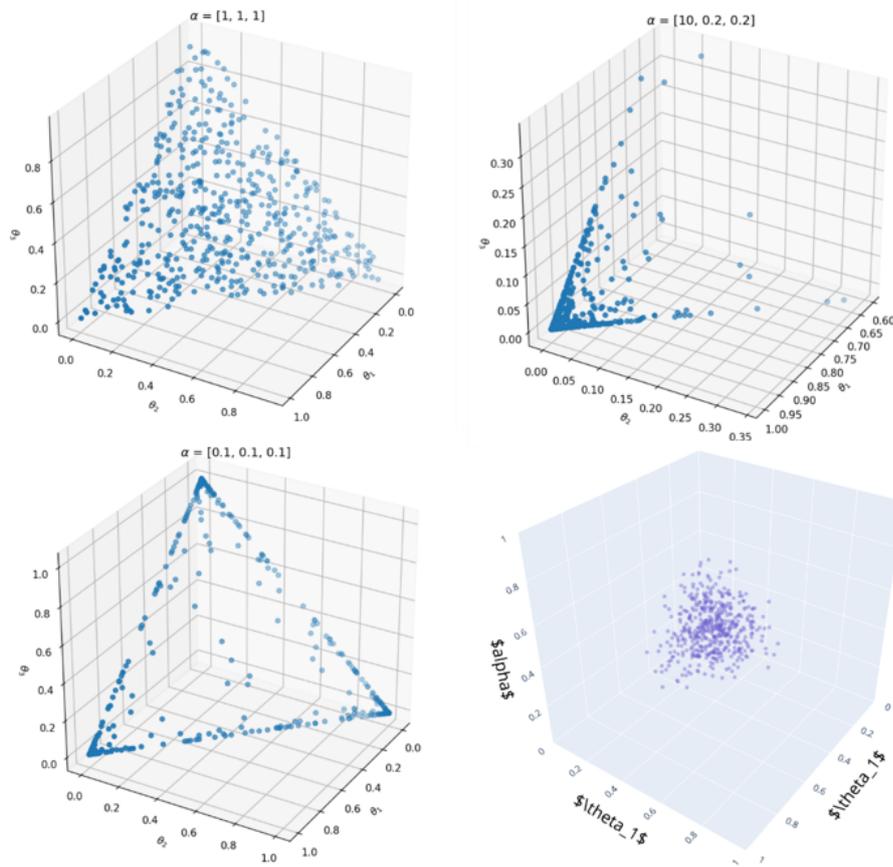


Figure 4.27: Visualization of Dirichlet distribution: $\alpha = [1, 1, 1]$; $\alpha = [10, 0.2, 0.2]$; $\alpha = [0.1, 0.1, 0.1]$; $\alpha = [10, 10, 10]$.

each α corresponds to one random variable in the vector, if one alpha is larger than others the distribution will shift to its corresponding corner in the simplex.

The difference between the traditional GMM model and GMM with Dirichlet process prior is how to choose the number of clusters. In this case, data visualization is convenient because the RGB colour space is 3D. One could roughly decide on the number of clusters needed to interpret the data. The orientation of each cluster depends on the its corresponding 3D covariance matrix. The traditional GMM rely on Expectation-Maximization (EM algorithm) with predetermined number of clusters. However, the EM algorithm can suffer from singularities, where some clusters may not have enough data points to effectively estimating the covariance matrices. As

shown in Figure 4.28, the ellipsoid represents a Gaussian visualized with 2.7 standard deviation. It can be seen that components at the corners are very small. The shape is almost squashed into 2D ellipse because one dimension in its covariance matrix is close to zero. The huge ellipsoid in the middle of the plot gives a much wider distribution than the data distribution.

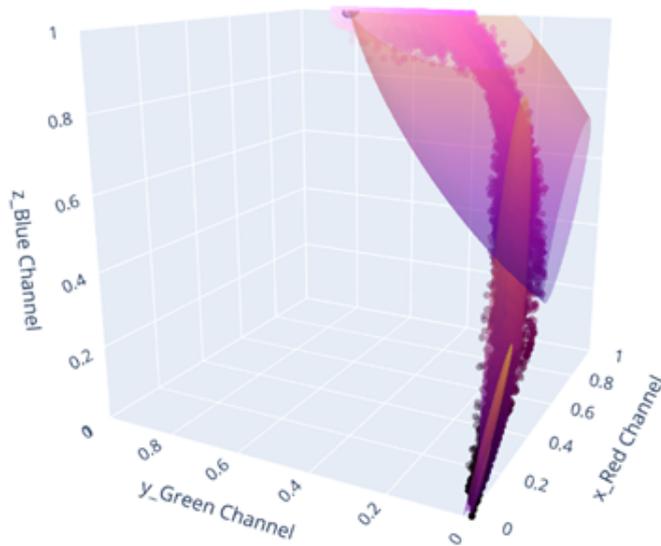


Figure 4.28: Visualization of fitting GMM with EM algorithm.

Using Bayesian model with Dirichlet process prior helps for a better fitting. As a Bayesian method, more parameters are needed than traditional GMM because of the inclusion of a prior distribution. The choice for the prior is defined through Dirichlet process, which is a generalization of the Dirichlet distribution to have infinite components. In other words, the DPGMM assigns data points into Gaussian clusters based on the probability of the data point in the present cluster. 'Infinite components' means the model can adapt to the data points distribution and accordingly assign Gaussian as many as needed to cover the whole of the data distribution. Figure 4.29 shows the fitting of DPGMM visualized using 2.7 standard deviation to construct the ellipsoids.

Bayesian Gaussian mixture models with a Dirichlet process prior

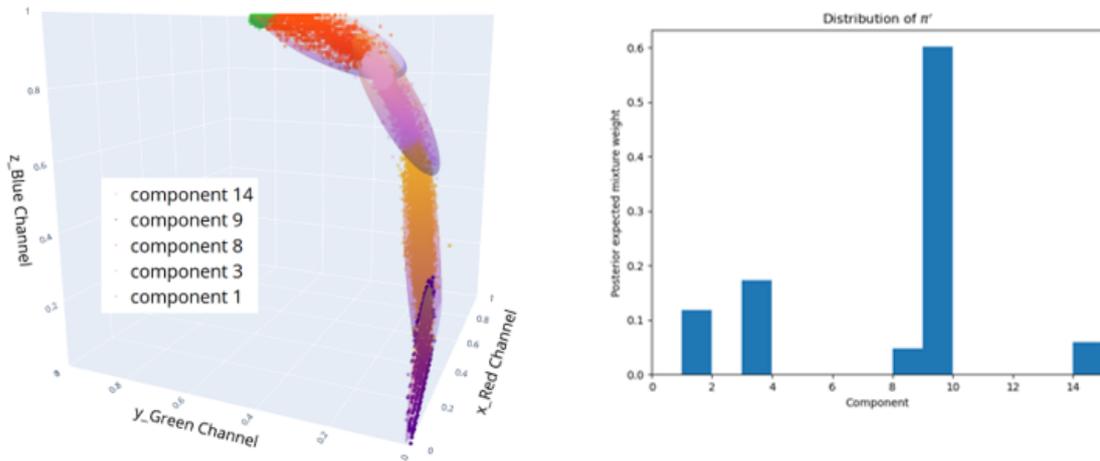


Figure 4.29: Visualization of fitting GMM with Dirichlet process prior, where 5 components are enough to fit the data. Right: weights assigned to each Gaussian component.

To eliminate the false positive examples, the ROI corresponding colour image will be checked using the fitted DPGMM. Specifically, the R, G, B values of each pixel in the patch will be fed into the model. If the new RGB data point falls within the ellipsoid showed in Figure 4.29, it will be marked as a fire pixel. The ROI patch filtering is then based on the percentage of pixels in the patch satisfies the colour model. The threshold is set to 0.2 to tolerate noise. Although the robustness of this method needs further tests, the false positives in the training and testing dataset used in this research were all rejected.

4.3.3 CNN for classification

Training a CNN model was also attempted to discriminate fire from non-fires. Both methods employed the same ROI extraction strategy. Comparing to SVM, the CNN model directly uses resized ROI patched of size 227x227 as input. Same as with the SVM classifier, the same training data is gathered in-house using 950 nm cut-on camera. Since there is no online database available, data augmentation was used

to enrich the data pool size by horizontal flipping and random gamma to mimic the intensity variation. Finally, the dataset contains 13878 fire examples and 15150 non-fire examples.

The CNN model structure is illustrated in Figure 4.30. Inspired by AlexNet in [142], with limited resource and only binary classification, this model takes similar but much simpler structure which contains three convolutional layers and one fully connected layer. The model used ReLU as nonlinear activation function; batch normalization was applied to the convolutional layers to re-centre the output from each layer; adaptive weight update was adopted by varying the learning rate; the initialization method used was Xavier. In total, the model contains 3936 trainable parameters.

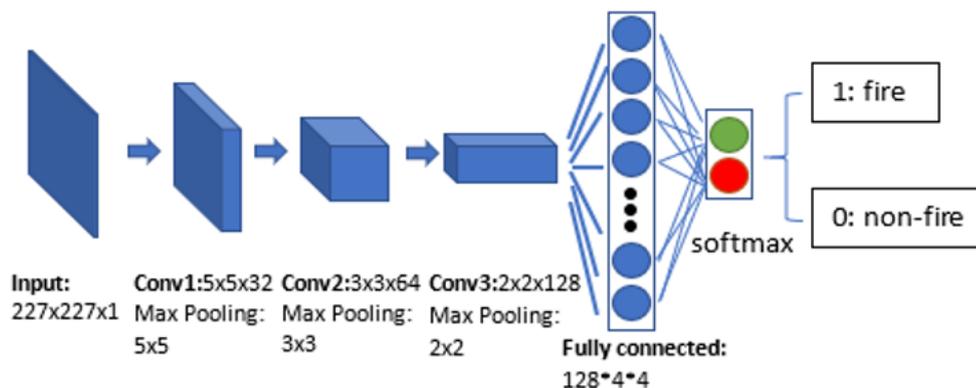


Figure 4.30: Illustration of CNN model structure: 3 convolution layers with max pooling, one fully connected layer and softmax classifier.

The test set detection rate of both CNN and SVM used is plotted in Figure 4.31 with respect to distance to camera. It can be seen that the detection rate of CNN drops significantly as the fire is further away from the camera. The SVM classifier is only slightly affected. Figure 4.32 shows the history of loss error and accuracy during the training at each epoch. The training accuracy was 100 percent, which is an indication of over-fitting given the test accuracy is much lower. The trained CNN is

far from optimal, which requires further fine-tuning to reduce over-fitting, re-structure and more diverse dataset. Despite of the enlarged dataset by data augmentation, the CNN performance is inferior to that of SVM. The generalization of SVM classifier is better, given that only a few hundreds of training examples used. Moreover, training data used for the SVM were images of 5m from camera. It still performance very well for unseen fire images at greater distances, showing the effectiveness of the feature extraction stages.

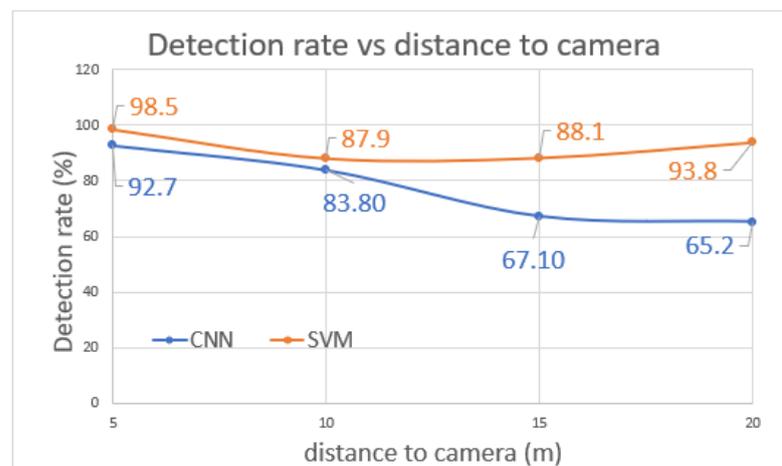


Figure 4.31: CNN and SVM detection rates vs distance to camera.

The comparison between CNN and SVM used in this case reflects the practical considerations of choosing the appropriate machine learning algorithm as previously discussed. Provided with a well-defined dataset, neural networks will give very solid performance, but the fine-tuning may entail extensive experimentation. The fine-tuning process is more empirical than definitive that no clear guideline on how the structure should be for a specific case. One might resort to transfer learning to retrain the existing model to speed up the process. However, in specific situations, these choice might not always be available due to the different type of data for different tasks. SVM on the other hand is very effective in cases that the dataset is limited.

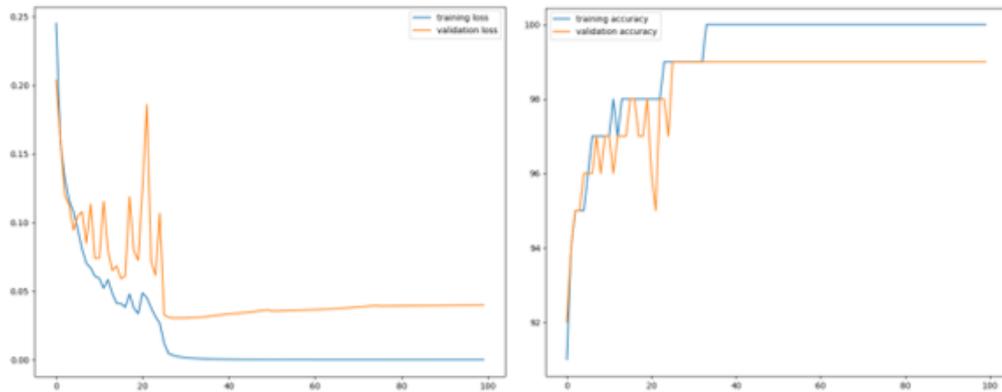


Figure 4.32: Training graphs of CNN. Left: loss error vs number of epoch; right: accuracy vs number of epoch. Blue: training set; orange: validation set.

The training process is so much faster as only \mathbf{w} and b needs to be found, which is proportionate to number of features in each example.

4.4 Summary

The trained SVM is very effective in distinguish fire from non-fire objects. The false positive examples seem to have more texture and pixel variations. The most obvious example for this situation is light reflection from the ground. This could happen especially when the ground if made of grits, which can be reflective and grainy. The false negative examples all appear less bright comparing to other fire examples for each case. The main reason is the disturbance from ambient air flow, which reduced the texture and contrast of a flame image. Notwithstanding more missed detections in the 10m test case, the performance is still satisfactory. Because in real-time systems, many frames will be checked within a second, the fire will be detected in time. Moreover, this method does not rely on temporal feature, which make it suitable for moving camera situation.

The Bayer CFA is not only transmissive in the visible spectrum, but also in NIR spectrum at a lesser degree. The blue colour filter allowed more NIR photons to reach

the sensor than those of the red colour filters. The NIR photons leaked into red and blue channel without the IR blocker gives the image a pink taint. For cases with NIR cut-on filters, the colour distribution trend showed some consistency, especially for the case with 950 nm cut-on filter. Moreover, the exposure seemed not to influence the colour distribution. The ability to see NIR gives this camera sensor an edge over the common colour cameras for fire detection. Comparing to the use of monochrome LWIR from the fire detection literature, the colour can be used as an extra layer of filtering to facilitate the image processing pipeline.

Two popular machine learning algorithms for classification were attempted in the case of fire detection. They all can give very good performance given that the hardware nowadays are better optimized for machine learning tasks. The question of which one to use has a solution in the practical considerations for specific task. One advantage of using Pi camera system is the system configurability that the camera setting can be fixed and work consistently. Many fire detection research reported that illumination change affects system robustness as the camera system's auto white balance shifted the colour significantly. In addition, the Pi system as a single-board computer offers on-board computability which enables easy deployment. Furthermore, the two camera module and the Pi platform only cost less than 100 GBP. The much lower cost with NIR capability really makes it an appealing choice for fire detection, especially compared with any other types of IR cameras. Finally, the combination of flame physics, algorithms and multi-spectral capability has the potential for robust fire detection systems, comparing to many other research that solely treating fire detection as an algorithm problem.

Chapter 5

NIR combustion diagnostics

5.1 Introduction

The hydrocarbon flame spectroscopy in the IR spectrum contains lots of exploitable information for diagnostics as shown in Figure 5.1. In the field of combustion diagnostics for lean premixed flames, NIR diode-laser based absorption spectroscopy was well studied in literature, where the concentration of combustion products such as H_2O , CO_2 , CO and CH_4 , etc can be inferred by the integral integrated line absorption. In addition, gas temperature can also be calculated by the ratio of line intensities at selected absorption wavelength. The knowledge of the distribution, concentration and gas temperature is particularly useful in combustion control applications. Particularly, CO , CO_2 and H_2O as the main combustion products in premixed lean hydrocarbon flames are effective in determining the combustion efficiency. Moreover, the temperature of the flame is a key indicator of combustion thermal efficiency. The vibrational overtones and combination bands of these species are usually discretely distributed with overlapping in the infrared spectrum.

However, carefully selected narrow bands can make room for measuring individual species with much less interference from other species. Figure 5.2 shows the line absorption strength from HITRAN database for CO , CO_2 and H_2O , where non-

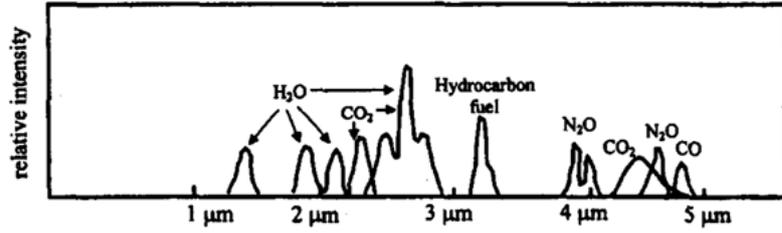


Figure 5.1: Emission relative intensity of hydrocarbon flames[122].

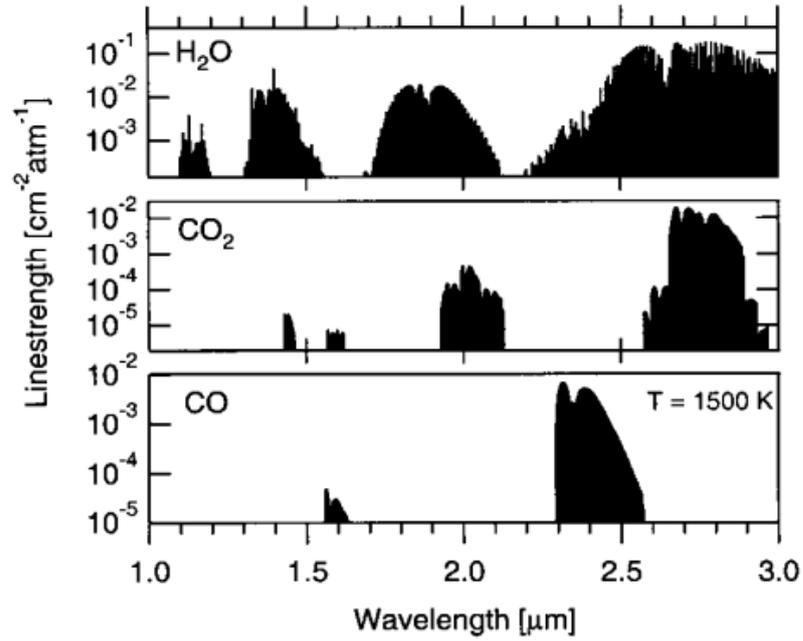


Figure 5.2: Line absorption strength of H_2O , CO_2 and CO at 1500K.

overlapping regions for these species can be utilized for NIR absorption spectroscopy studies. [123] exploited the NIR absorption spectrum, in particular for CO at $2.3 \mu\text{m}$, H_2O at $1.343, 1.392, 1.799 \mu\text{m}$, to measure species concentrations. In addition, the temperature was inferred from the ratio of selected bands.

The popularity of NIR absorption spectroscopy rather than the mid and long wavelength IR for combustion diagnostics may be partially ascribed to the advancement of sensor technology. For example, NIR detectors, such as InGaAs detectors usually operating from 900 nm to $2.5 \mu\text{m}$, is a common choice because of its relative low cost and can be operated at room temperature. In addition, the natural

advantage of photosensitive conductor is fast response comparing to the common thermal detector type of long wavelength IR sensors, enabling its capability for high temporal resolution applications. Meanwhile, with the CMOS fabrication technology compatibility, ROIC can be integrated giving it fast readout and relative low-cost.

In contrast to the active instrumentation using NIR laser absorption spectroscopy, the passive instrumentation whereby NIR radiations are passively gathered by NIR detectors, is much less used in literature. The main reason could be the overlapping emission spectrum of the CO_2 and H_2O . Nevertheless, NIR passive imaging can be very useful in some situations.

5.2 NIR spectrum of a flame

The emission spectrum generated in flames can be categorized to 3 types: (a). transition of electrons from one stationary energy state to another (electronic transition), which corresponds to single spike in the spectrum; (b). in addition to the electronic transition in 1, the change in the internal vibrational and rotational energy of the molecules, which usually produces multiple banded spectrum; (c) through emission or absorption by solid particles or droplets, or by ionization or recombination, which generate continuous spectrum [123]. The sensitivity of silicone-based CMOS sensors is capable of acquire the line-of-sight spectrum of water molecules as well as carbon dioxide in the NIR spectrum up to 1100 nm. Their emission in the NIR is only a fraction of the banded spectrum throughout the whole IR spectrum. Such emitted radiation is not generated by any change of quantum state but by the change of vibrational and rotational energy.

Most combustion is an exothermic chemical process, which involves energy conversion and heat transfer. The released energy interacts with the ambiance through three manners: conduction, convection and radiation. In a real scenario, most of the energy is lost by conduction to the burner and by convection with surrounding cool air. The part of energy by radiation is accounted by the emission spectrum of the flame, which usually takes a wide range of the spectrum from UV to far-IR. The radiation from combustion under different conditions have different banded spectrum and peaks. Generally, the energy emission in VIS and UV accounts for no more than 0.4 percent of the total combustion energy. However, the radiation energy in the infrared spectrum can take as high as 20 percent of the total [123].

In the case of diffusion hydrocarbon flames, the presence of soot from incomplete combustion will have an emission spectrum more resembling the Plancks blackbody

radiation, which covers a wide spectrum. The soot radiation spectrum dominates over that of the premixed part at the root of the flame. In the case of premixed lean hydrocarbon flames, however, main observations of the emission spectrum come from excited intermediates, such as OH^* , C_2^* , CH^* , etc. Extensive research has been published, establishing the correlations between the combustion properties and the emission of those intermediates. Different from the dominant peaks in VIS and UV of the intermediates, which is generated during the combustion chemical process. The main emission in IR comes from the combustion product carbon dioxide and water vapour.

5.3 Passive NIR imaging of fuel-lean hydrocarbon flames

In this research, simultaneous visible and near-IR imaging of premixed methane and hydrogen flames is utilized to study the flame behaviours and structure. Various equivalence ratios are compared for the case of methane and hydrogen premixed flames. In addition, a colour analysis is made to show that the colour is dependent on the temperature of flames as well as the optical property of Bayer colour filter.

To avoid any confusion, two types of raspberry pi cameras were used: 1. Sony IMX219 without IR blocker (pi camera V2); 2. Omnivision OV5647 (pi camera V1). The first sensor is sensitive to near IR up to 1100 nm but still have the Bayer Colour array on the sensor. Whereas the second sensors Bayer colour filter is removed chemically by in-house developed procedures, which means the whole spectral sensitivity range of the sensor is usable without any influence by in-built components. The removal of the Bayer filter for the second sensor turns it to a monochromatic sensor, thereby every pixel receives equal radiation. However, the combination of near-IR sensitivity and Bayer colour filter of the first sensor will introduced false colour because of its optical property. As the silicon-based CMOS image sensor is more prevalent than ever, there can be NIR leakage due to the difference of integration of IR blocker from different manufacturers. The study of such false colour can be instructive of dealing with problems as such.

5.3.1 Fuel-lean premixed flames imaging in 300-1100 nm

Figure 5.3 shows the time averaged premixed methane flame of 10 images over the sensors spectral sensitivity range (300-1100nm). The camera setting is fixed to 15 fps with ISO 800. It should be noted that the irregular dark spots are caused by

the debris of the Bayer colour filter from the chemical removal. The image reveals the inclusion of line-of-sight emission of the intermediate species, such as CH^* , C_2 , OH^* , and combustion product of H_2O and CO_2 . The inner cone bright outline mainly comes from CH^* and C_2^* , which have strong peaks at 430 nm and 515 nm, respectively. The surrounding fuzzy region of the cone corresponds to the rotational and vibrational bands of H_2O and CO_2 , which stretches much further upwards than the visible appearance of the premixed flame.

It is worth mentioning that the imaging technique should be restricted for qualitative measurements in terms of radiation signal strength, unless a thorough spectral response of the sensor over the sensitivity range is calibrated with respect to the absolute value. Referring to the sensor spectral response in Figure 3.11, the spectral response is nonlinear, which is much stronger in NIR than in the visible range.

Narrow band filters with FWHM = 10 nm (Full Width-Half Max), were applied to the sensor at emission peaks of three species of CH^* , C_2^* , OH^* . As Figure 5.4 illustrated, the sensor sensitivity at 309nm is very weak, which resulted in a weak and fuzzy cone. The image for C_2^* is noticeably sharper than that of CH^* , indicating a higher concentration at the shell of the cone.

Near infrared cut-on filter was also applied to isolate the emission of H_2O and CO_2 as shown in 5.5. From the cut-on wavelength to sensitivity range of the sensor, hot combustion products were visualized. It is obvious that as the bandwidth decreases, the resultant image is dark because of less integration of photons. It is also noticeable that the sharpness drops from the wider bandwidth to narrow bandwidth. The most apparent change in terms of flame structure is the hot CO_2 and H_2O , which stretches much longer than the normal visible flame. The necking effect can be observed at the

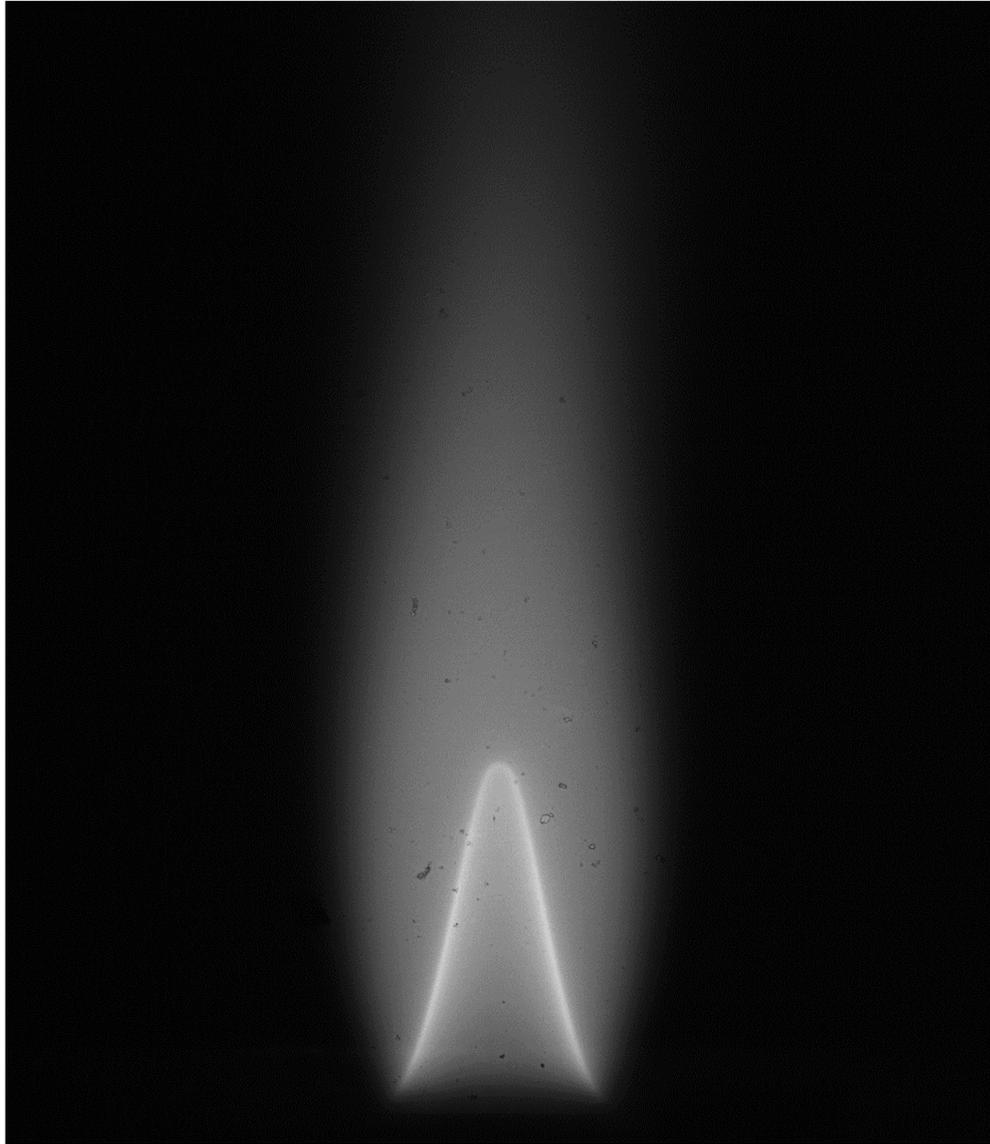


Figure 5.3: Lean premixed methane flame in complete sensitive spectrum of Si-based CMOS sensor (300-1100 nm).

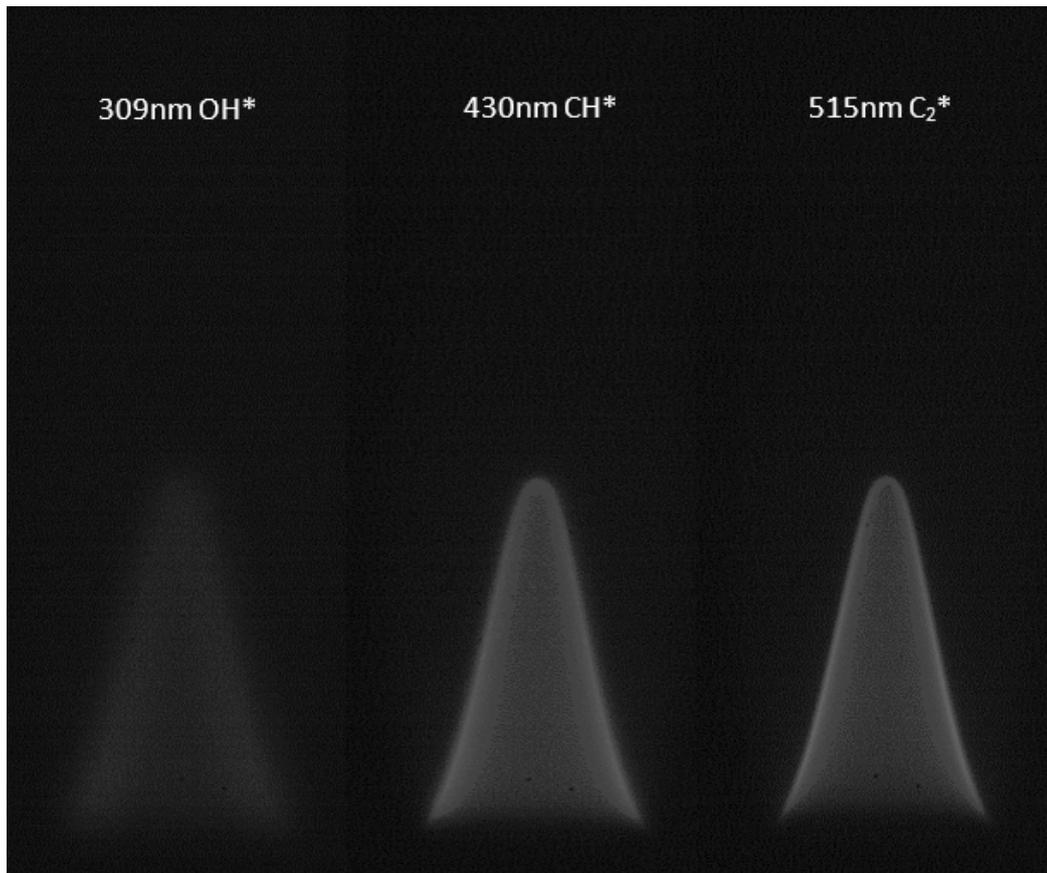


Figure 5.4: Narrow band imaging with lean-premixed methane flames. Left: 309 nm for OH^* ; middle: 430 nm for CH^* ; right: 515 nm for C_2^*

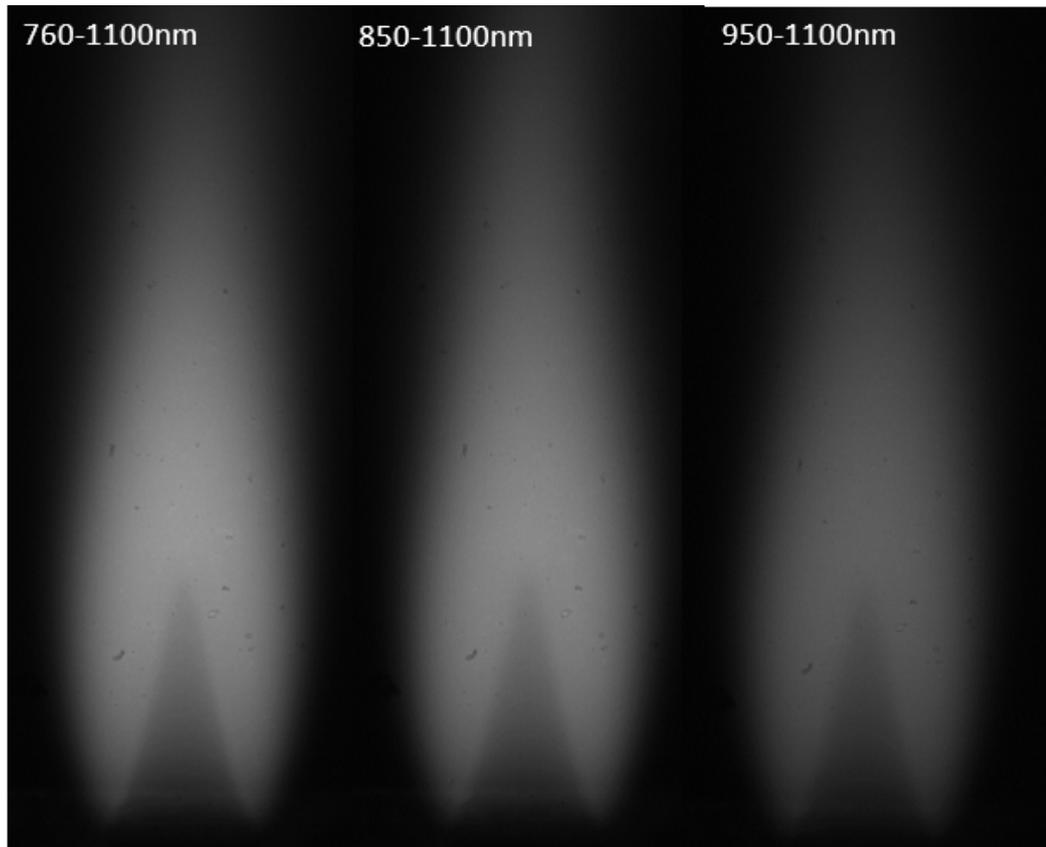


Figure 5.5: Short band NIR imaging with lean premixed hydrocarbon flames. Left: 760-1100 nm; middle: 850-1100 nm; right: 950-1100 nm.

mid length of the flame. Moreover, the cone in the core of the flame comes from the lesser integration of line-of-sight emission of the combustion products, because the fuel and air mixture is not reacting to produce H_2O and CO_2 .

5.4 Effect of Bayer CFA on hydrocarbon flame colour

The effect of near infrared radiation of spectral range up to 1100nm on the resultant flame colour is studied through two cases: 1. Methane diffusion flames (VIS versus 300-1100 nm); 2. Premixed methane and air flames (VIS versus 850-1100 nm). In both cases, the sample flame images were taken simultaneous.

In Figure 5.6 top left, the extra NIR radiation of the diffusion methane flame received by the pixels render the image with a purplish tint. It is visually noticeable that the brightness is larger comparing to the VIS image. The hue colour scale used here is [0, 180]. The purplish colour corresponds to hue range of 140 to 155. By contrast, the yellowish to red colour of visible diffusion flame having a hue range of 0 to 30 and 160 to 180, which is the emission of soot particles. The weak premixed bluish flame at the root corresponds to CH^* and C_2^* gives hue range of 110 to 130. There is also a purplish colour with hue range of 130 to 160 at the edge and in the vicinity of the joint of diffusion and premixed region in the root. Because in these regions, both soot emission and premixed bluish radiation are weak as the integration depth is smaller, the line-of-sight integration of both generates such colour.

Referring to Figure 5.7, a considerable amount of infrared radiation reached the pixels through all colours of the Bayer colour filter. Especially in the blue channel, a significant amount of pixels shifted from low values to higher value comparing to that of the visible image. A much larger portion of pixels having values in range of 150 to saturation, whereas no saturated pixels in the blue channel of visible flame image. It can also be seen that a great portion of low range pixels in red channel were shifted towards the right of the histogram.

In the premixed methane flame case, the effect of NIR from hot CO_2 and H_2O on

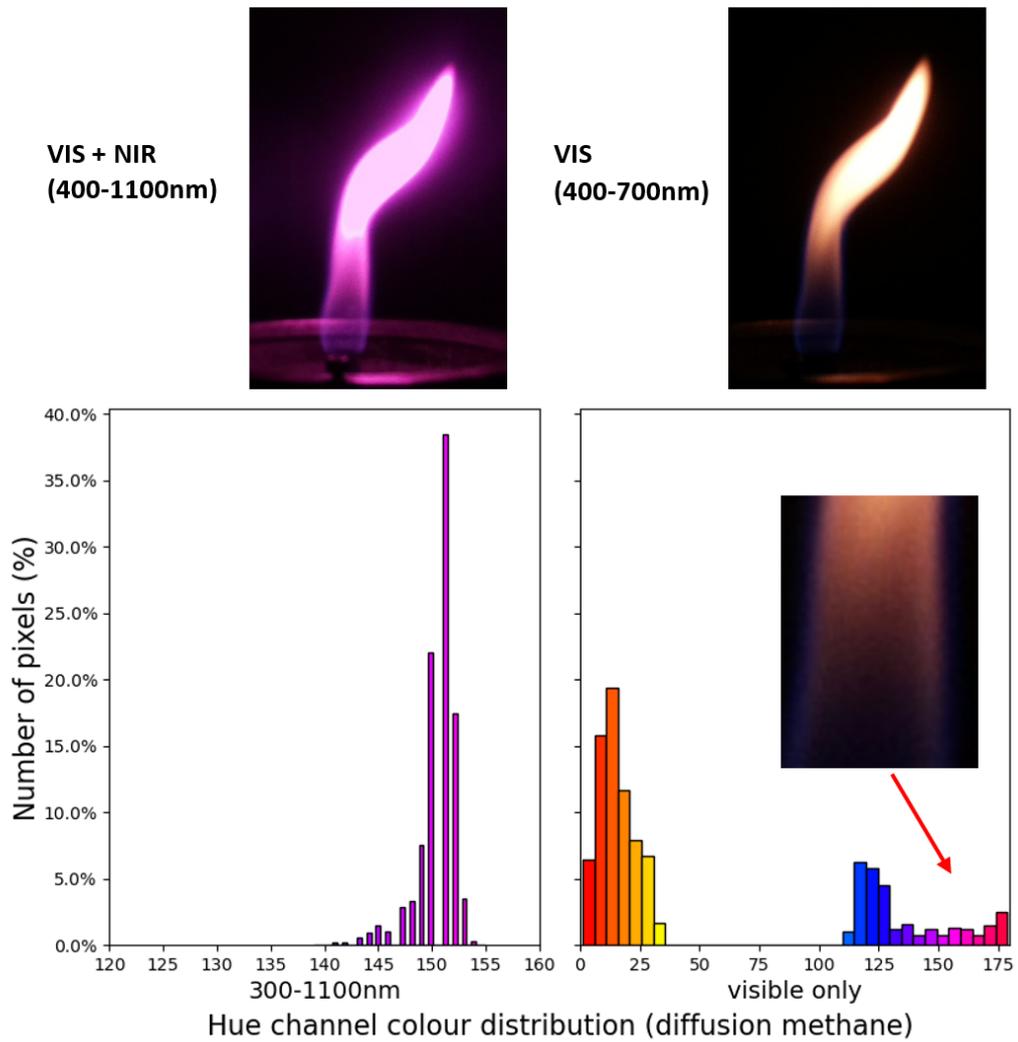


Figure 5.6: Diffusion methane flame imaging VIS vs VIS+NIR and hue channel distribution.

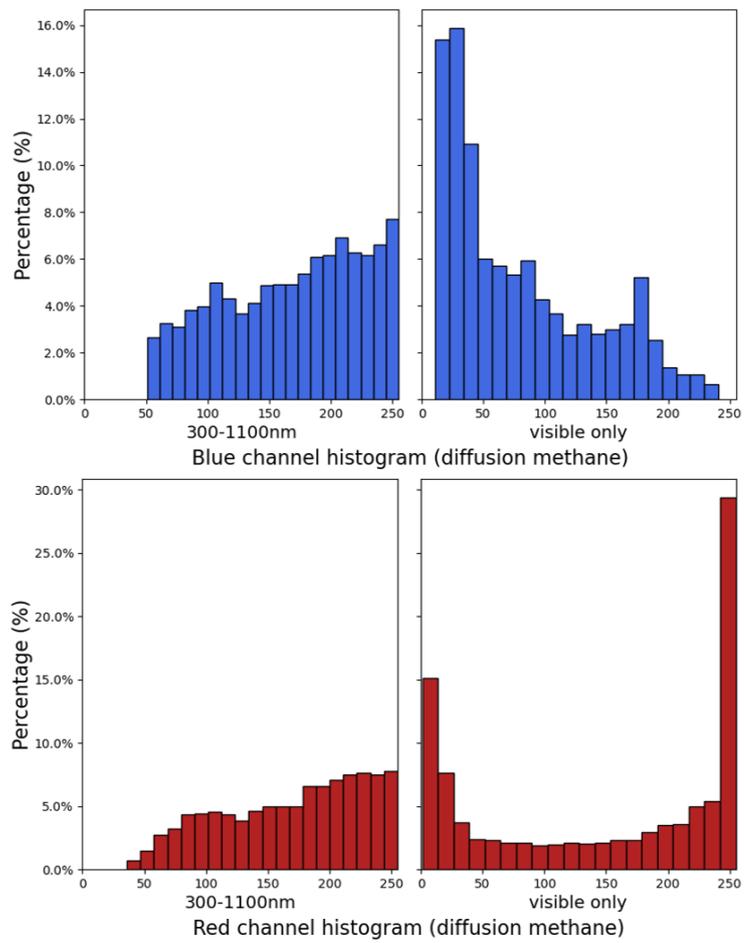


Figure 5.7: Diffusion methane flame blue and red channel histograms: VIS vs VIS+NIR.

the colour of flame image is also studies across the range of equivalence ratio in the fuel-lean condition. The radiation from CO_2 and H_2O is isolated by applying a cut-on filter at 850nm, which give the sensor sensitivity range of 850 to 1100nm. Meanwhile, the visible premixed methane flame is simultaneously captured for comparison. The camera setting is fixed at exposure time of 15.5 ms with ISO 400 at 30fps for radiation intensity benchmarking.

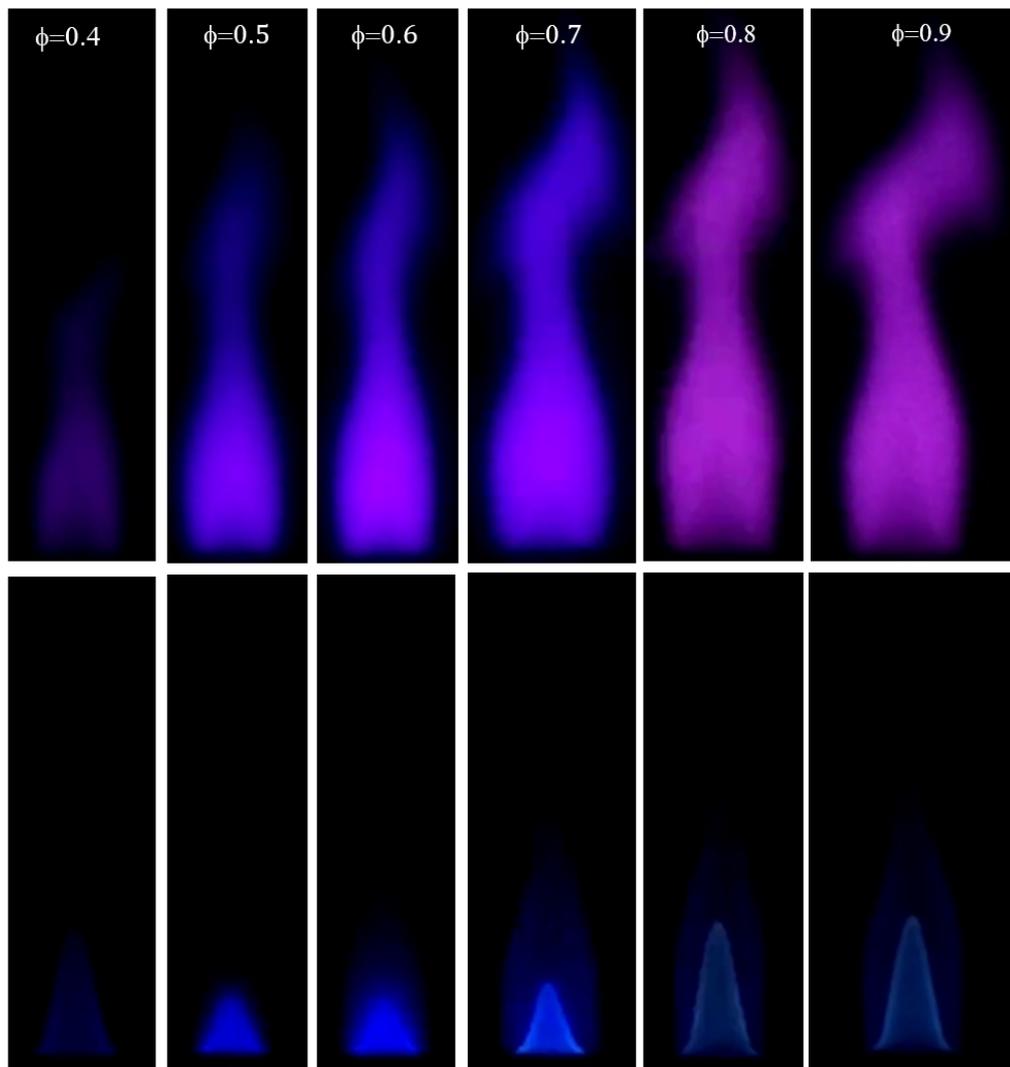


Figure 5.8: Premixed methane flame imaging in fuel-lean condition: VIS (bottom) vs VIS+NIR (top).

Figure 5.8 shows the synchronized dual band imaging at premixed equivalence

ratio from 0.4 to 0.9. The total volumetric flow rate for methane case as well as hydrogen case are fixed at 3.0 L/min. There is a discernible shift of colours as the equivalence ratio increases towards stoichiometric. Because of the cut-on filter fitted to the camera, the images are the captured CO_2 and H_2O in the spectral range of 850 to 1100 nm. The wobble of the hot combustion products is stronger as the equivalence ratio increases. The hue channel histogram of flames at each equivalence ratio is plotted in Figure 5.9. Overall, the flame colours shift from blue (120-125) to purple (up to 145) until the equivalence ratio reaches 0.9.

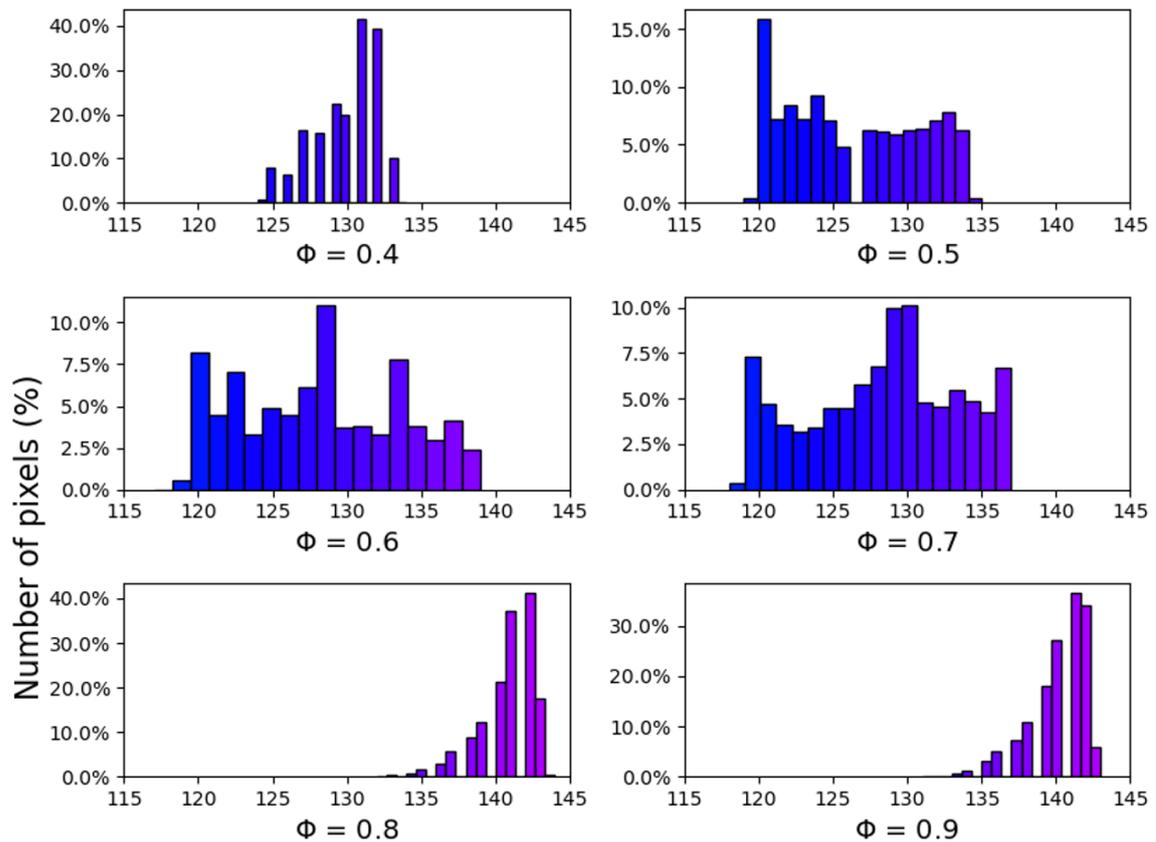


Figure 5.9: Premixed methane flame colour shift w.r.t equivalence ratio: (850-1100 nm).

In comparison, the visible flame colours are very consistent and centred at hue of 120, as shown in Figure 5.10. At equivalence ratio of 0.7, 0.8 and 0.9, a peak

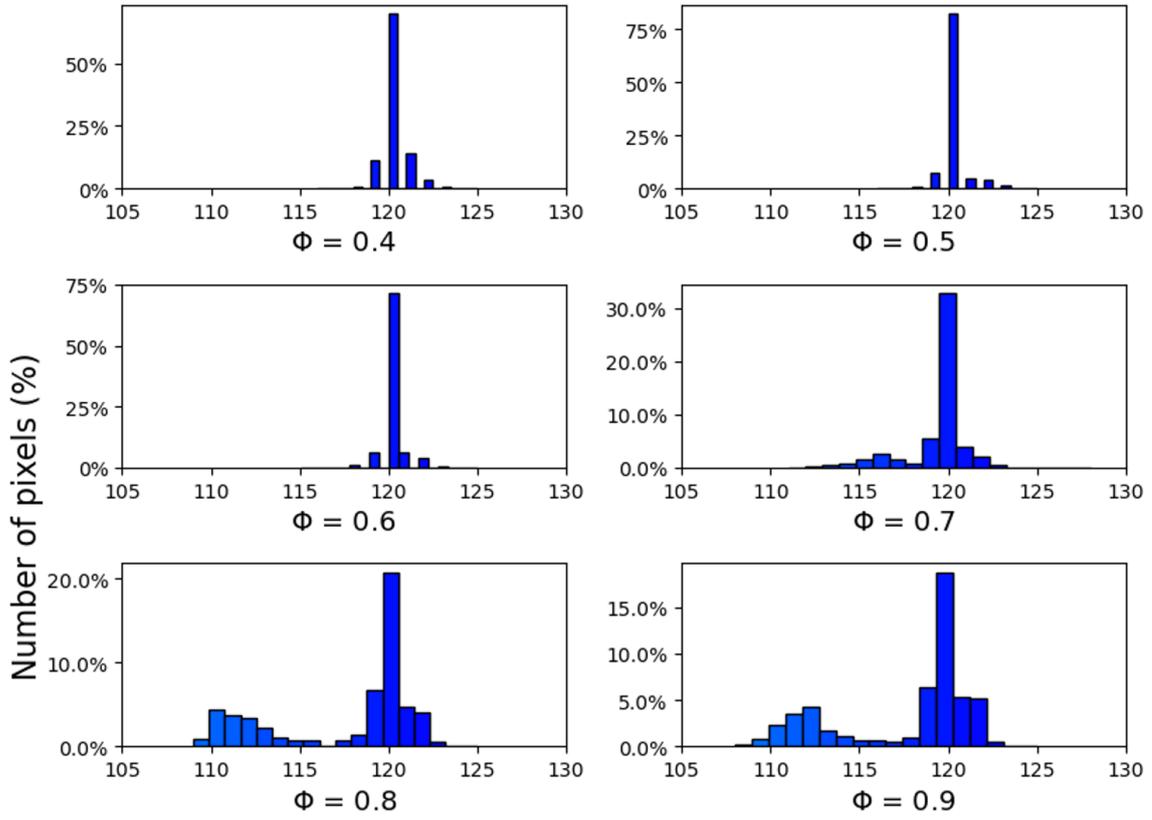


Figure 5.10: Premixed methane flame colour shift w.r.t equivalence ratio: visible spectrum.

of light blue centred approximately within 110-115. This should be caused by the mixed emission from CH^* and C_2^* , which highly concentrates at the vicinity of the cone where the combustion starts [144].

To consolidate that the tinted colour appearance is due to the Bayer colour filter having significant transmission of the NIR radiations, the colour channel histogram in RGB colour space is plotted for the case of equivalence ratio of 0.8 in Figure 5.12. It is obvious that all channels in the NIR case have histograms shifted towards the right comparing the that of the visible case. Figure 5.11 gives the official spectral response of this camera sensor with IR blocker. The integrated Bayer colour filter throttles all three colours near 700 nm, which is presumably cause by the IR blocker. The actual spectral response should have peaks in the region of 700 to 1100 nm.

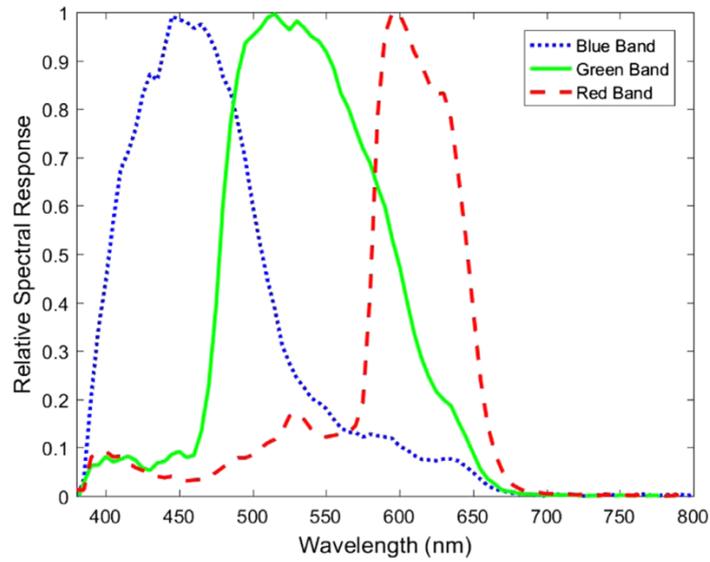


Figure 5.11: Camera Bayer CFA spectral response in visible spectrum [145].

Because the Bayer colour filter are long-pass filters with peaks at each desired colour wavelength, it is expected that the red channel has a considerable response in the spectrum of 850 to 1100 nm. Nonetheless, the comparison of channel intensity histograms prove that the blue and green filter has higher transmission of NIR than that of the red filter.

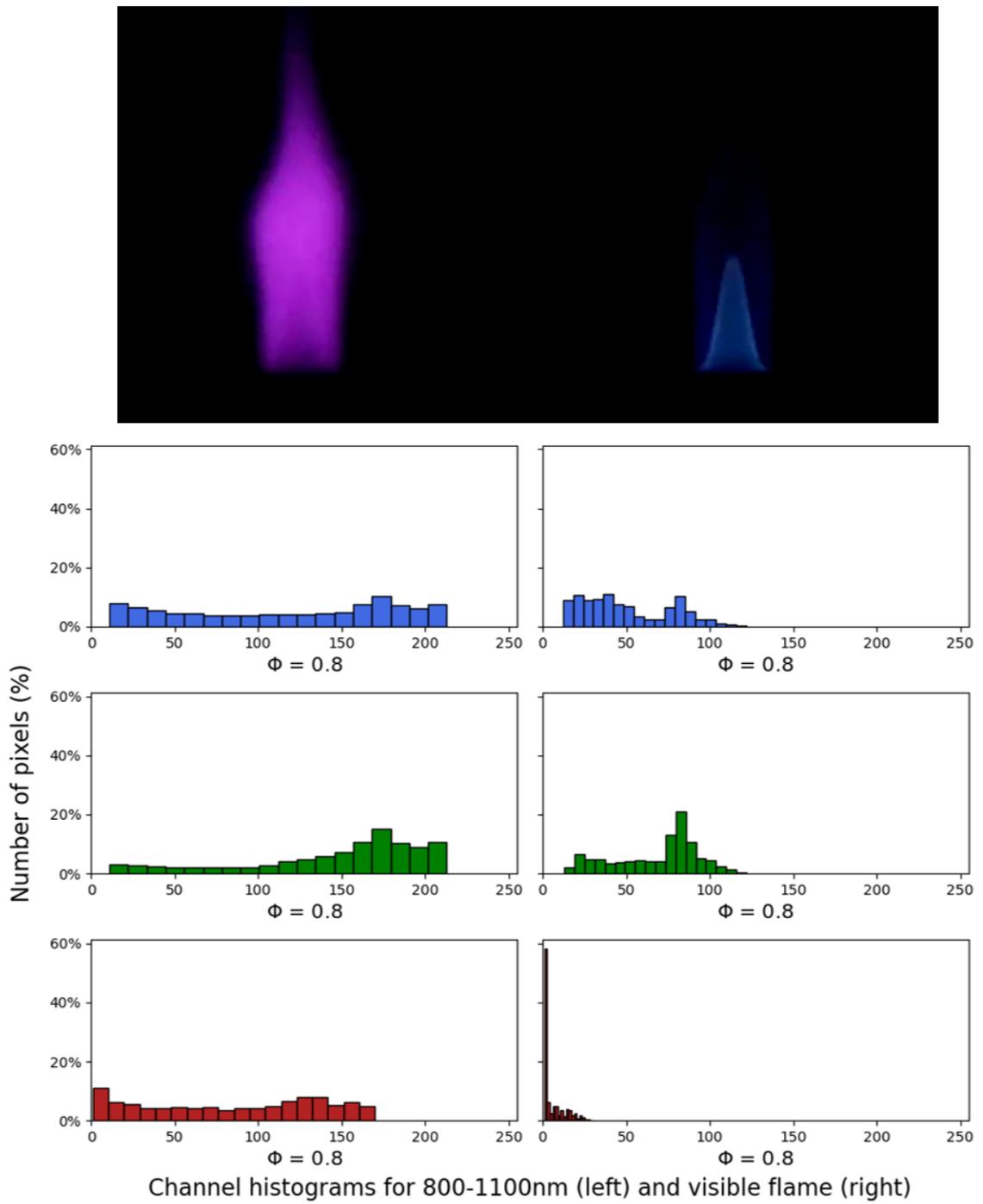


Figure 5.12: Premixed methane flame R,G,B histograms at equivalence ratio 0.8.

5.5 Premixed flame with NIR imaging

Premixed flames are widely used in industrial applications. The control of such flames to achieve stability is of great importance. There are many factors affecting the flame stability, some of which have quantified variables to relate the phenomenon to theoretical modelling, for instance the Lewis number in Eq(5.1), which is the ratio of thermal diffusivity and mass diffusivity [146]. Also, Damkohler numbers Da defined as the ratio of reaction rate to mixing rate. The instability is also coupled with concentration of reactants. [147] showed that if both reactants have similar proportion, given the Lewis number is larger than unity, flame pulsation is likely to happen. In the regime of premixed flame, the Lewis number of lean reactants will introduce diffusive-thermal instability [148]. Moreover, the effect of Lewis number and equivalence ratio on the cause of flame pulsation has been studied experimentally, which showed good agreement with theoretical modelling [149].

$$Le_i = \frac{\alpha}{D_{ij}} \quad (5.1)$$

Instability is also caused by stretch rate of the flame. These parameters bearing different names are independent as mathematical quantities. However, they are closely related in terms of physical meaning. Contrasting to premixed hydrocarbon flames, hydrogen premixed flames is more sensitive to stretch, which is more likely to cause instability [149]. The coupling effect of fuel diffusion and thermal conduction can also promote thermal-diffusion instability in the case of stationary planar flame. Because instability arose dependently on diffusive process, which is dominantly affected by combustion chemical reaction [150].

Thermal radiation loss is another significant aspect on flame instability. The

radiation energy emission is usually larger from the downstream hot combustion product than the upstream reacting gases [151]. This loss of energy contributes to the lower maximum temperature comparing to the adiabatic flame temperature. Therefore, the ability of direct imaging of CO_2 and H_2O provides a valuable tool for premixed flame diagnostics at a very accessible cost.

5.5.1 Premixed methane and hydrogen flame pulsation with NIR imaging

To study the pulsation of premixed flames through NIR imaging, the experiment was set up to visualise the combustion product of CO_2 and H_2O . The experiment layout is shown in Figure 5.13. Flow meters were used to control the equivalence ratio. And a short band optical filter of 850-1100nm was applied to the camera.

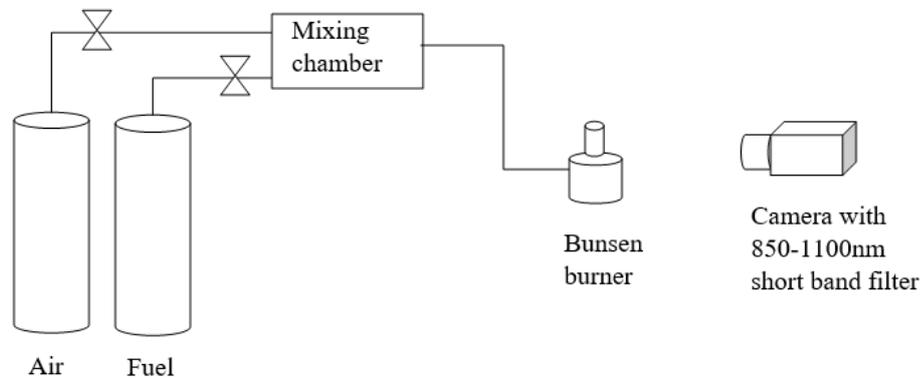


Figure 5.13: Experiment setup for NIR combustion product visualisation.

Time series data were extracted from images captured with spectral range of (850-1100 nm), for two cases: premixed methane and air, premixed hydrogen and air. For premixed methane flames, the equivalence ratio was kept at 0.4, 0.5, 0.6, 0.7, 0.8, 0.9. For premixed hydrogen, it was kept at 0.5, 0.7 and 1.0. As illustrated in Figure 5.14, a patch of pixels in blue channel for both cases were extracted for analysis. The patch was centred at pixel location of (225, 425) for methane, where the necking occurred.

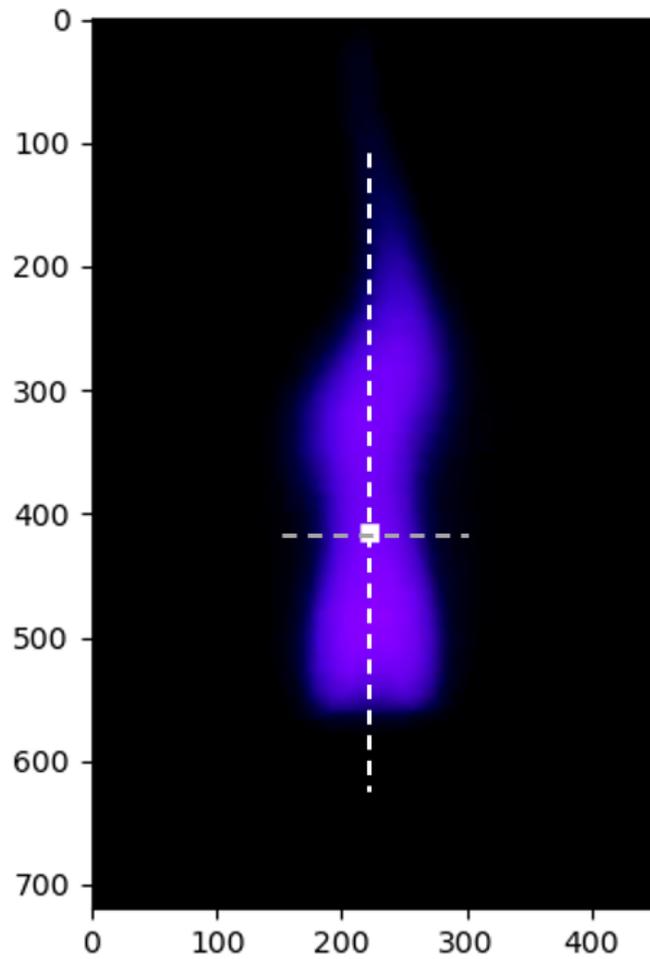


Figure 5.14: pixel value extraction lines for monitoring. Vertical: centre line; horizontal: necking position; square patch: intensity monitoring.

It is centred at location of (320, 450) for hydrogen flame. To analyse the pulsation of the flame, the flame length at the vertical centreline and horizontal position were also recorded. The vertical stretching and horizontal wobble of hot CO_2 and H_2O are good indicator for quantitative measurement of pulsation with respect to equivalence ratio. Image sequence were gathered at 30 fps with fixed exposure and ISO, thereby comparative intensity analysis can be made.

The power spectral density (PSD) of the mean value of the extracted patch at various equivalence ratios are plotted in Figure 5.15. All equivalence ratio cases

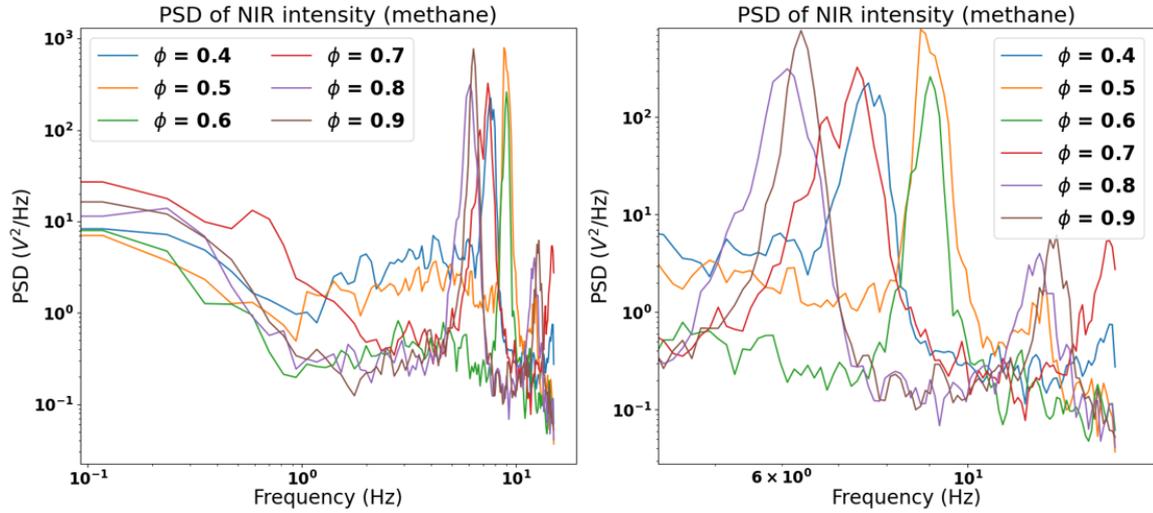


Figure 5.15: PSD of NIR intensity variation in premixed methane flame. Right panel is the enlarged view on the second peak near 11 Hz.

exhibited similar trends. The most energy containing region is at the low frequency part. However, for all cases, a dominant peak is located around frequency of 8 Hz. The zoomed in plot on the right panel of Figure 5.15 also indicate a second weaker peak centred 11 Hz. By observation, the intensity variation frequency is roughly disproportionate to the equivalence ratio. The variation of intensity is due to the intermittent rising of hot combustion product.

The PSD of the extracted length with respect to time of the vertical centreline and horizontal line showed great consistency for all equivalence ratios as shown in Figure 5.16. Both curves peak at 8 Hz, which is in agreement with the PSD of intensity. A second peak locates around 11 Hz. By comparison, the power strength of the peaks, vertical pulsation has greater amplitudes comparing to the horizontal. This is expected because of the buoyancy effect trying to stretch the hot combustion products upwards. The flame wobbles due to the vertical stretching. Previous research have shown that the oscillation of pressure [152], equivalence ratio and chemiluminescence spectral emission [153], have similar frequency spectrum as obtained by monitoring

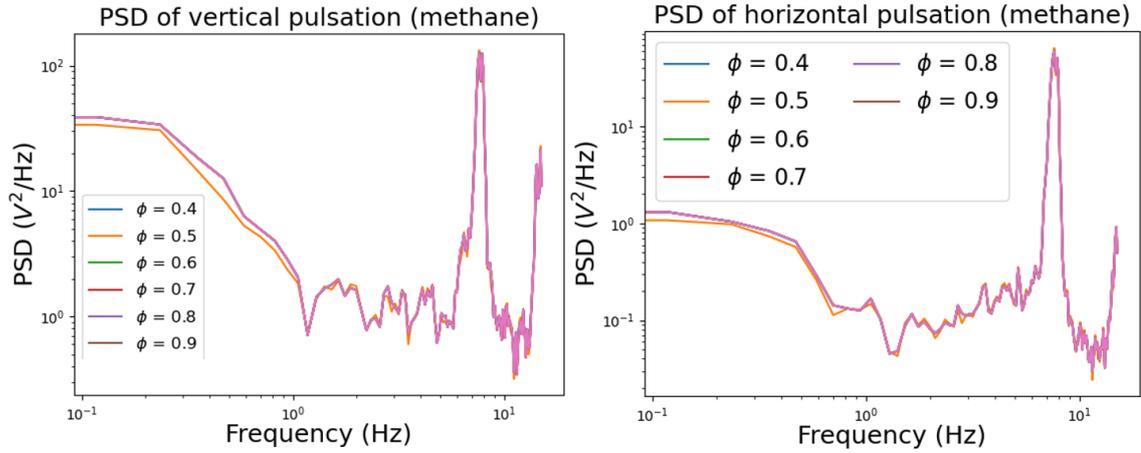


Figure 5.16: PSD of NIR premixed methane flame pulsation. Left: vertical; right: horizontal.

the spectrum of H_2O and CO_2 . [143] also reported the fluctuation in infrared emission is consistent with the simultaneously measure acoustic response. Even though those research focused on the turbulence flames in combustors, it validates the method of monitoring hot combustion products as an effective tool for combustion instrumentation.

In the hydrogen premixed flame case, frequency spectrum is not as so obviously structured as that for methane premixed flames. A predominant peak was only observed at stoichiometric, which centres at 5 Hz. Another less strong peak locates neat 10Hz. At equivalence ratio of 0.5, the intensity variation spectrum spans wider than the rest of the cases, with irregular power distribution over the frequency range from 1 to 10 Hz. Overall, the oscillation are reduced by the decrease in equivalence ratio for hydrogen premixed flames.

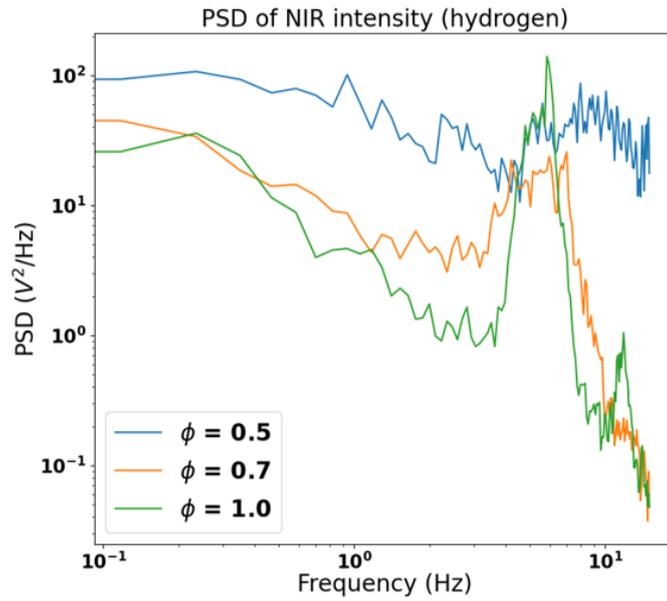


Figure 5.17: PSD of NIR intensity variation premixed hydrogen flame. Blue: $\phi=0.5$, yellow: $\phi=0.7$, green: $\phi=1.0$.

The fluctuation PSD of the extracted centreline and horizontal lines for hydrogen premixed flame shows great consistency for all equivalence ratios. Referring to Figure 5.18, most of the frequency spectral power is contained in the low frequency region. A decay is observed from approximately 0.4 Hz to the end of the spectrum. Observed from the image sequence of premixed hydrogen flames, for the lean premixed cases, the flame tends to stay vertical, and no periodic wobble appeared as in the methane cases. The sporadic wobble may be caused by the ambient air flow or fluctuation from the burner pipe. Only when the equivalence ratio reaches stoichiometric, the flame exhibited periodic stretch and wobble.

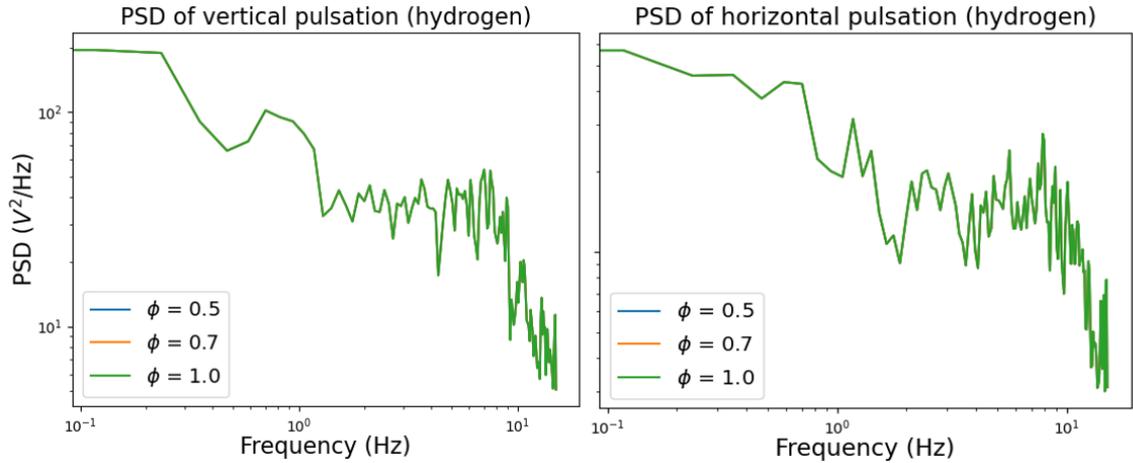


Figure 5.18: PSD of NIR premixed hydrogen flame pulsation. Left: vertical; right: horizontal.

5.5.2 Premixed methane and hydrogen flame evolution with NIR imaging

The evolution of one pulsation cycle of methane flame are plotted in Figure 5.19 and Figure 5.20 at each equivalence ratio. The colour contour is based on the blue channel image, which is normalized by 255. At equivalence ratio of 0.4, the flame shows the weakest NIR emission. The excess of air in the lean mixture suppresses the flame temperature. Near the reaction zone, for equivalence ratio of 0.5, 0.6, 0.7, all revealed strong NIR emission. It should be mention that the pixels are saturated for these three cases. This is compromise for intensity benchmarking as the image sensors dynamic range is limited. For equivalence ratio from 0.8 and 0.9, cooling is observed as there is very small region of saturation or no saturation at all. This may be caused by more intense mixing of the hot combustion products with ambient air due to stretching and wiggling. By visual inspection, the flame has greater stretching and waving as equivalence ratio approaches stoichiometric.

The camera setting ISO for imaging hydrogen flames was set to 100 instead of 400 for methane case, and other settings were kept the same. The NIR radiation

from premixed hydrogen flames is much stronger than that of the methane flames, indicating higher temperature distribution of the flame vicinity. In contrast to the shape methane flame, hydrogen premixed flame did not show significant wobble and necking.

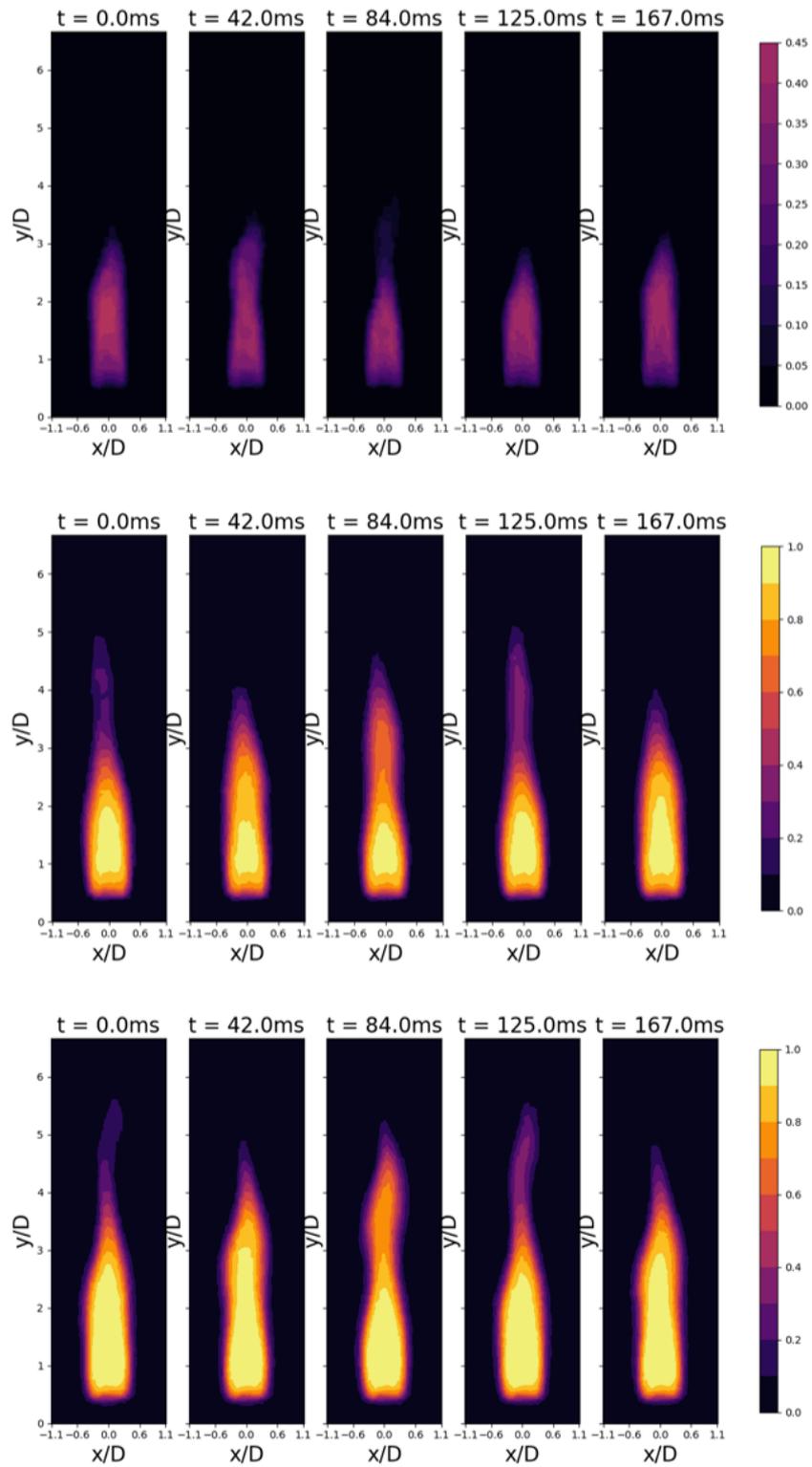


Figure 5.19: Colour contour of premixed methane flames at fuel-lean condition. Top to bottom: $\phi = 0.4, 0.5, 0.6$.

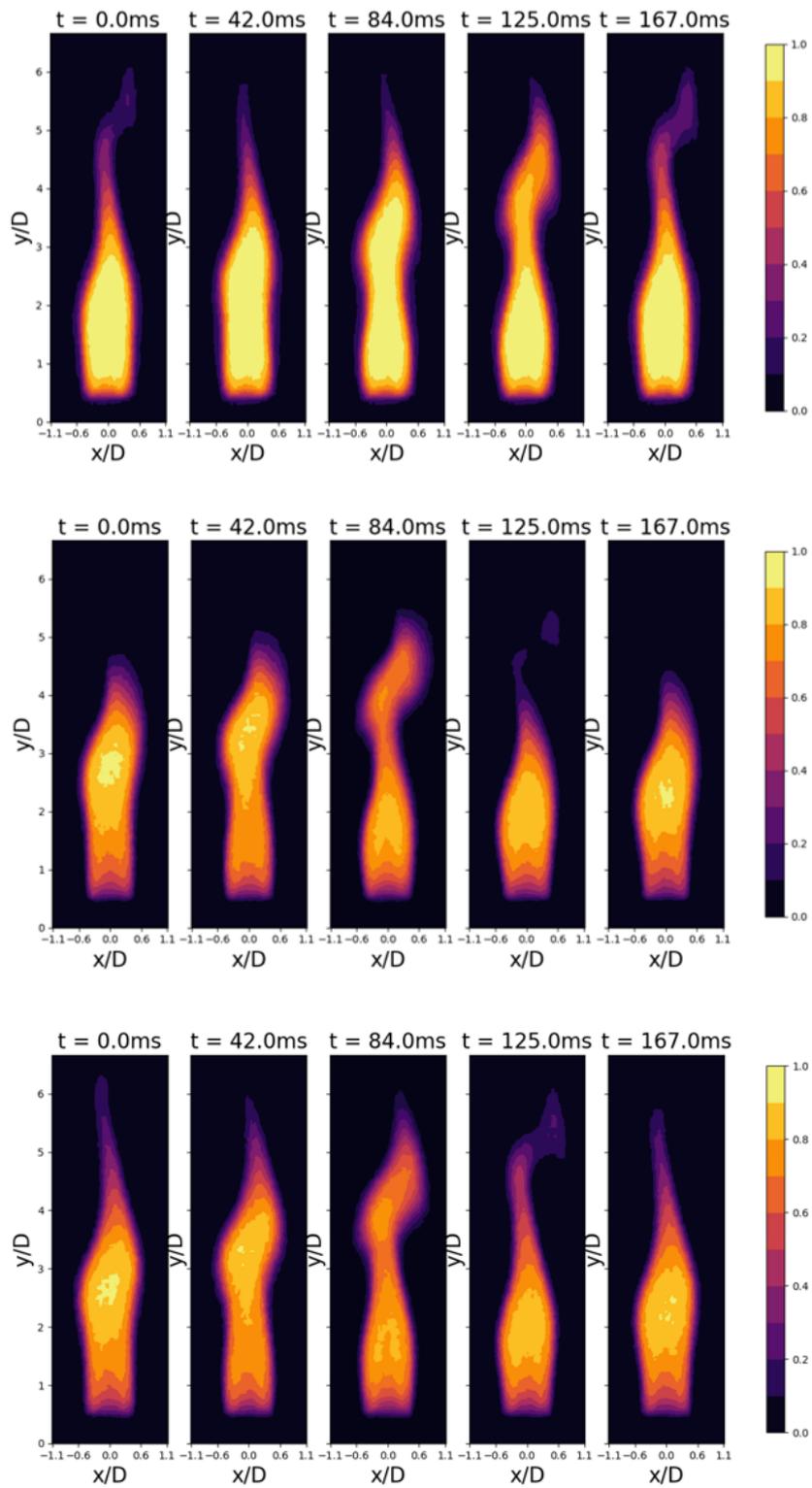


Figure 5.20: Colour contour of premixed methane flames at fuel-lean condition. Top to bottom: $\phi = 0.7, 0.8, 0.9$.

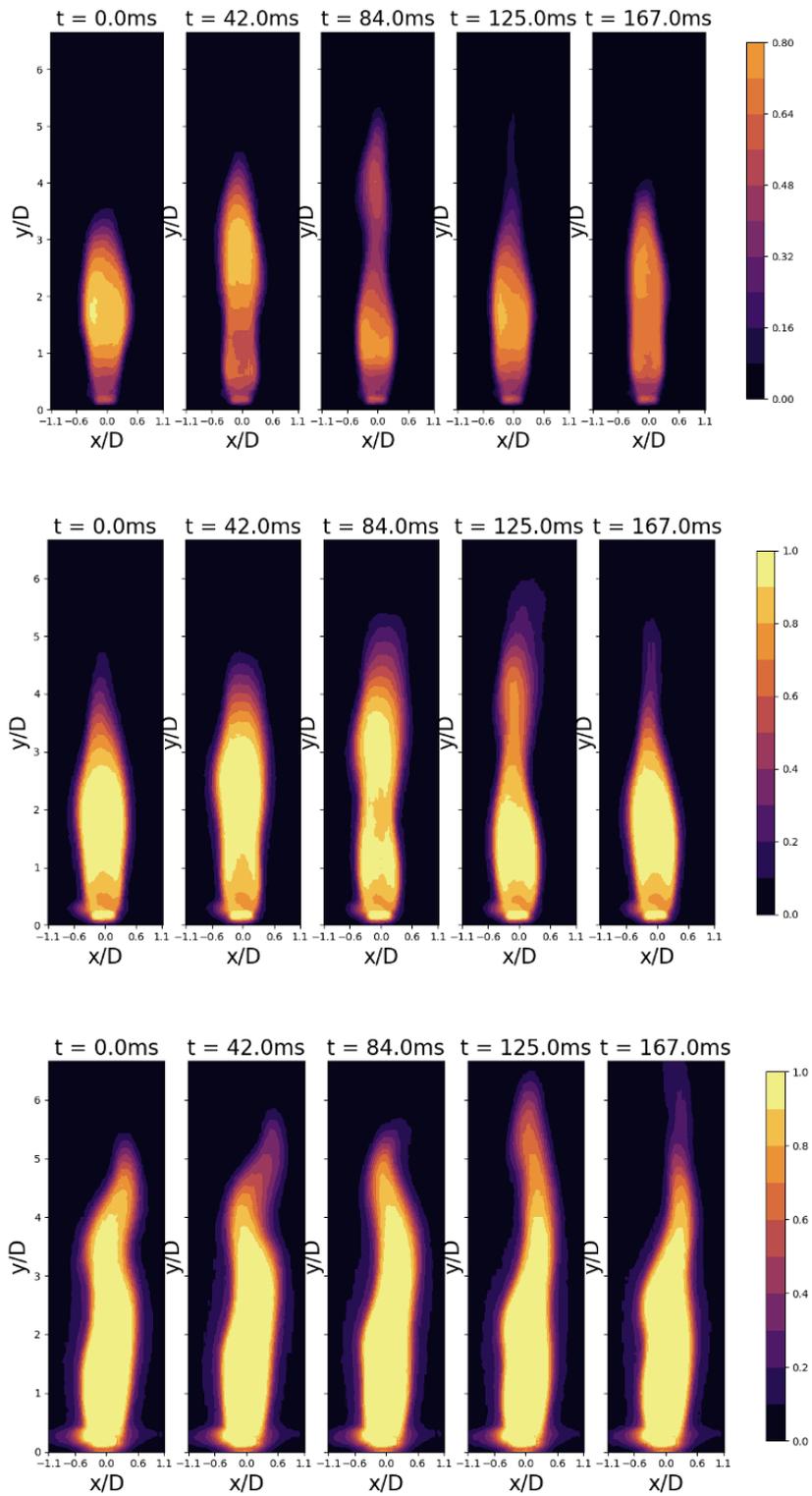


Figure 5.21: Colour contour of premixed hydrogen flames at fuel-lean condition. Top to bottom: $\phi = 0.5, 0.7, 1.0$.

The time averaged intensity distribution of premixed methane flames at equivalence ratio 0.7, 0.8, 0.9 over 4 seconds (120 frames) is illustrated in Figure 5.22, which is superimposed by the outline of visible flame cone. It can be seen that the high intensity region at the centre of the contour plot shrinks with the increase of equivalence ratio. The high intensity outline corresponds to the CH^* and C_2 radicals which occurs at the boundary of reaction. Research have reported that these two radical are better correlated with temperature profile that with OH^* in some situations [154]. This is in agreement with the overlapping of cone tips and high intensity contour region. The layered protruding contour from the flame root to the high intensity region reveals the preheating of the fuel and air mixture before entering the reaction zone. Moreover, region of high NIR intensity region stretches along with the flame cone, with respect to the equivalence ratio.

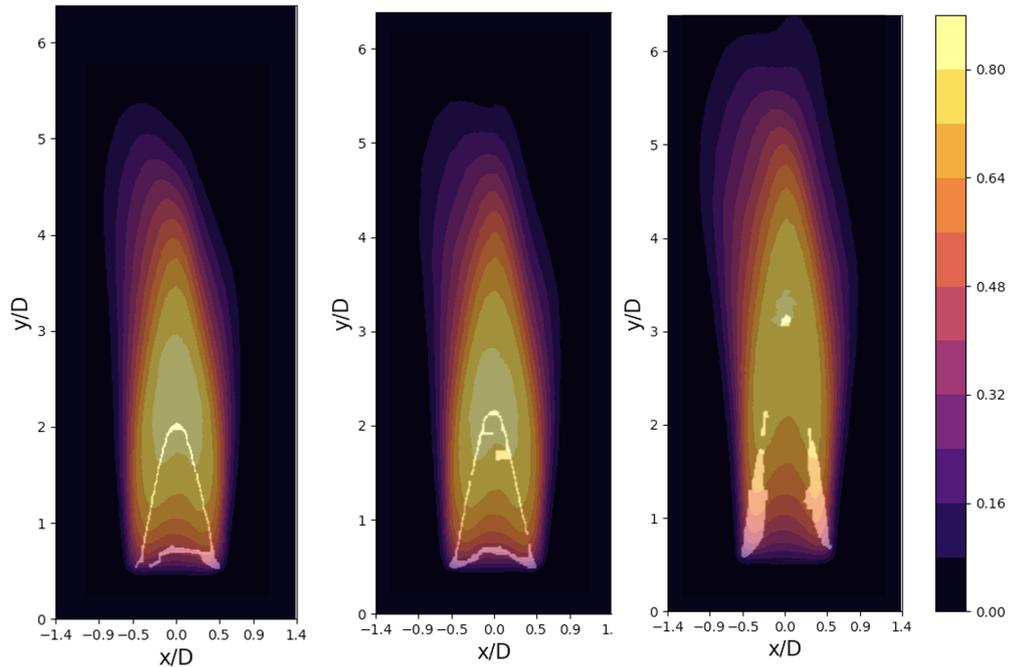


Figure 5.22: Time averaged premixed methane flames for 120 frames with CH^* and C_2^* outline. Left to right: $\phi = 0.7, 0.8, 0.9$.

5.6 Usage of NIR imaging in solid fuel combustion

5.6.1 Background

As natural wood processed and reclaimed wood is extensively used in construction and furniture of domestic household as well as in public environment, the fire safety regarding to these materials is of great importance. Wood combustion is a highly complex process, which involves different substances in the wood subjected to thermal decomposition that active chemical reaction will eventually lead to combustion.

Wood is mainly made of lignin, polyoses and cellulose, which is a kind of polymer but made by nature. These are intertwined together in a cellular form during the plants growing process. Under thermal decomposition, wood materials can undergo different combustion states, such as smoking, pluming, smouldering and flame-out. Figure 5.23 illustrates the components of plant biomass and wood, where cellulose, polyoses (hemicellulose) and lignin contributes the major weight of the wood.

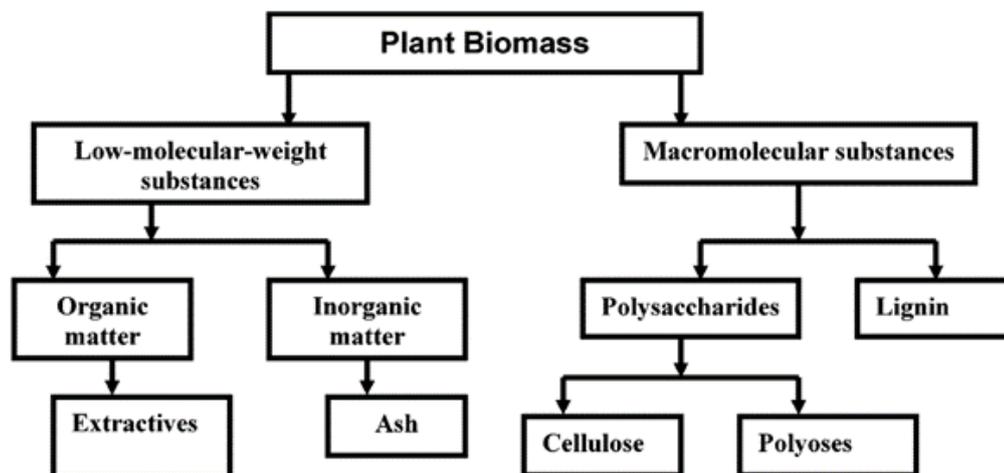


Figure 5.23: Illustration for biomass and wood constituents [155].

Cellulose accounts for approximately 40-50 percent of the weight, which also responsible for the hardness of the material [156]. Figure 5.24 is the basic constituent

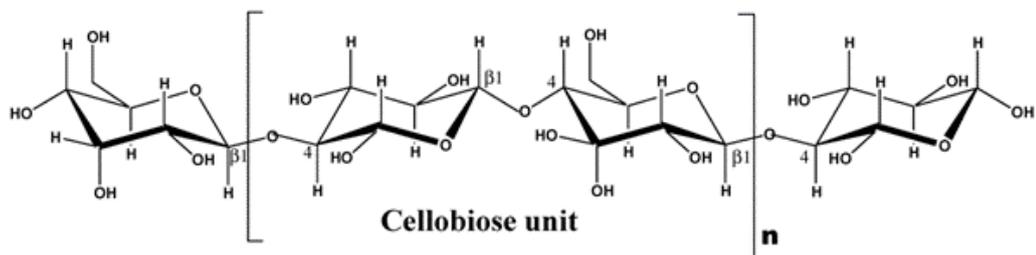


Figure 5.24: Molecular structure of cellulose polymer [157].

of cellulose polymer, called cellobiose unit. The cellobiose unit is a combination of two glucose anhydride, which is formed by breaking the bond of water to it. The cellobiose unit is replicated 5000-10000 times to form a long chain. Such chains will then be bonded by hydrogen bonds. Afterwards, these chains intertwine to form the fibres [157].

Another main constituent is the hemicellulose or polysoses, which can take up to 35% of the weight [156]. The building block of the hemicellulose is heteropolysaccharide, which consists of several polymerized monosaccharides. The repetition of the units is around approximately 150, which is much smaller than that of the cellulose. Simpler structure makes it less stable than the cellulose with a decomposition temperature of 200-260 Celsius which is also easier to produce volatiles [158].

The third element is lignin, which contributes to 16-33 percent of the weight depending on the type of the wood [159]. Unlike cellulose and hemicellulose, the structure of lignin is random which is made of cross-linked resin. Because of its amorphous structure, many possible linkages between different building units of other constituents exists. Lignin acts as a glue to bind other fibrous constituents [160]. The decomposition temperature of lignin has a range of 280-500 Celsius [158]. The decomposition of lignin produces liquid products called pyroligneous acid, which has

about 20% aqueous constituents and around 15% tar. In addition, it also produces flammable gases, such as methane, ethane and carbon monoxide, which in total accounts for 10% of weight of lignin [155].

The chemistry involved in wood combustion process consists of thousands of reactions, where as much as hundreds of chemical substances interact [161]. Such thermal decomposition is called pyrolysis, which in short wood is decomposed to combustible gases and consumed by combining with oxygen, leaving only solid charcoal. The mechanism of pyrolysis is significantly dependent on temperature. A reported critical temperature of 300 Celsius, below or above which the pyrolysis will bias towards charring and gasification. Wood fire is always orange because of the soot formation by the amalgamation of free carbons from the charring of the wood. Near the root of the flame, there can be bluish flame from carbon monoxide, methane and ethane. The lack of oxygen in the wood as it is charring promotes carbon monoxide generation.

The sustained flaming of wood is a self-supplying loop as illustrated in Figure 5.25, where the wood is heated by the flame to generate combustible gases through pyrolysis. Those gases mix with air and then transported into the flame by buoyancy. Fire will sustain and propagate by heating the adjacent virgin wood in order to produce enough gases for the flame. The heating of the wood mainly takes two manners: convection and radiation; the dominance of which varies from case to case, depending on the flame scale and the setup condition.

The scale of the fire is an essential factor for the dominance of the type of heat transfer to sustain the cyclic combustion process. In the case of wall fire, radiation grows to become stronger than convective heat flux as the flame increase in size [162].The radiation of a wood fire is contributed from particulate, condensed tar,

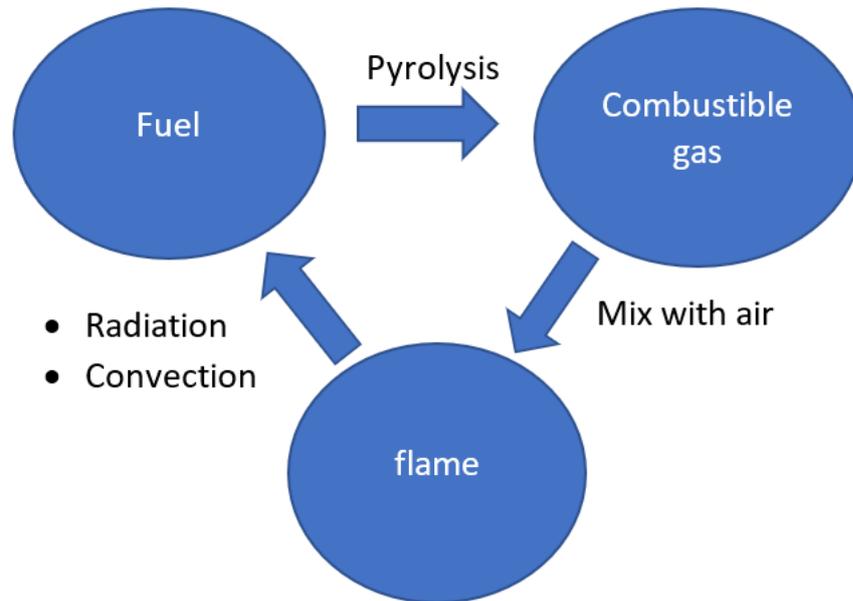


Figure 5.25: The loop of self-sustained wood combustion.

water as wood moisture and other combustion products of gases from pyrolysis which are mainly CO_2 and H_2O . They emit radiation in different part of the spectrum. Solid particles usually have a radiation curve resembles the blackbody radiation, having an emission over a wide spectrum. Whereas the emission of CO_2 and H_2O falls into the infrared spectrum. The radiation from both gaseous products and solid particulate also depends on their concentration [163]. In the case of pool fire, the heat transfer becomes radiation-lead for pan size larger than 20 cm [164]. If the flame scale is further increased, the flame temperature will rise because of soot blockage. The prolific soot prevents effective radiation loss therefore increases flame temperature [163]. As a result, the energy loss due to radiation is reduced. Figure 5.26 shows the radiative energy starts to drop dramatically after the pan diameter reached approximately 5 meters. Apart from the material properties of the wood, geometric properties can play a vital role in promoting or suppressing fire propagation. Heat transfer by convection is greatly affected by the geometry setup of the scenarios. Basically, convection is

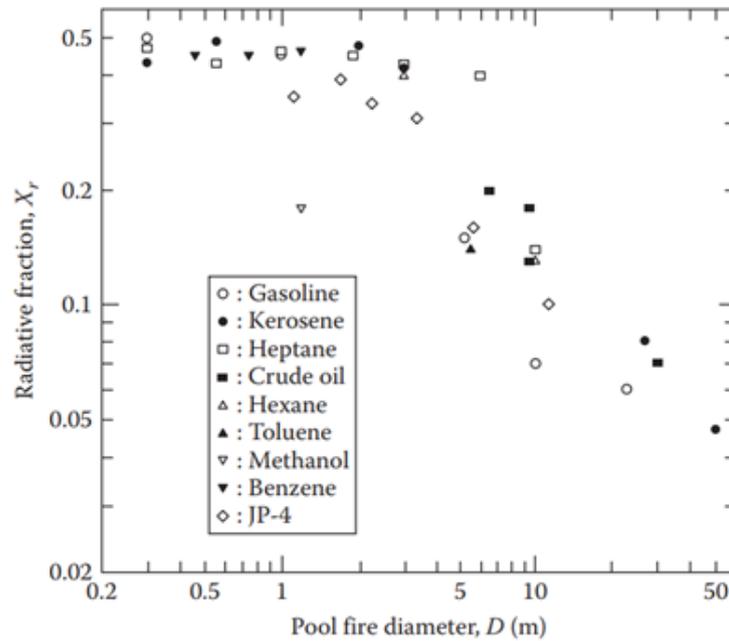


Figure 5.26: Radiation energy versus flame size for pool fire [165].

the interaction between flow field and energy transport. Hot fluid will rise due to buoyancy effect resulted from change in density. Hence, the combustion of wood is more likely to sustain as the unburn fuel is in the downstream of the hot fluid, which helps to preheat for drying and then pyrolysis. In contrast, heat transfer by radiation is in the electromagnetic form, and is omni-directional.

Smouldering can happen in a low temperature situation. Temperature drop or a change in oxygen supply can cause can invert flaming into smouldering or vice versa. In the smouldering process, significant production of carbon monoxide is observed, taking more than 10% of the total fuel weight. The sustenance of smouldering relies on air diffused into the reaction zone, as illustrated in Figure 5.27. The reaction rate is dictated by the supply of oxygen within a certain limit, after which the excess of oxygen initiate flaming.

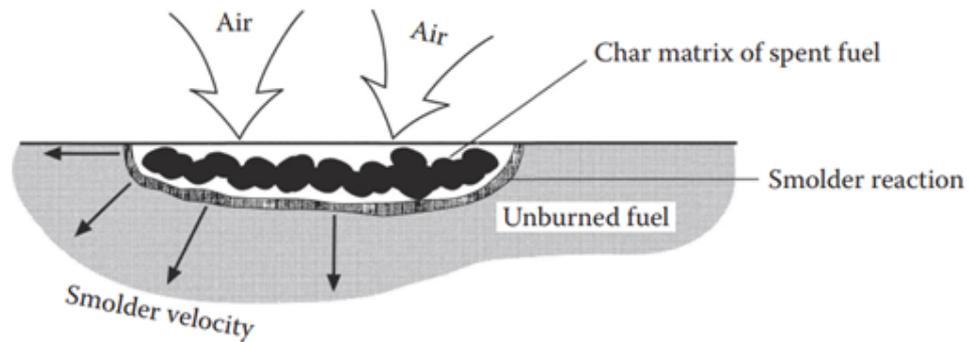


Figure 5.27: Schematic diagram of smouldering in wood [1].

5.6.2 Fire propagation on wood with combined imaging system

In the test case of this research, flame propagation on inclined oak wood is studied with synchronized imaging technique combining visible, NIR and schlieren. The dominance of convective preheating was confirmed to be the main mechanism for prolonged flame propagation. Even though the wood rod samples are pre-dried to have consistent moisture content, two other factors can significantly influence the outcome of the test, namely wood density variations in individual samples and grain orientations. Qualitative analysis of the physical phenomenon is based on repetitive experiments to mitigate randomness. The mechanism of fire propagation in this test setup will be investigated.

The schematic diagram of the experiment setup is shown in Figure 5.28. The cylindrical wood sample is clamped onto a stand, where the angle of inclination θ can be controlled. Cylindrical oak wood was used, which has a length of 30 cm and cross-sectional diameter of 9 mm. A pilot flame was used for ignition at the lower end. For a controlled ignition, the pilot flame was produced at fixed equivalence ratio and volumetric flow rate. The duration for ignition is 20 seconds for all tests. The multi-spectrum imaging was combined with schlieren visualization for

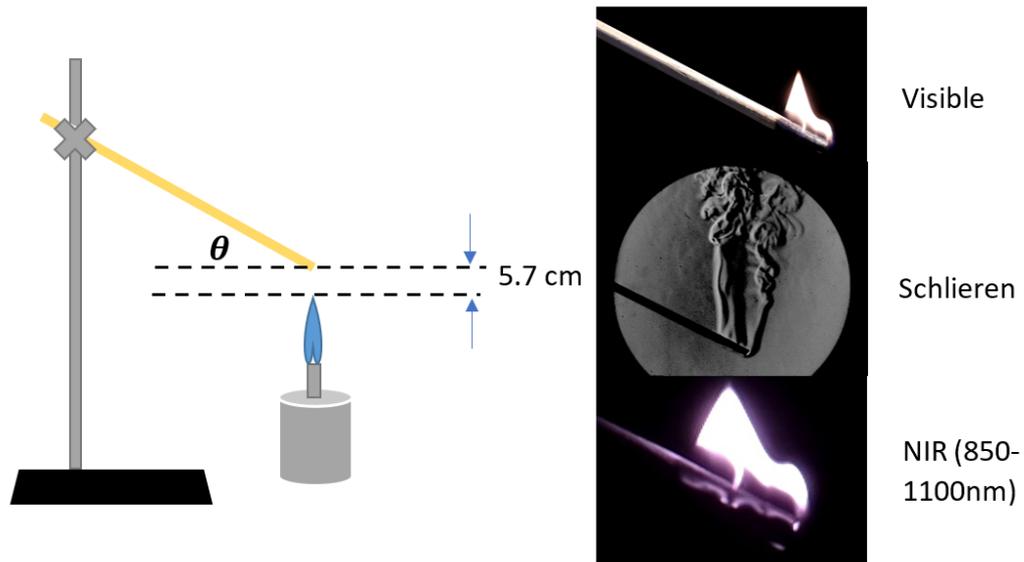


Figure 5.28: Schematic diagram of the experiment setup.

combustion diagnostics. Specifically, the visible images were used to observe charring; the schlieren visualized the hot flow surrounding the wood sample; finally, the NIR sensors were able to locate the regions possibly undergone pyrolysis as it can see radiation of object with temperature as low as 375 Celsius.

The layout of the imaging system is shown in Figure 5.29. In the Z-type schlieren, the test space is in the middle of the parallel light path, which has double the focal length of the parabolic mirror (3 metres). A stereo raspberry pi camera system with a visible and short banded NIR sensor. The pi system is placed along the parallel path but slightly below the line of sight aiming upwards to avoid blocking the schlieren light path. Testing sets were made at inclination θ of 0, 20, 25, and 30 degrees. Each test case was repeated for 10 times to mitigate random influence.

Table 5.1: Summary of fire propagation in test cases at different inclinations.

Angle θ	Average propagated length (cm)	Indefinite propagation probability
20°	4.89	0%
25°	4.85	36%
30°	5.49	64%

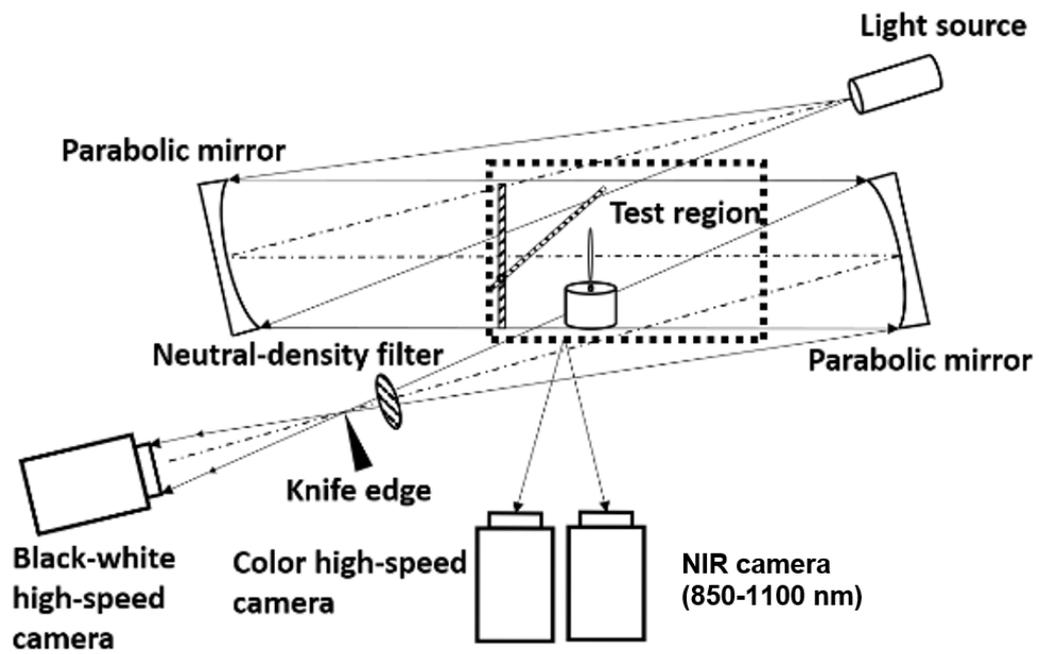


Figure 5.29: Layout of combined imaging system.

The overall trend of all tested groups was summarized in Table.5.1. Because at $\theta = 0^\circ$, the flame cannot propagate lengthwise and extinguish very soon, therefore excluded from the summary. In general, the burning lifetime and propagated length vary greatly. Such variation mainly comes from the inconsistency of the wood material, especially the density, grain alignment with the length and porousness. The averaged propagated length has a significant increase at $\theta = 30^\circ$ because of the longer preheating length. Therefore, it is more likely for the self-supplying process to survive. Limited by the length of the wood sample, cases of indefinite propagation were counted. In total, 64% percent of all test cases of $\theta = 30^\circ$ propagated through the whole length. The grain orientation can enhance the combustion. For example, in cases that the grains lead to upstream, the combustible gases from pyrolysis are naturally transported to the reaction zone. Whereas, if the grains lead to the lower end, from which the gases are expelled and then ignited without contributing to the pre-heating.

5.6.2.1 Some observations and discussion

For illustration, the case with $\theta = 30^\circ$ is provided in Figure 5.30. The combined visualization of flame propagation in both visible vs schlieren and NIR vs schlieren is presented. The schlieren helps to visualize the hot air flow generated by the flame. The NIR imaging enables the visualization of pyrolysis area.

By observation, the hot flow caused by the flame encompasses a longer length, which pre-heat the downstream virgin wood by convection. The larger the inclination angle θ , the longer the pre-heating. As a result, the fire is more likely to propagated. The right column shows the heated wood at time instant of 30 seconds apart. The glowing regions on the burning wood indicates a temperature zone above 375 Celsius. This is the temperature zone that the wood is subjected to significant mass loss [166]. The cell structure of the material is degraded, which leads to dimensional change. More specifically, the crystal structure of cellulose and lignin will be completely destroyed rendering an undefined structure [167]. This destruction of the polymer structure is accompanied with carbonization within 300-500 Celsius. As the temperature increased, the developed carbon structure will produce bio-morphous carbons [168]. The shrinkage of the wood dimension is obvious as it burns. This can be seen at $t = 60s$ of Figure 5.30 that the lower end is much thinner and slightly bent upwards. The use of NIR sensors helps to identify the temperature regions corresponding to different stages of pyrolysis, which can be an effective tool for combustion diagnostic inference.

The combined imaging technique reveals the different wood combustion regimes. Firstly, the NIR imaging reveals the pyrolysis zone at a temperature above 375 Celsius, as illustrated by region As in Figure 5.31. The decomposition of lignin and polyoses and degradation of the material ultrastructure seem to coexist. In the left panel, the

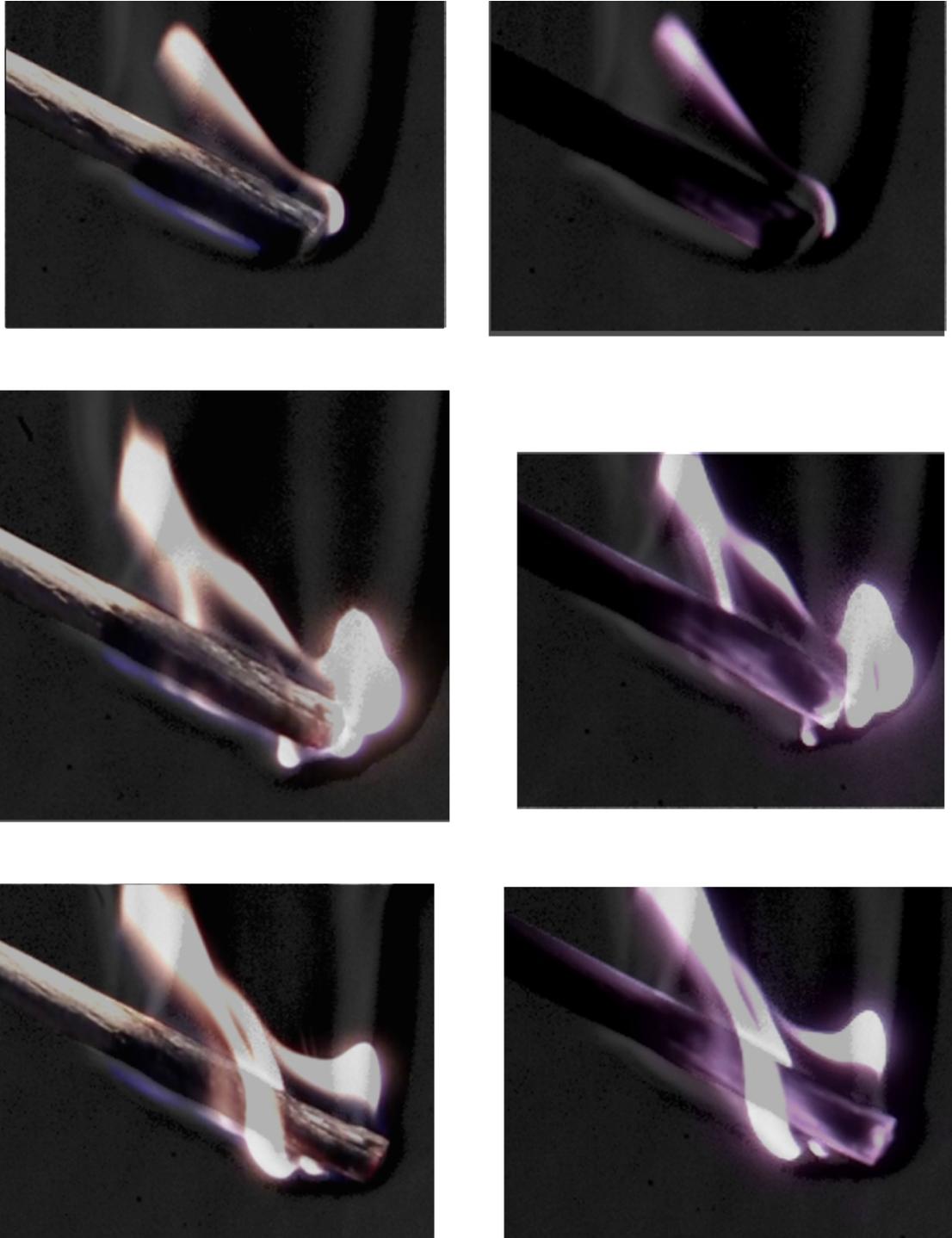


Figure 5.30: Combined imaging with NIR, visible and schlieren for $\theta = 30^\circ$. Top to bottom row: $t = 0\text{s}$; $t = 30\text{s}$; $t = 60\text{s}$.

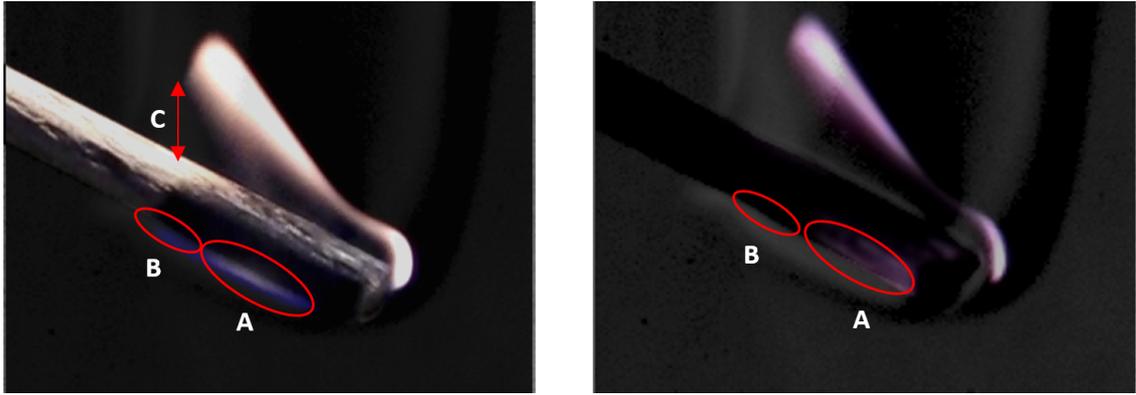


Figure 5.31: Visualization of different regimes of wood combustion. Region A: >375 Celcius \rightarrow degradation and carbonization \rightarrow free carbon \rightarrow orange flame; region B: <375 Celcius \rightarrow Hemicellulose and lignin decomposition only \rightarrow mostly combustible gases \rightarrow blue premixed flame; C: flame lift-off.

colour of flame at region A is mixed blue with oranges. The orange colour should be caused by the generated free carbon from polymer structural destruction, which entail a temperature range of 300-500 Celsius. This overlaps with the capability of the NIR imaging on the right panel at region A.

In region B, it is highly likely that the polyoses and lignin undergo thermal decomposition, which produce most of the gas generation. The temperature is likely to have a temperature range of 200 to 300 Celsius. This is deduced from the non-glowing part but still within the hot air flow field. Moreover, it can also be inferred by the absence of soot emission in the region A of the left panel with only bluish flame presented. Because sooty flames come from the free carbon released after the structural deconstruction at temperature above 350 Celsius, which is exemplified by region A. The inference is in consensus with previous studies, which stated the temperature for the pyrolysis of lignin and polyoses to produce combustible gases are 260 [169] and 240 Celsius [170]. The flame maximum lift-off is indicated by C. The flame lift-off the most significant at the propagation front in all test cases. A possible mechanism to explain this phenomenon is that the excess of produced combustible

gases at the front in combination with buoyancy effect from the combustion increases the flow velocity, which then gives rise to the lift-off. The lift-off is observed in all cases when there is a corresponding underneath bluish flame (as in region B).

The temperature gradient along A and B from IR imaging, indicates the different stages of pyrolysis the wood is undergone. The closer to the high front of the flame, the lower the temperature, but more combustible gases generated. The closer to the lower end, the higher the temperature. Because the fuel has already been consumed for propagation, the lignin and polyoses left inside has a lower rate of conversion to gases. Therefore, the lift-off is smaller. The lift-off height along the length seems disproportionate to the temperature gradient long the length. The lift-off seems to be larger as the inclination θ is larger. This is expected as θ increases, more pre-heated length immersed in the hot flow, hence more combustible gases produced.

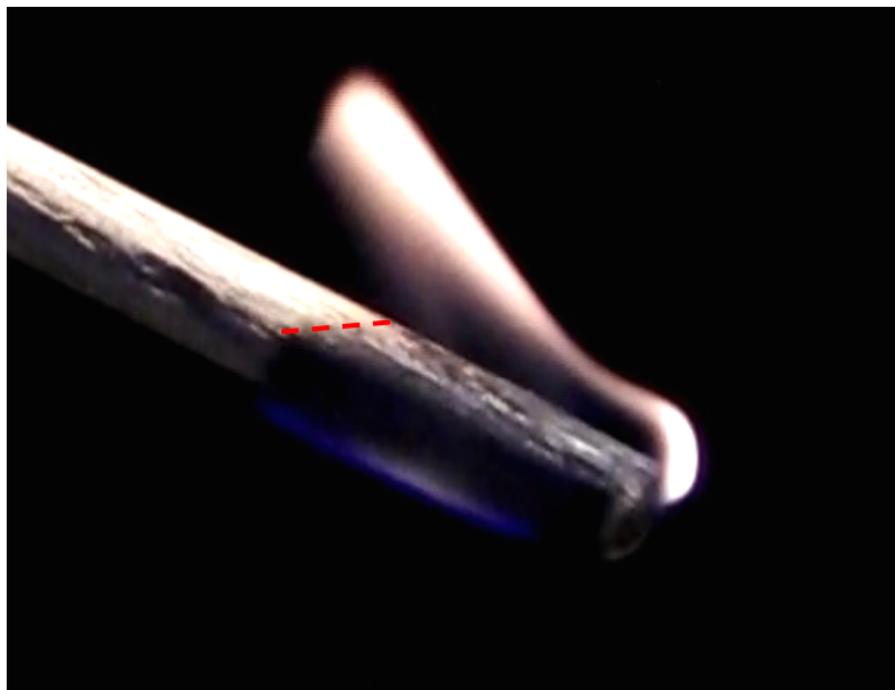


Figure 5.32: Illustration for the dominance of convection or radiation.

Convection or radiation dominant for fire propagation in this case? The main heat

transfer in this case is convection as discussed previously. Despite of contribution by radiant heat transfer from the luminous flame, the drive for flame propagation is by convective heat transfer. Studies have shown that the radiation will become dominant when the flame scale reaches 0.2-0.3 m and beyond [171]. In this research, in the most violent case, the flame was no larger than about 10cm. Whether the convection or radiant heat transfer is dominant may be deduced from the images. Referring to Figure 5.32, the charring boundary is highlighted by the dash line for that instance. As the flame propagates, the charring boundary roughly stays at such an orientation. This seems to be the hot flow separation region, where the hot mixture is then transported by buoyancy to form the diffusion flame. The flow separation leaves a lower temperature region above the fuel. As the hot flow immerses more of the upstream wood, the whole cross-section of the wood will be burnt. It is difficult to qualitative measure the significance of convection and radiation. The radiation is believed to have some contribution at much lesser degree comparing to that of convection. The evidence is the orientation of the charring boundary (highlighted by dashed line): if the radiation is significant, the region above dashed line in Figure 5.32 will most likely be burnt into char rather than leaving a boundary illustrated by the dash line.

5.6.3 Fire propagation on wood with combined imaging system with fan

In this test case, a fan and a diffuser were applied to the test setup for $\theta = 30^\circ$, as shown in Figure 5.33. The fan has adjustable speeds at 5 levels: 1.1, 2.1, 3.2, 4.2, 5.3 m/s . The fan is turned on right after the pilot ignition is turned off. The aim is to study the effect of controlled air flow on the flame propagation.

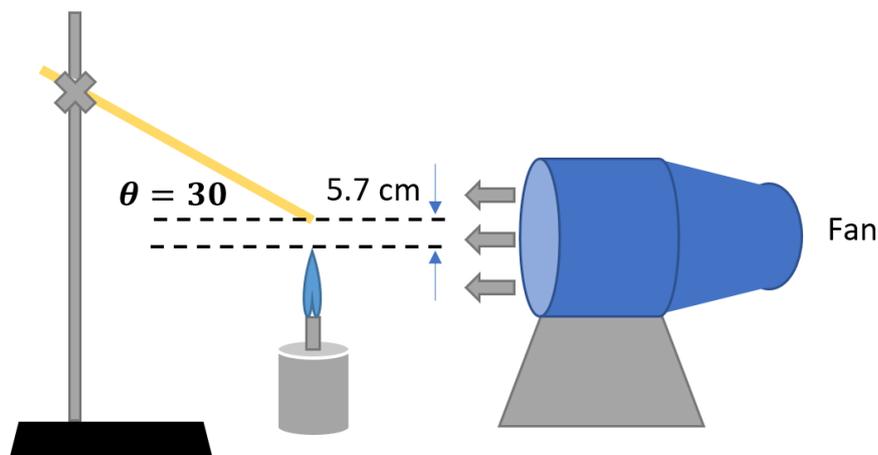


Figure 5.33: 2-D Schematic diagram of the experiment setup with fan.

At each wind speed, repetitive tests were established to eliminate randomness. Table.5.2 gives the summary of the test case. At speed 1, all test cases can propagate along the whole length of the sample. As the speed increase, the rate of indefinite propagation drops. As maximum wind speed, the preheating is disturbed strongly so that the flame extinguishes very soon. For the measure of average time taken to reach 8 cm, only cases that the flame successfully propagated over 8 cm were considered. It is very clear that at speed 1, the applied wind greatly enhances propagation and combustion. Not only all test cases are can indefinitely propagate, but the propagation speed is notably larger comparing to the no fan case. At higher wind speed, even though the propagation is faster in some cases, in more cases the flame extinguishes

Table 5.2: Summary of fire propagation in test cases at different inclinations.

$\theta = 30^\circ$	Indefinite propagation probability	Average time taken to reach 8 cm (s)	phenomenon
No fan	65%	165	-
Speed 1	100%	86	violent flaming, much longer preheating
Speed 2	50%	97	violent flaming, cooling observed, flame propagates underneath starts to appear
Speed 3	55%	73	More cases with flame only propagate long the underneath
Speed 4	17%	81	Mostly flame stays underneath the sample and very small
Speed 5	0%	-	Significant cooling, extinguish shortly

due to the cooling and shifted preheating by the wind. Referring to Figure 5.34, the flame at the same time instant after ignition turned off, demonstrated the difference in combustion conditions. The schlieren reveals the effect of wind speed on preheating. More concretely, at speed 1, the wind helps to bend the hot air flow towards upstream fuel, which increases the preheating length. In contrast, at speed 4, the wind speed is too high that the hot air flow is soon disturbed and dissipated to the lower left region, decreasing the contact of preheating. Despite the hot air is still exist underneath the sample, the resident time of the hot air is very small to effectively preheat the unburnt fuel upstream. Moreover, the underneath thin flame as shown in Figure 5.34 (d) started to appear at speed 2 and became dominant for higher speed cases, which reveals the cooling effect.

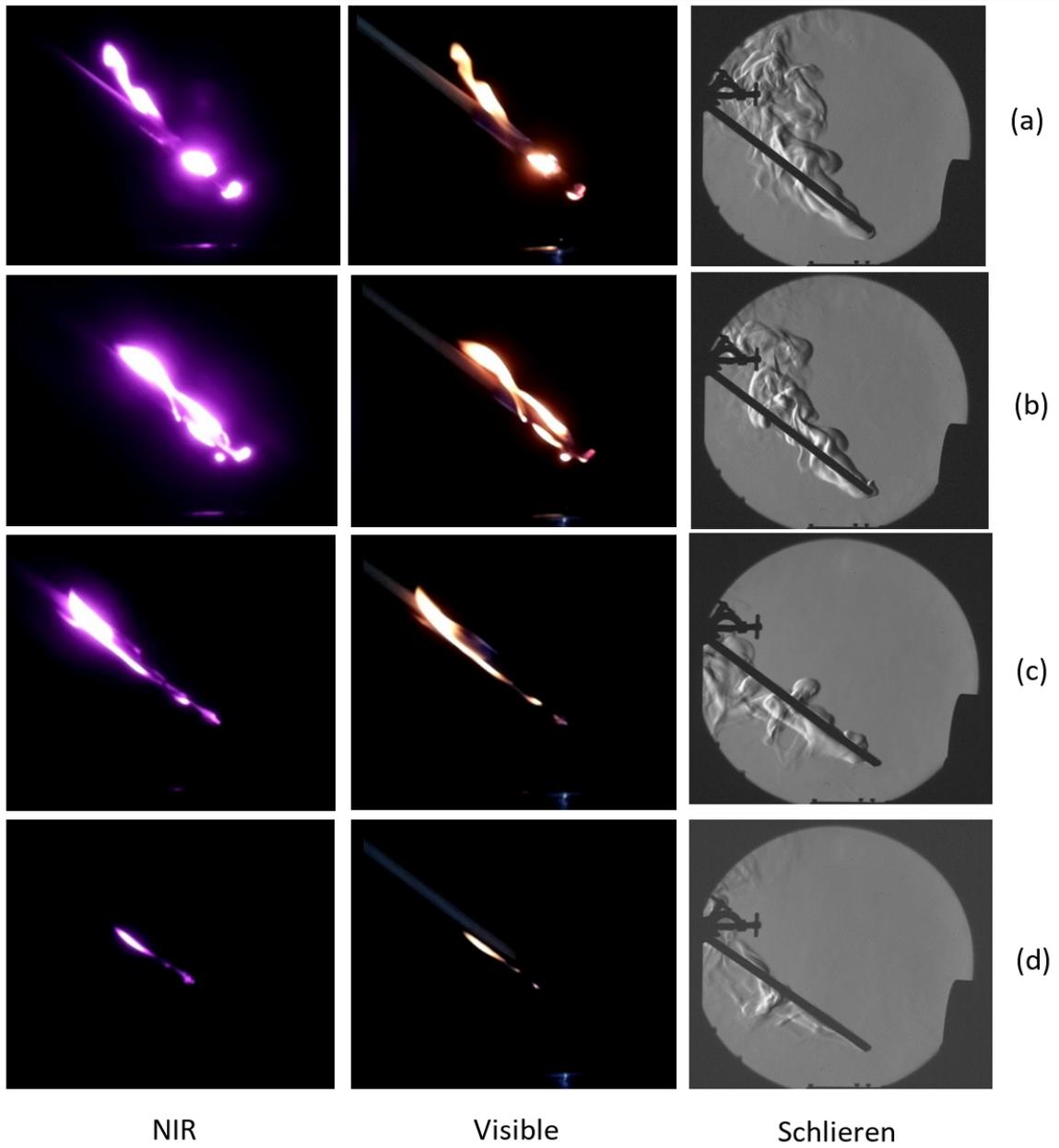


Figure 5.34: Combined imaging at different fan speeds. (a): speed 1, (b): speed 2, (c): speed 3, (d): speed 4.

The cooling of the applied wind can be quantified through schlieren imaging. At the core of schlieren imaging, the contrast image of the flow field is obtained by the using a knife edge blocking the deflected rays. Because light will deflect if passing through medium with different density, which gives different refractive index n . The deflected angle ε with respect to the optical axis is proportionate to the change of refractive index. As a result, different density regions in a schlieren image have corresponding local change of intensity ΔI comparing to the ambient undisturbed region. The relationship can be established as in Eq(5.2).

$$\frac{\Delta I}{I} = \frac{f_2}{a} \int_{\zeta_2}^{\zeta_1} \frac{1}{n} \frac{\partial n}{\partial y} dz \quad (5.2)$$

$$n - 1 = K\rho \quad (5.3)$$

Combining Gladstone-Dale relation in Eq(5.3), which is the expression to calculate density by refractive index, the ratio of intensity change, and undisturbed intensity can be expressed as Eq(5.4) [172]:

$$\frac{\Delta I}{I} = \frac{Kf_2}{a} \int_{\zeta_2}^{\zeta_1} \frac{\partial \rho}{\partial y} dz \quad (5.4)$$

K is Gladstone-Dale constant. f_2 is focal length of the second parabolic mirror, a is the height of light source image at the focal plane of the second parabolic mirror. ζ_1 and ζ_2 are the coordinates, the difference of which is the depth of test area along the optical axis. As shown in the schematic diagram of Figure 5.35, the deflection angle ε can be positive or negative with respect to the coordinate system. Depending on the orientation of the knife edge, positive gives rise to the increase in intensity, while negative gives decrease in intensity as these deflected rays are blocked.

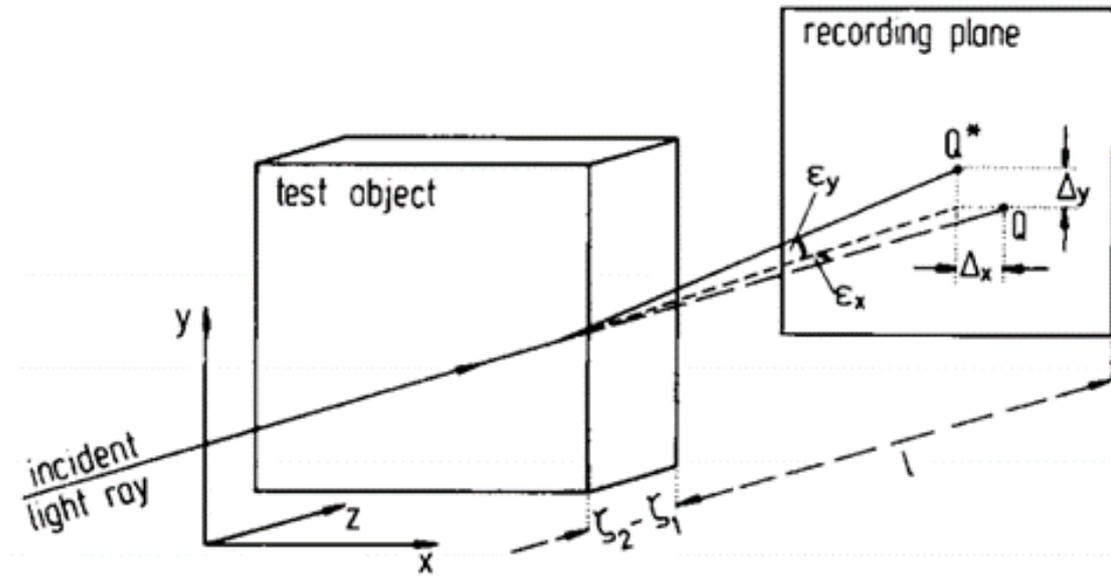


Figure 5.35: Illustration of deflected light rays caused by density change in schlieren system.

The calculation of local density is possible but requires rigorous calibration and setup. However, the change in density is easily obtainable according to Eq(5.4), where change in density is proportional to the change in intensity. The change of density is coupled with the significance of cooling from applied wind. To this end, a patch of pixels in the underneath of the sample where the wood is undergone pyrolysis is monitored over 60 frames of one second for each case. Since there can be positive and negative change in intensity, the absolute value is calculated. Finally, the spatial average is taken for each patch.

$$\Delta I = |I_{patch} - avg(I_{ambience})| \quad (5.5)$$

The extracted patch locations and processed results are illustrated in Figure 5.36. The average change in intensity is the largest at speed 1, while it is very similar for higher wind speed cases. At speed 1, the wind is not strong enough to cool the hot

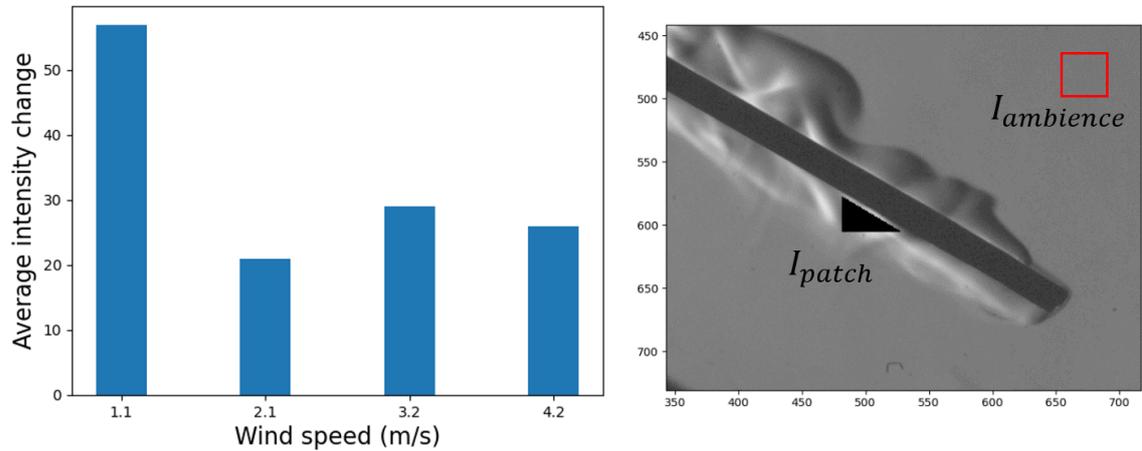


Figure 5.36: Illustration of patch intensity change from schlieren image. Left: average intensity change ΔI ; right: patches for monitoring.

air but to enhance the combustion by bending the hot air towards the upstream fuel. As a result, the hotter the air flow, the larger the difference in intensity and then the larger change in density. At higher wind speeds, the hot air from wood combustion is cooled by the rapid mixing with ambient air. The reduced temperature gives less change in density. Therefore, the refractive index is close to that of the ambient air. The change in intensity is smaller. In addition, the histogram of intensity change of the monitored patch over the period is plotted in Figure 5.37. It is obvious that at speed 1 larger change in density because of enhanced combustion is evident. In contrast, at higher speeds the histogram has a peak around 20, showing the cooling by wind to prohibit wood combustion in this case.

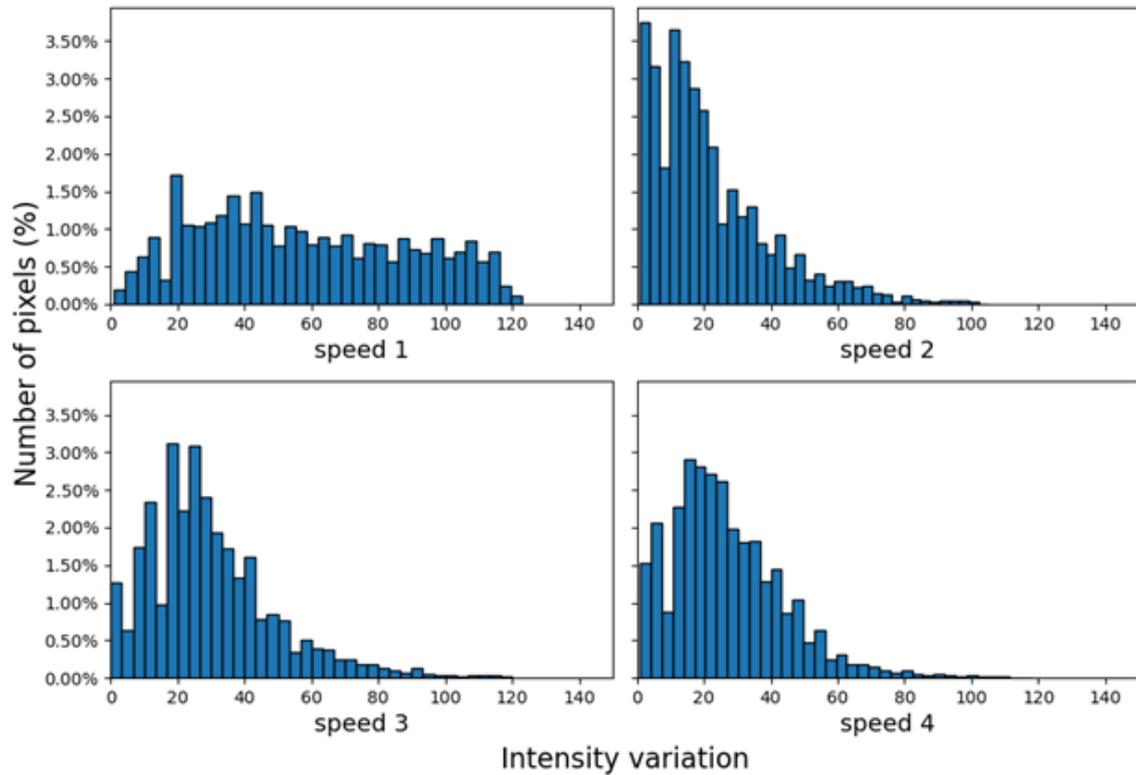


Figure 5.37: Histogram of pixel intensities in the extracted patch underneath the wood sample over 60 frames for each speed.

5.7 Summary

The modified sensor is capable of accurate chemiluminescence instrumentation with improved spectral response. It also has its limitations. For example, its nonlinear spectral response can limit its use to low temporal resolution situations, because the radiation intensity is much weaker in narrow-band imaging. In addition, for absolute spectral irradiance measurement, the sensor requires accurate calibration for benchmarking.

The NIR sensitivity of the silicon-base CMOS sensor is exploited for the imaging of hot combustion products of H_2O and CO_2 . As previously discussed, the intensity is predominantly from H_2O . The fluctuation of the flame as well as the interaction

of the flame and its combustion product can be observed and quantified. From the PSD, methane flames displayed strong fluctuations centred at 8 Hz; while for hydrogen flame, periodicity is only observed when the equivalence ratio reached stoichiometric. In addition, the hydrogen flames seem to be more sensitive to perturbations than methane flames. The immediate visualisation offered by the NIR capability can be used for premixed combustion of syn-gas, especially for the visualisation of regions of radiative heat loss.

Finally, the NIR thermal capability not only provides a handy visualisation of hot area, but can also be used for two-colour pyrometry with extended measurable range comparing to common visible camera.

Chapter 6

Conclusion and future work

6.1 Conclusion

This research explored the use of low-cost near infrared image sensors for fire detection and combustion diagnostics. The flame as a combustion phenomenon is a complex exothermic chemical and physical process, where molecules disintegrate and recombine to form new products together with the release of energy. Alongside this process, radiation is produced over a wide spectrum from UV to infrared. Some excited intermediates output radiations at specific spectrum with unique peaks. Other end products, such as carbon (*soot*), CO_2 , and H_2O dominate the infrared spectrum.

The use of near infrared spectrum (700-1100 nm) for flame and combustion study has always been overlooked, especially in the case of passive imaging. Two possible reasons may have caused the situation: firstly the overlapping of NIR emission from CO_2 , and H_2O may cause difficulties in instrumentation; secondly, the part of NIR spectrum from 700-1100 nm is not particularly targeted from the main sensor manufactures, which may also be subjected to the limitations of mainstream sensor material properties.

In this thesis, state-of-the-art fire detection was comprehensively reviewed. The overall structure of a video fire detection algorithm consists three main parts: region

of interest (ROI) extraction, feature selection and extraction and decision-making. There can be multiple sub-stages within each part. Many fire detection in literature only treat fire detection as an algorithm problem. It is believed by the author that a comprehensive knowledge of sensor hardware as well as a bespoke algorithm is the key to robust and informative fire detection. In addition, video fire detection has very limited extension to the infrared spectrum throughout literature. Not only is the captured flame signal different from our intuition, but the interaction between the environment and the infrared spectral radiation is less studied. All of which helps to shed lights on the detection and improve the robustness of fire detection system.

The usability of near infrared spectrum can also be valuable for flame diagnostics, especially for lean premixed flames where the radiation disturbance from soot is minimal. The excellent near infrared spectral sensitivity make it very promising in combustion products (CO_2 , and H_2O) monitoring. Because of the infrared spectrum of these two products contributes a great portion of the radiant energy of the flame, their visualization and instrumentation can be used for combustion diagnostics, especially for thermal induced instability. Moreover, the near sensitivity gives limited thermal capability that it can visualize the temperature distribution of objects having temperature as low as 375 Celsius. This was conveniently used to visualize wood pyrolysis regions in the study of flame propagation of solid fuel combustion. The main aspects of the research have been concluded as follows:

- The modification and characterization of CMOS image sensor. The Bayer colour filter array is commonly used to generate colour mosaic from which the complete visible spectral colour can be reproduced by interpolation. With the integrated Bayer CFA, the quantum efficiency of the sensor pixels is reduced. In addition, the IR blocker also integrated in the camera lens system to prevent the near

infrared rays from reaching the sensors. The removal of Bayer CFA and IR blocker unlocks the full potential of silicon based CMOS image sensor, enabling it to sense UV to near infrared spectral radiation (300-1100 nm). The test showed great linearity of camera response with respect to ISO and exposure. The spectral response of the sensor is non linear. The use of the sensor for instrumentation of absolute spectral irradiance requires normalization of its spectral response.

- The image property of NIR fire image has been extensively studied. Similar to visible camera, the NIR camera also relies on the reflected NIR radiation from the scene to form an image. Difference is also very noticeable, especially from vegetation, sky and clouds because of the different scattering effect of NIR wavelength comparing to that of visible lights. The combined effect of higher sensitivity in the NIR spectrum and fire radiation property give rise to the huge contrast between flame and the background. The camera system has high configurability that the fixed setting can circumvent the camera's auto white balance and adaptive exposure, therefore eliminate the effect of illumination in the scene. Tests showed that the image maintained its high contrast between foreground fire and background even with very strong sun reflection in the surroundings. In an indoor situation with lighting, most of the scene were nearly dark and the light sources were very dim comparing to fire. Simple empirically determined threshold is effective in the ROI extraction.
- For fire recognition, spatial texture features were used, namely the image entropy, homogeneity and intensity variance, which are seldom used. Probably because texture is greatly affected by illumination and the exposure of an image.

In many research, fire detection algorithms were proposed using common visible camera, none of which shed lights on the camera hardware setting and control. This is potentially problematic as the adaptive exposure depends on the scene illumination. In addition, in the case of using public surveillance system for fire detection, the control over the camera setting may also be limited. Moreover, in the fire detection of using long wavelength IR microbolometer detector, frequent mention of fire image saturation have been noticed. For these situations, texture based feature can be very unreliable as the spatial intensity variation cannot be consistently preserved. Therefore, in research utilizing long wavelength IR cameras, movement based feature were extracted, which measures the contour variation of the saturated foreground fire. Hence, it can be concluded that the choice of feature extraction largely depends on camera hardware and application scenarios. A comprehensive knowledge on both hardware and algorithms are beneficial in designing robust video fire detection systems.

- On machine learning based classification algorithms, kernel support vector machine (SVM) and convolutional neural network (CNN) were tested. It is worth reiterating that most algorithms in theory performs similarly. The right question to ask is whether the algorithm is practically realizable given the current resources. The point has been demonstrated in this research. Neural networks usually have much more trainable parameters in addition to the hyper-parameters in fine tuning. Therefore, the optimization process requires large amount of data to achieve better function fitting between the input and output. Notwithstanding the training process has been drastically accelerated owing to the AI enhanced processing units, more time can be dedicated in model architecture experimentation and training techniques, which are an iterative

process with extensive efforts. In the case of SVM however, it is a sparse model where only a subset of the training examples (support vectors) are used to find the optimal decision plane, offering much easier training and deployment. This is ideal for the case with limited data resources. The performance of a SVM for a classification task depends on the features. More relevant and distinctive features of the objects of interest were extracted, more likely the classification performance will be good. For example, the features of each class are highly intertwined or clustered in the feature space, the nonlinear kernel SVM will find a highly nonlinear decision surface that tries to distinguish each class. Because of the distribution of the features from each class are so similar (intertwined), it is very likely that the optimal decision plane will have much reduced performance.

- The combination of NIR and visible spectrum with Bayer colour filter array (Bayer CFA) produces purple to pinkish colour of flames, showed a consistent colour distribution for both diffusion and premixed flames. This feature has been utilized in fire detection as an extra colour filtering to increase system robustness by false positive reduction. Gaussian mixture model with Dirichlet process prior was employed to model the colour of the flame. Comparing to conventional GMM, the non-parametric Bayesian approach generates a much better fit of the colour distribution. Such false colour is caused by the addition of near infrared radiation into blue and red pixels. This is immune to false fire as in advertisement or on display, which only emits visible lights. In addition, the system is also capable of detecting non-luminous flames.
- In the field of gas combustion diagnostics, the NIR sensitivity allows the camera

to capture the emission of combustion products (CO_2 and H_2O). Test cases of lean premixed methane/hydrogen and air flame showed the convenience of imaging of flame pulsation, which is difficult by imaging in visible spectrum. The visualization of hot combustion products can be applied to more complex combustion conditions to study thermal interaction of flame with flow field. In addition, the removal of Bayer CFA provides the capability of precise narrow band imaging with improved pixel quantum efficiency from UV to NIR spectrum (300-1100 nm) at a much less cost comparing to other NIR camera sensors such as InGaAs cameras.

- The sensitivity of NIR of such sensors is capable of thermal imaging to a lesser degree comparing to mid and long wavelength cameras, but with much less cost and much higher spatial resolution. In the case study, the multi-spectral sensitivity of the camera sensor was exploited in conjunction with schlieren imaging forming a comprehensive visualization technique for solid fuel fire propagation. The low temperature regions corresponding to wood pyrolysis and the visualization of convective hot flow field explained the mechanism of flame propagation with respect to sample inclination. Despite the lesser capability of thermal imaging comparing to longer wavelength thermal cameras, the NIR camera sensor will find its ground in solid fuel combustion diagnostics with incomparable cost-effectiveness. Moreover, the correlation between intensity variation and media density change has also been studied through schlieren imaging test case, where the cooling effect can be quantified.

6.2 Future work

Future work can be established by following along the two main bifurcations: Multi-spectrum fire detection and NIR combustion diagnostics.

- The robustness of the fire detection system needs further tests. Real world situation can be replicated in laboratory. Particularly, short wavelength infrared is prone to affected by thermal reflections. If the fire is close to wall or some other kinds of surfaces, the reflection will have a great effect on the fire images. The reflectance varies on different surfaces, which all depends on the material and finishing. The interaction of NIR radiation with various common materials need further investigation. Therefore, a range of real world surfaces need to be tested in regards to the change of fire image and to test detection algorithm. Moreover, in a museum situation, glass is extensively used. If there is a fire, there can be multiple reflections of the fire which may introduce more interference. Furthermore, the NIR radiation may bouncing between glass or glossy surfaces that could potentially render further disturbance. Such real world condition can be emulated to help further study.
- Further filtering to reject false positive classification is worth investigation using Kolmogorov-Smirnov statistical test (KS test) in addition to the DPGMM colour model. This can be though as a statistical texture information because the flame in NIR only mode at controlled setting has a distinctive spatial distribution as well as pixel intensity distribution. With enough samples (fire images), the empirical cumulative distribution function can be obtained. For a potential fire image patch, its distribution can be compared using KS test whereby the maximum distance in the two distribution function spaces can be

computed. Then corresponding threshold needs to be found to reject the null hypothesis that the new distribution (potential fire patch) is drawn from the empirical cumulative distribution.

- The stereo cameras are useful for inferring the flame size and distance to the camera. Stereo camera model operates on two cameras having the same visible camera, where image registration and matching from both image is readily available. However, in the stereo imaging system with cameras operating at different spectrum, image registration can be a challenge. The image registration between NIR and NIR+Visible fire image can be studied. This particularly promising for future automated robot firefighter to locate fire and navigate through space. With on-board computation capability of raspberry pi system, the ease of system integration can make the cost of prototyping of future automated fire fighting very affordable.
- It is well known that long wavelength IR radiation can penetrate smoke as the scattering effect is dramatically reduced for long wavelength . In comparison, the NIR wavelength is much shorter, which has limited smoke penetration. Therefore, to what degree does the smoke affect fire image is a topic worth study. For experimentation, smoke can be generated in an enclose space where the density of smoke can be controlled. Fire imaging test can be established accordingly.
- The use of NIR for combustion products imaging of lean premixed flame can be further extended to more complex situations. For example, the imaging can be applied to tubular premixed flame propagation in quartz tube. The visualization of CO_2 and H_2O could bring new insights to the internal pressure

and flame oscillations. Moreover, the excellent NIR sensitivity can be utilized for real-time temperature measurements by two-colour methods, with one selected band for each camera, offering great spatial resolution.

Appendix A

Fire detection algorithm

The code produced in this research can be found in the link below:

<https://drive.google.com/drive/folders/1KcAVp3IiqsuYzxS0RsWVML44TJUNHJPa?usp=sharing>

See the comments in the file for explanation.

References

- [1] “Notre-dame came far closer to collapsing than people knew. this is how it was saved. - the new york times.” [Online; accessed 2021-04-08].
- [2] J. G. Quintiere, *Principles of fire behavior*. Albany, N.Y: Delmar Publishers, 1998.
- [3] S. A. Dyer, *Survey of Instrumentation and Measurement*. IEEE press, 2001.
- [4] W. A.Schmidt-Ott, H.Burtscher, “Electrostatic fire detector,” *Fire Safety journal*, vol. 17, pp. 423–430, 1991.
- [5] T. Shields, *Building and fire*. Silcock, London: Longman Academic, Scientific & Technical, 1987.
- [6] B. J. Meacham and V. Motevalli, “Characterization of smoke from smoldering combustion for the evaluation of light scattering type smoke detector response,” *Journal of Fire Protection Engineering*, vol. 4, no. 1, pp. 17–32, 1992.
- [7] M. A. Jackson and I. Robins, “Gas sensing for fire detection: Measurements of co , co_2 , h_2 , O_2 , and smoke density in european standard fire tests,” p. 25.
- [8] E. Budnick and E. W. Forsell, “In situ tests of smoke detection systems for telecommunications central office facilities,” *Hughes Associates*, 10 1990.

- [9] L. C. Speitel, "Fourier transform infrared analysis of combustion gases," *Journal of Fire Sciences*, vol. 20, no. 5, pp. 349–371, 2002.
- [10] J. Hodgkinson and R. P. Tatam, "Optical gas sensing: a review," *Measurement Science and Technology*, vol. 24, p. 012004, 1 2013.
- [11] D. Popa and F. Udrea, "Towards integrated mid-infrared gas sensors," *Sensors*, vol. 19, p. 2076, 5 2019.
- [12] X. Liu, S. Cheng, H. Liu, S. Hu, D. Zhang, and H. Ning, "A survey on gas sensing technology," *Sensors*, vol. 12, pp. 9635–9665, 7 2012.
- [13] M. Sieger and B. Mizaikoff, "Toward on-chip mid-infrared sensors," *Analytical Chemistry*, vol. 88, no. 11, pp. 5562–5573, 2016.
- [14] L. Fleming, D. Gibson, S. Song, C. Li, and S. Reid, "Reducing n_2o induced cross-talk in a ndir co_2 gas sensor for breath analysis using multilayer thin film optical interference coatings," *Surface and Coatings Technology*, vol. 336, pp. 9–16, 2018.
- [15] P. Barritault, M. Brun, O. Lartigue, J. Willemin, J.-L. Ouvrier-Bufferet, S. Pocas, and S. Nicoletti, "Low power co_2 ndir sensing using a micro-bolometer detector and a micro-hotplate ir-source," *Sensors and Actuators B: Chemical*, vol. 182, pp. 565–570, 6 2013.
- [16] M. Dong, C. Zheng, S. Miao, Y. Zhang, Q. Du, Y. Wang, and F. Tittel, "Development and measurements of a mid-infrared multi-gas sensor system for co , co_2 and ch_4 detection," *Sensors*, vol. 17, no. 10, p. 2221, 2017.

- [17] T. Vincent and J. Gardner, “A low cost mems based ndir system for the monitoring of carbon dioxide in breath analysis at ppm levels,” *Sensors and Actuators B: Chemical*, vol. 236, pp. 954–964, 2016.
- [18] A. Pusch, A. De Luca, S. S. Oh, S. Wuestner, T. Roschuk, Y. Chen, S. Boual, Z. Ali, C. C. Phillips, M. Hong, S. A. Maier, F. Udrea, R. H. Hopper, and O. Hess, “A highly efficient cmos nanoplasmonic crystal enhanced slow-wave thermal emitter improves infrared gas-sensing devices,” *Scientific Reports*, vol. 5, no. 1, p. 17451, 2015.
- [19] A. D. Luca, S. Z. Ali, R. Hopper, S. Boual, J. W. Gardner, and F. Udrea, “Filterless non-dispersive infra-red gas detection: A proof of concept,” pp. 1220–1223, 2017 IEEE 30th International Conference on Micro Electro Mechanical Systems (MEMS).
- [20] Q. Tan, L. Tang, M. Yang, C. Xue, W. Zhang, J. Liu, and J. Xiong, “Three-gas detection system with ir optical sensor based on ndir technology,” *Optics and Lasers in Engineering*, vol. 74, pp. 103–108.
- [21] C. Ranacher, C. Consani, A. Tortschanoff, R. Jannesari, M. Bergmeister, T. Grille, and B. Jakoby, “Mid-infrared absorption gas sensing using a silicon strip waveguide,” *Sensors and Actuators A: Physical*, vol. 277, pp. 117–123, 7 2018.
- [22] N. P. Ayerden, G. de Graaf, P. Enoksson, and R. F. Wolffenbuttel, “A highly miniaturized ndir methane sensor,” (Brussels, Belgium), p. 98880D, SPIE Photonics Europe, 2016. [Online; accessed 2021-03-15].

- [23] C. Massie, G. Stewart, G. McGregor, and J. R. Gilchrist, “Design of a portable optical sensor for methane gas detection,” *Sensors and Actuators B: Chemical*, vol. 113, no. 2, pp. 830–836, 2006.
- [24] P. Su, Z. Han, D. Kita, P. Becla, H. Lin, S. Deckoff-Jones, K. Richardson, L. C. Kimerling, J. Hu, and A. Agarwal, “Monolithic on-chip mid-ir methane gas sensor with waveguide-integrated detector,” *Applied Physics Letters*, vol. 114, no. 5, p. 051103, 2019.
- [25] L. Tombez, E. J. Zhang, J. S. Orcutt, S. Kamlapurkar, and W. M. J. Green, “Methane absorption spectroscopy on a silicon photonic chip,” *Optica*, vol. 4, no. 11, p. 1322, 2017.
- [26] J. Cholin, “The current state of the art in optical fire detection,” *Plant/Operations Progress*, vol. 8, no. 1, pp. 12–18.
- [27] E. Jacobson, “Finding novel fire detection technologies for the offshore industry,” *Fire*, p. p. 26, 2000.
- [28] A. Chenebert, T. P. Breckon, and A. Gaszczak, “A non-temporal texture driven approach to real-time fire detection,” pp. 1741–1744, 2011 18th IEEE International Conference on Image Processing, 2011. ISSN: 2381-8549.
- [29] G. Healey, D. Slater, T. Lin, B. Drda, and A. D. Goedeke, “A system for real-time fire detection,” pp. 605–606, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 6 1993. ISSN: 1063-6919.
- [30] T.-H. Chen, P.-H. Wu, and Y.-C. Chiou, “An early fire-detection method based on image processing,” vol. 3, pp. 1707–1710 Vol. 3, 2004 International Conference on Image Processing, 2004. ICIP '04., 2004. ISSN: 1522-4880.

- [31] J. Chen, Y. He, and J. Wang, "Multi-feature fusion based fast video flame detection," *Building and Environment*, vol. 45, pp. 1113–1122, 5 2010.
- [32] X. Qi and J. Ebert, "A computer vision-based method for fire detection in color videos," *International journal of imaging and robotics*, pp. 22–34, 2009.
- [33] G. Marbach, M. Loepfe, and T. Brupbacher, "An image processing technique for fire detection in video images," *Fire Safety Journal*, vol. 41, no. 4, pp. 285–289, 2006.
- [34] L. Rossi, M. Akhloufi, and Y. Tison, "On the use of stereovision to develop a novel instrumentation system to extract geometric fire fronts characteristics," *Fire Safety Journal*, vol. 46, pp. 9–20, 1 2011.
- [35] T. Celik, "Fast and efficient method for fire detection using image processing," *ETRI Journal*, vol. 32, no. 6, pp. 881–890, 2010.
- [36] B. U. Toreyin, "Fire detection algorithms using multimodal signal and image analysis," *PhD thesis*, 2009.
- [37] Z. Lingkai and L. Li, "Video flame detection method based on improved fast robust feature," *Journal of Physics: Conference Series*, vol. 1518, p. 012065, 2020.
- [38] W. P. Iii and M. Shah, "Flame recognition in video q," *Pattern Recognition Letters*, p. 9, 2002.
- [39] P. V. K. Borges and E. Izquierdo, "A probabilistic approach for vision-based fire detection in videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 5, pp. 721–731, 2010.

- [40] Z. Li, L. S. Mihaylova, O. Isupova, and L. Rossi, "Autonomous flame detection in videos with a dirichlet process gaussian mixture color model," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 1146–1154, 2018.
- [41] K. Dimitropoulos, P. Barmpoutis, and N. Grammalidis, "Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 2, pp. 339–351, 2015.
- [42] K. Chen, Y. Cheng, H. Bai, C. Mou, and Y. Zhang, "Research on image fire detection based on support vector machine," pp. 1–7, 2019 9th International Conference on Fire Science and Fire Protection Engineering (ICFSFPE), 2019.
- [43] R. A. Khan, J. Uddin, and S. Corraya, "Real-time fire detection using enhanced color segmentation and novel foreground extraction," pp. 488–493, 2017 4th International Conference on Advances in Electrical Engineering (ICAEE), 2017. ISSN: 2378-2692.
- [44] X. Wang, Y. Li, and Z. Li, "Research on flame detection algorithm based on multi - feature fusion," (Chongqing, China), pp. 184–189, 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), IEEE, 2020. [Online; accessed 2020-12-06].
- [45] Z. Xu and J. Xu, "Automatic fire smoke detection based on image visual features," pp. 316–319, 2007 International Conference on Computational Intelligence and Security Workshops (CISW 2007), 2007.

- [46] R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L. Wixson, “A system for video surveillance and monitoring,” p. 69.
- [47] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, “Contour based smoke detection in video using wavelets,” pp. 1–5, 2006 14th European Signal Processing Conference, 2006. ISSN: 2219-5491.
- [48] T. Celik, H. Demirel, H. Ozkaramanli, and M. Uyguroglu, “Fire detection using statistical color model in video sequences,” *Journal of Visual Communication and Image Representation*, vol. 18, no. 2, pp. 176–185, 2007.
- [49] Z. Xiong, R. Caballero, H. Wang, A. M. Finn, M. A. Lelic, and P.-Y. Peng, “Video-based smoke detection: Possibilities, techniques, and challenges,”
- [50] I. Kolesov, P. Karasev, A. Tannenbaum, and E. Haber, “Fire and smoke detection in video with optimal mass transport based optical flow and neural networks,” (Hong Kong), pp. 761–764, 2010 17th IEEE International Conference on Image Processing (ICIP 2010), IEEE, 2010. [Online; accessed 2020-11-13].
- [51] F. Gomez-Rodriguez, B. C. Arrue, and A. Ollero, “Smoke monitoring and measurement using image processing: application to forest fires,” (Orlando, FL), p. 404, AeroSense 2003. [Online; accessed 2021-03-25].
- [52] Z. Wei, X. Wang, W. An, and J. Che, “Target-tracking based early fire smoke detection in video,” pp. 172–176, 2009 Fifth International Conference on Image and Graphics, 2009.

- [53] P. V. K. Borges, J. Mayer, and E. Izquierdo, “Efficient visual fire detection applied for video retrieval,” pp. 1–5, 2008 16th European Signal Processing Conference, 2008. ISSN: 2219-5491.
- [54] J. J. Huseynov, S. Baliga, A. Widmer, and Z. Boger, “Infrared flame detection system using multiple neural networks,” in *2007 International Joint Conference on Neural Networks*, pp. 608–612, 2007.
- [55] K. Wang, Y. Zhang, J. Wang, Q. Zhang, B. Chen, and D. Liu, “Fire detection in infrared video surveillance based on convolutional neural network and svm,” in *2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP)*, pp. 162–167, 2018.
- [56] I. Kolesov, P. Karasev, A. Tannenbaum, and E. Haber, “Fire and smoke detection in video with optimal mass transport based optical flow and neural networks,” in *2010 IEEE International Conference on Image Processing*, pp. 761–764, 2010.
- [57] K. Muhammad, J. Ahmad, and S. W. Baik, “Early fire detection using convolutional neural networks during surveillance for effective disaster management,” *Neurocomputing*, vol. 288, pp. 30–42, 2018.
- [58] W. Lee, S. Kim, Y.-T. Lee, H.-W. Lee, and M. Choi, “Deep neural networks for wild fire detection with unmanned aerial vehicle,” in *2017 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 252–253, 2017.
- [59] P. V. K. Borges and E. Izquierdo, “A probabilistic approach for vision-based fire detection in videos,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 5, pp. 721–731, 2010.

- [60] B. U. Töreyn, “Fire detection in infrared video using wavelet analysis,” *Optical Engineering*, vol. 46, no. 6, p. 067204, 2007.
- [61] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [62] R. Berwick, “An idiots guide to support vector machines (svms),” p. 27.
- [63] W. Pitts, “Some observations on the simple neuron circuit,” *The bulletin of mathematical biophysics*, vol. 4, no. 3, pp. 121–129, 1942.
- [64] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, no. 6088, pp. 533–536, 1986. number: 6088 publisher: Nature Publishing Group.
- [65] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [66] X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier neural networks,” p. 9.
- [67] “What is the relation between logistic regression and neural networks and when to use which?,” 2021. [Online; accessed 2021-03-19].
- [68] “Perceptron,” 2019. [Online; accessed 2021-03-18].
- [69] “Neural networks introduction to machine learning.” [Online; accessed 2021-03-19].
- [70] “What is softmax regression and how is it related to logistic regression?,” 2021. [Online; accessed 2021-03-19].

- [71] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998. event: Proceedings of the IEEE.
- [72] “Cs 230 - convolutional neural networks cheatsheet.” [Online; accessed 2021-03-20].
- [73] K. P. Bennett and E. J. Bredensteiner, “Duality and geometry in svm classifiers,” p. 8.
- [74] C. J. C. Burges, “A tutorial on support vector machines for pattern recognition,” *SUPPORT VECTOR MACHINES*, p. 47.
- [75] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” *arXiv:1502.03167 [cs]*, 2015. arXiv: 1502.03167.
- [76] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” p. 30.
- [77] “Advances in detectors: Hot ir sensors improve ir camera size, weight, and power,” 1 2014. [Online; accessed 2021-03-23].
- [78] S. H. Black, T. Sessler, E. Gordon, R. Kraft, T. Kocian, M. Lamb, R. Williams, and T. Yang, “Uncooled detector development at raytheon,” (Orlando, Florida, United States), p. 80121A, SPIE Defense, Security, and Sensing, 5 2011. [Online; accessed 2021-03-23].
- [79] J. Caulfield, J. Curzan, J. Lewis, and N. Dhar, “Small pixel oversampled ir focal plane arrays,” (Baltimore, Maryland, United States), p. 94512F, SPIE Defense + Security, 6 2015. [Online; accessed 2021-03-23].

- [80] R. K. McEwen, D. Jeckells, S. Bains, and H. Weller, “Developments in reduced pixel geometries with movpe grown mct arrays,” (Baltimore, Maryland, United States), p. 94512D, SPIE Defense + Security, 6 2015. [Online; accessed 2021-03-23].
- [81] K. Wang, Y. Zhang, J. Wang, Q. Zhang, B. Chen, and D. Liu, “Fire detection in infrared video surveillance based on convolutional neural network and svm,” (Shenzhen), pp. 162–167, 2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP), IEEE, 7 2018. [Online; accessed 2020-11-13].
- [82] S. Verstockt, A. Vanoosthuyse, S. Van Hoecke, P. Lambert, and R. Van de Walle, “Multi-sensor fire detection by fusing visual and non-visual flame features,” in *Image and Signal Processing* (A. Elmoataz, O. Lezoray, F. Nouboud, D. Mammass, and J. Meunier, eds.), vol. 6134, pp. 333–341, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. collection-title: Lecture Notes in Computer Science DOI: 10.1007/978-3-642-13681-8₃9.
- [83] “Thermal imaging fire detection algorithm with minimal false detection,” *KSII Transactions on Internet and Information Systems*, vol. 14, 5 2020. [Online; accessed 2020-11-13].
- [84] I. Bosch, S. Gomez, R. Molina, and R. Miralles, “Object discrimination by infrared image processing,” in *Bioinspired Applications in Artificial and Natural Computation* (J. Mira, J. M. Ferrández, J. R. Álvarez, F. de la Paz, and F. J. Toledo, eds.), vol. 5602, pp. 30–40, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. collection-title: Lecture Notes in Computer Science DOI: 10.1007/978-3-642-02267-8₄.
- [85] A. S.-A. José Meseguer, Isabel Pérez-Grande, *Spacecraft thermal control*. Woodhead Publishing, 2012.

- [86] R. K. Willardson and E. Weber, *Uncooled Infrared Imaging Arrays and Systems*. Academic Press, 1ed ed., 10 1997.
- [87] A. Rogalski, “Progress in focal plane array technologies,” *Progress in Quantum Electronics*, vol. 36, pp. 342–473, 3 2012.
- [88] A. Rogalski and M. Razeghi, “Narrow gap semiconductor photodiodes,” p. 12.
- [89] D. F. King, J. S. Graham, A. M. Kennedy, R. N. Mullins, J. C. McQuitty, W. A. Radford, T. J. Kostrzewa, E. A. Patten, T. F. McEwan, J. G. Vodicka, and J. J. Wootan, “3rd-generation mw/lwir sensor engine for advanced tactical systems,” (Orlando, FL), p. 69402R, SPIE Defense and Security Symposium, 4 2008. [Online; accessed 2019-03-28].
- [90] S. Iwabuchi, Y. Maruyama, Y. Ohgishi, M. Muramatsu, N. Karasawa, and T. Hirayama, “A back-illuminated high-sensitivity small-pixel color cmos image sensor with flexible layout of metal wiring,” pp. 1171–1178, 2006 IEEE International Solid State Circuits Conference - Digest of Technical Papers, 2 2006. ISSN: 2376-8606.
- [91] D. Durini, *High performance silicon imaging: fundamentals and applications of CMOS and CCD sensors*. 2020. OCLC: 1143015377.
- [92] A. Rogalski, *Infrared detectors*. Boca Raton: Taylor & Francis, 2nd ed ed., 2011.
- [93] “Uncooled infrared imaging market commercial & military applications,” tech. rep., 2011.
- [94] A. Rogalski, P. Martyniuk, and M. Kopytko, “Challenges of small-pixel infrared detectors: a review,” *Reports on Progress in Physics*, vol. 79, p. 046501, 4 2016.

- [95] J. A. Ratches, “Current and future trends in military night vision applications,” *Ferroelectrics*, vol. 342, pp. 183–192, 10 2006.
- [96] M. O. Oladimeji, M. Turkey, M. Ghavami, and S. Dudley, “A new approach for event detection using k-means clustering and neural networks,” (Killarney, Ireland), pp. 1–5, 2015 International Joint Conference on Neural Networks (IJCNN), IEEE, 7 2015. [Online; accessed 2020-11-13].
- [97] X. Yan, H. Cheng, Y. Zhao, W. Yu, H. Huang, and X. Zheng, “Real-time identification of smoldering and flaming combustion phases in forest using a wireless sensor network-based multi-sensor system and artificial neural network,” *Sensors*, vol. 16, p. 1228, 8 2016.
- [98] M. E. Webber, J. Wang, S. T. Sanders, D. S. Baer, and R. K. Hanson, “In situ combustion measurements of co, co₂, h₂o and temperature using diode laser absorption sensors,” *Proceedings of the Combustion Institute*, vol. 28, pp. 407–413, 1 2000.
- [99] A. G. Gaydon, *The Spectroscopy of Flames*. Dordrecht: Springer Netherlands, 1974. DOI: 10.1007/978-94-009-5720-6.
- [100] F. A. Williams, “Department of aerospace engineering, university of california at san diego, la joila, calif., u.s.a.,” vol. 8, p. 3.
- [101] V. Vilimpoc, L. P. Goss, and B. Sarka, “Spatial temperature-profile measurements by the thin-filament-pyrometry technique,” *Optics Letters*, vol. 13, p. 93, 2 1988.
- [102] P. B. Kuhn, B. Ma, B. C. Connelly, M. D. Smooke, and M. B. Long, “Soot and thin-filament pyrometry using a color digital camera,” *Proceedings of the Combustion Institute*, vol. 33, no. 1, pp. 743–750, 2011.

- [103] W. M. Pitts, “Thin-filament pyrometry in flickering laminar diffusion flames,” *Symposium (International) on Combustion*, vol. 26, pp. 1171–1179, 1 1996.
- [104] J. V. Pastor, J. M. García-Oliver, A. García, C. Micó, and S. Möller, “Application of optical diagnostics to the quantification of soot in n-alkane flames under diesel conditions,” *Combustion and Flame*, vol. 164, pp. 212–223, 2 2016.
- [105] Q. V. Nguyen, R. W. Dibble, C. D. Carter, G. J. Fiechtner, and R. S. Barlow, “Raman-lif measurements of temperature, major species, oh, and no in a methane-air bunsen flame,” p. 12.
- [106] S. Kostka, S. Roy, P. J. Lakusta, T. R. Meyer, M. W. Renfro, J. R. Gord, and R. Branam, “Comparison of line-peak and line-scanning excitation in two-color laser-induced-fluorescence thermometry of oh,” *Applied Optics*, vol. 48, p. 6332, 11 2009.
- [107] S. Roy, T. R. Meyer, R. P. Lucht, V. M. Belovich, E. Corporan, and J. R. Gord, “Temperature and co2 concentration measurements in the exhaust stream of a liquid-fueled combustor using dual-pump coherent anti-stokes raman scattering (cars) spectroscopy,” *Combustion and Flame*, vol. 138, pp. 273–284, 8 2004.
- [108] K. Sun, R. Sur, X. Chao, J. B. Jeffries, R. K. Hanson, R. J. Pummill, and K. J. Whitty, “Tdl absorption sensors for gas temperature and concentrations in a high-pressure entrained-flow coal gasifier,” *Proceedings of the Combustion Institute*, vol. 34, pp. 3593–3601, 1 2013.
- [109] S. Li, A. Farooq, and R. K. Hanson, “H2o temperature sensor for low-pressure flames using tunable diode laser absorption near $2.9 \mu\text{m}$,” *Measurement Science and Technology*, vol. 22, p. 125301, 10 2011. publisher: IOP Publishing.

- [110] X. Zhou, X. Liu, J. B. Jeffries, and R. K. Hanson, "Selection of near H_2O absorption transitions for in-cylinder measurement of temperature in ICE engines," *Measurement Science and Technology*, vol. 16, pp. 2437–2445, 12 2005.
- [111] R. M. Mihalcea, D. S. Baer, and R. K. Hanson, "Advanced diode laser absorption sensor for in situ combustion measurements of CO_2 , H_2O , and gas temperature," p. 7.
- [112] D. Blunck, S. Basu, Y. Zheng, V. Katta, and J. Gore, "Simultaneous water vapor concentration and temperature measurements in unsteady hydrogen flames," *Proceedings of the Combustion Institute*, vol. 32, no. 2, pp. 2527–2534, 2009.
- [113] S. Nakaya, T. Funahashi, Y. Asakami, I. Fujio, S. Takahashi, and M. Tsue, "Thermometry of combustion gas measuring two-band near-infrared emissions less than 1.1 m from water molecules," *Experimental Thermal and Fluid Science*, vol. 94, pp. 1–8, 6 2018.
- [114] G. Wang, X. Liu, S. Wang, L. Li, and F. Qi, "Experimental investigation of entropy waves generated from acoustically excited premixed swirling flame," *Combustion and Flame*, vol. 204, pp. 85–102, 6 2019.
- [115] J. Ji, Y. Sivathanu, and J. Gore, "Thermal radiation properties of turbulent lean premixed methane air flames," *Proceedings of the Combustion Institute*, vol. 28, pp. 391–398, 1 2000.
- [116] Y. Ju, G. Masuya, and P. D. Ronney, "Effects of radiative emission and absorption on the propagation and extinction of premixed gas flames," *Symposium (International) on Combustion*, vol. 27, pp. 2619–2626, 1 1998.

- [117] Y. Ju, G. Masuya, F. Liu, H. Guo, K. Maruta, and T. Niioka, "Further examinations on extinction and bifurcations of radiative CH_4/air and $\text{C}_3\text{H}_8/\text{air}$ premixed flames," *Symposium (International) on Combustion*, vol. 27, pp. 2551–2557, 1 1998.
- [118] S. Anand and K. Manjari.R.K, "Fpga implementation of artificial neural network for forest fire detection in wireless sensor network," (Chennai, India), pp. 265–270, 2017 2nd International Conference on Computing and Communications Technologies (ICCCCT), IEEE, 2 2017. [Online; accessed 2020-11-13].
- [119] "What is swir? | edmund optics." [Online; accessed 2021-03-29].
- [120] H. M. Heise, "Donald a. burns, emil w. ciurczak (eds.): Handbook of near-infrared analysis, 3rd ed.," *Analytical and Bioanalytical Chemistry*, vol. 393, pp. 1387–1389, 3 2009.
- [121] M. Vredenburg, P. Caspers, R. Hoogerbrugge, and D. Barends, "Choice and validation of a near infrared spectroscopic application for the identity control of starting materials.," *European Journal of Pharmaceutics and Biopharmaceutics*, vol. 56, pp. 489–499, 11 2003.
- [122] F. W. K. Iv, E. Lee, L. H. Kidder, and E. N. Lewis, "Near infrared spectroscopy: the practical chemical imaging solution," *Spectroscopy Europe*, p. 5, 2002.
- [123] Z. Xiaobo, Z. Jiewen, M. J. W. Povey, M. Holmes, and M. Hanpin, "Variables selection methods in near-infrared spectroscopy," *Analytica Chimica Acta*, vol. 667, pp. 14–32, 5 2010.
- [124] "Ingaas camera c12741-03." [Online; accessed 2021-03-29].
- [125] F. Von I/. MECKE and H. W. C. G. BALDWIN, "Warum erscheinen die H_2O im ultraroten licht hell?," *DIE NATLIRWISSENSCHAFTEN*, 1937.

- [126] K. Mangold, J. A. Shaw, and M. Vollmer, “The physics of near-infrared photography,” *European Journal of Physics*, vol. 34, pp. S51–S71, 11 2013.
- [127] “Blue sky and rayleigh scattering.” [Online; accessed 2021-03-29].
- [128] M. Vollmer, K.-P. Möllmann, and J. A. Shaw, “The optics and physics of near infrared imaging,” (Bordeaux, France), p. 97930Z, *Education and Training in Optics and Photonics: ETOP 2015*, 10 2015. [Online; accessed 2021-03-28].
- [129] “The bayer sensor strategy.” [Online; accessed 2021-03-30].
- [130] alex, “New and old raspberry pi camera comparison shots (1.3, 2.1 & noir),” 4 2016. [Online; accessed 2021-03-30].
- [131] T. Wilkes, A. McGonigle, T. Pering, A. Taggart, B. White, R. Bryant, and J. Willmott, “Ultraviolet imaging with low cost smartphone sensors: Development and application of a raspberry pi-based uv camera,” *Sensors*, vol. 16, p. 1649, 10 2016.
- [132] S.-L. Pan, “75 inventors: Chih-kang chiu, yung-ho;,” p. 10.
- [133] G. P. Anderson, J. H. Chetwynd, and E. P. She, “Afgl atmospheric constituent profiles (0.120km),” p. 46.
- [134] Y. Le Maoult, T. Sentenac, J. Orteu, and J. Arcens, “Fire detection a new approach based on a low cost ccd camera in the near infrared,” *Process Safety and Environmental Protection*, vol. 85, pp. 193–206, 1 2007.
- [135] C. Yuan, Z. Liu, and Y. Zhang, “Fire detection using infrared images for uav-based forest fire surveillance,” (Miami, FL, USA), pp. 567–572, 2017 International

- Conference on Unmanned Aircraft Systems (ICUAS), IEEE, 6 2017. [Online; accessed 2020-11-13].
- [136] J. D. Burnett and M. G. Wing, “A low-cost near-infrared digital camera for fire detection and monitoring,” *International Journal of Remote Sensing*, vol. 39, pp. 741–753, 2 2018.
- [137] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, pp. 610–621, 11 1973. event: IEEE Transactions on Systems, Man, and Cybernetics.
- [138] N. J. W. Morris, S. Avidan, W. Matusik, and H. Pfister, “Statistics of infrared images,” (Minneapolis, MN, USA), pp. 1–7, 2007 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 6 2007. [Online; accessed 2021-04-01].
- [139] C. M. Bishop, *Pattern recognition and machine learning*. Information science and statistics, New York: Springer, 2006.
- [140] S. McAllister, J.-Y. Chen, and A. C. Fernandez-Pello, *Fundamentals of Combustion Processes*. Mechanical Engineering Series, New York, NY: Springer New York, 2011. DOI: 10.1007/978-1-4419-7943-8.
- [141] B. J. Jackel, C. Unick, M. T. Syrjäsoo, N. Partamies, J. A. Wild, E. E. Woodeld, I. McWhirter, E. Kendall, and E. Spanswick, “Auroral spectral estimation with wide-band color mosaic ccds,” p. 24, 2014.
- [142] “raspicam - can raspberry pi noir camera see near infrared light (~1550nm)?” [Online; accessed 2020-12-18].