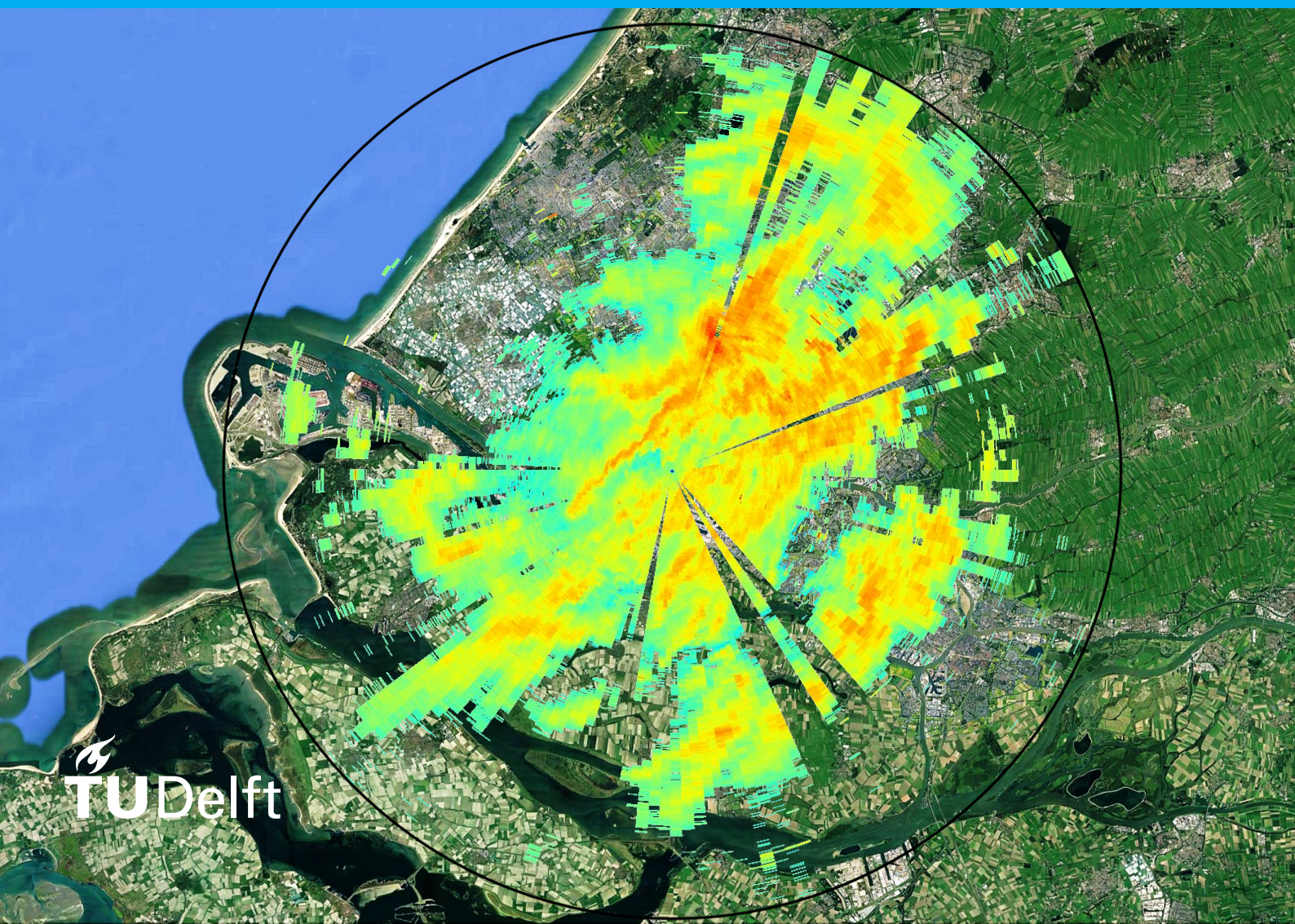


Weather radar data processing

A new technique to filter non-hydrometeors and a new methodology to evaluate filtering techniques without ground truth

C.Chen



Weather radar data processing

**A new technique to filter non-hydrometeors
and a new methodology to evaluate filtering
techniques without ground truth**

by

Cheng Chen

to obtain the degree of Master of Science
at the Delft University of Technology,

Student number: 4898249

Project duration: April, 2020 – January, 2021

Thesis committee: Ir. C. M. H. Unal,
Dr. A. C. P. Oude Nijhuis,
Prof. dr. ir. H. W. J. Russchenberg,
Dr. O. A. Krasnov,

TU Delft, Chair and supervisor
SkyEcho, co-supervisor
TU Delft
TU Delft

Preface

Before I came to study in TU Delft, I was still confused about what I am interested in. I concern the environmental issues and fancy the IT technology but I could not find a way to integrate them. During my bachelor study of environmental engineering, extensive chemical laboratory works made me decide to change my major. After I came to study TU Delft, I had the chance to try new things to find more insights. Most of my compulsory lectures are about hydrology and atmosphere. I even attempted to choose some electives in the Architecture Faculty. Finally, remote sensing caught my interests.

2020 is a hard year for everyone. And this project was done during this period. I went back to China to celebrate the Chinese New Year with my family members in the middle of January. However, the virus outbreak forced me to terminate my holiday in advance. Then after I came back to the Netherlands for a short period, the lockdown occurred here too. Therefore, this project was done totally in my room in Delft. Besides, as a freshman of remote sensing, this project is difficult for me at the beginning. The first time I got to know remote sensing was in the 4th quarter of my first year. Therefore, neither solid physical nor mathematical background knowledge did I have. I had to spend huge efforts to overcome those difficulties.

Though the period and the project are both hard, I am lucky to receive supports from many people. I would like to thank my supervisors Christine Unal and Albert Oude Nijhuis first. Christine helped me to initialize this project, review and give me detailed feedback to fix my drafts constantly. There would not be this thesis without her help. Albert supplied me all the data and helped me to understand the code and practical working. Also, thank Yann from SkyEcho for helping me to understand the project at first and give me ideas for future career planning in radar areas.

Thank my friends from the same master track. The moment of studying and preparing exams together day and night will never be forgotten.

Thank all of my friends from Delft Student Fellowship and my host family, Mr. and Mrs. Chen, for your assistance when I came to the Netherlands the first time. You treat me like your real family and teach me a lot of life skills.

Finally, thank my parents who give me all the financial support for my tuition and life cost in the Netherlands. It would be impossible to come to study here without your help.

*Cheng Chen
Delft, January 2021*

Contents

Abstract	1
Nomenclature	1
1 Introduction	5
1.1 Background	5
1.2 Problem statement	6
1.3 Research objectives	6
1.4 Research Limitations	7
1.5 Thesis roadmap	7
2 Literature Review	8
2.1 Spectral polarimetric processing	8
2.2 Image segmentation	9
2.3 Conclusions	11
3 Radar Specifications and System Overview	12
3.1 FMCW modulation pattern	12
3.2 System specifications	16
3.3 Radar location	17
4 Data processing and filtering	18
4.1 Processing pipeline	18
4.2 Range-Doppler Spectrogram Acquisition.	19
4.3 Description of two specific filters	19
4.3.1 Default Rijmond filtering	20
4.3.2 Object-orientated spectral polarimetric (OBSpol) filtering	20
4.4 Radar ray profile and PPI.	21
4.4.1 Radar Profile	21
4.4.2 PPI	22
5 Methodology	24
5.1 Image Feature Observation and Quantification.	24
5.1.1 Examples of Spectrogram Observations	26
5.1.2 Features Quantification	28
5.2 Histogram based Thresholding Enhancement	31
5.2.1 Spectral polarimetric parameter sensitivity analysis	31
5.2.2 Threshold magnitude determination	33
5.3 A Jensen–Shannon Distance-based thresholding filter.	34
5.3.1 Step 1: Adaptive spectral polarimetric thresholding	34
5.3.2 Step 2: Removing ground clutter by notch filter.	36
5.3.3 Step 3: Reconstruct missing precipitation	36
5.3.4 Step 4: Removing narrow Doppler power spectra	37

5.3.5	Examples of implementing JSD-based Filtering	37
5.4	Criteria for good segmentation	43
5.4.1	Criteria in spectral domain	43
5.4.2	Criteria in time domain.	43
5.4.3	Schematic of good and bad segmentation results	45
5.5	Evaluation	49
5.5.1	Subjective evaluation	49
5.5.2	Objective evaluation	49
5.6	Compare and assess different filtering methods	57
6	Results	59
6.1	Subjective evaluation results	59
6.1.1	Clear air case	59
6.1.2	The rainy cases	60
6.2	Objective Evaluation.	62
6.2.1	The clear air case evaluation	62
6.2.2	The two rainy cases evaluation	62
6.3	Filter performance: summary and discussions	65
7	Conclusions and recommendations	69
7.1	Conclusions	69
7.2	Contribution	70
7.3	Recommendations	70
	Bibliography	74

Abstract

Accurate rainfall intensity measurements are important to many applications like weather monitoring and forecasting, economics, urban design, agriculture and so on. A weather radar is one of the instruments to measure rainfall by reflected microwave signal from raindrops. However, the rainfall rate estimated by weather radar has various errors as it is an indirect measurement of rain. One type of errors relates to the measurement of other objects than raindrops and noise. These objects are termed clutter and represent buildings, trees, airplanes, birds, insects, ... One type of techniques mitigates clutter and noise using polarimetric measurements in the spectral domain. This methodology, spectral polarimetric processing, can be reinforced by digital image processing techniques. Image segmentation is the image processing technique chosen due to its similar objectives with spectral polarimetric processing. There are two major challenges for this project. The first one is to search and design effective methods with the help of image segmentation techniques to remove clutter and noise while retaining precipitation. Most of the image segmentation studies were done to tackle images from other areas which have different features from compared to Range-Doppler spectrograms. The methods from image segmentation should be thoroughly tested before being applied. Another problem to be resolved is to assess the techniques applied without any ground truth data. And those methods used in previous studies may not be suitable for testing large data sets. A framework is proposed to understand and quantify the features of polarimetric spectrograms which can be useful for tackling the two challenges. Then, a new filtering technique, criteria for good results and evaluation methods are proposed based on the framework. Three different filtering techniques are tested by the evaluation methods in different scenarios. Finally, some points for future research are also discussed.

Nomenclature

LIST OF ACRONYMS

ATSR	Alternate Transmission Simultaneous Reception
CV	Computer Vision
DFT	DiscreteFourier Transform
DL	Deep Learning
FFT	Fast Fourier Transform
FMCW	Frequency Modulated Continuous Wave
IF	Intermediate Frequency
JSD	Jessen-Shannon Distance
NaN	Not a number
OBSpol	Object orientated spectral polarimetric
PPI	Plan position indicator
RHI	Range height indicator
Rx	The received signals
STSR	Simultaneous Transmission Simultaneous Reception
Tx	The transmitted signals

LIST OF SYMBOLS

α	Azimuth angle
$\delta\phi$	Phase change
Δf_{IF}	Change of IF frequency
Δr	Range resolution
Δv	Doppler velocity resolution
λ	The wavelenght of the transmitted signal
$\overline{P_{cla}(X)}$	The mean distribution of spectral parameter X in the clear air case
ϕ	Phase of the radar signal
ϕ_0	Initial phase
ρ_{co}	Copolar correlation coefficient
τ	Time delay
B	Bandwidth
C	Connectivity
C_r	Calibration constant for radar
c	Speed of light
$\overline{C_{co}(r)}$	Co-polar correlation
$D_{KL}(P Q)$	Kullback–Leibler Divergence
f_{IF}	IF frequency
f_s	The sampling frequency
H	Shannon's entropy
H_d	The boundary contrast
H_r	The expected region entropy
H_v	The region entropy
$JSD(P Q)$	Jessen-Shannon Distance between P and Q
JSD_D	The difference between JSD_H and JSD_U
JSD_H	Jessen-Shannon Distance between the distribution of extracted hydrometeors and $\overline{P_{cla}(X)}$
JSD_U	Jessen-Shannon Distance between the distribution of extracted non-hydrometeors and $\overline{P_{cla}(X)}$
L_{DR}^{HH}	Linear Depolarization Ratio (HH)
L_{DR}^{VV}	Linear Depolarization Ratio (VV)
L_{in}	The region of inner-line object
L_{out}	The region of outer-line object
M^G	The label mask after the notch filter
M^{nd}	The label mask after removing narrow Doppler spectra
$M^{s\rho_{co}}$	The label mask after $s\rho_{co}$ thresholding
$M^{step\ x}$	The label mask of the x-th step
$N_{Doppler}$	The number of Doppler resolutions
N_{range}	The number of resolution volumes of range
N_{sweeps}	The number of sweeps
N_{su}	The number of single unconnected echo bins of one radar ray
N_{ws}	The number of weak and strong echo bins of one radar ray
q	The structural element for dilation operation
r	Range

r_{max}	Maximum scanning range
R_{nd}	The ratio for narrow Doppler spectra removing
R_{rm}	Removal percentage
R_{su}	Ratio of single unconnected echo bins
R_{ws}	Ratio of weak and strong echoes
R_C	The region of extracted non-hydrometeors
R_H	The region of extracted hydrometeors
sZ_{XY}	Spectrogram of Reflectivity
S	Slope of the Sawtooth modulation
$s\rho_{co}$	Spectrogram of copolar correlation coefficient
S_{XY}	The complex spectrogram from transmitted polarization X and reception polarization Y
S_C	The pixel numbers of R_C
S_H	The pixel numbers of R_H
S_I	The total pixel numbers of the input spectrogram
sL_{DR}^{HH}	Spectrogram of L_{DR}^{HH}
sL_{DR}^{VV}	Spectrogram of L_{DR}^{VV}
$sX(r, v)$	Any spectral polarimetric parameter
sZ_{DR}	Spectrogram of Differential Reflectivity
T_{opt}	The optimized $s\rho_{co}$ threshold
T_C	Duration of one chirp
v	Doppler velocity
v_{max}	Maximum Doppler velocity
Z_{DR}	Differential reflectivity
Z_{XY}	Reflectivity transmitting by Y and received by X

Introduction

1.1. Background

Accurate rainfall intensity measurements are important to many applications like weather monitoring and forecasting, economics, urban design, agriculture and so on. A weather radar is one of the instruments to measure rainfall by reflected microwave signal from raindrops. And weather radar has larger spatial coverage than those granular measurements provided by rain gauges. However, the rainfall rate estimated by weather radar has various errors as it is an indirect measurement of rain. One of the reasons behind is that the radar antenna receives unwanted, non-meteorological signals named clutter. In an urban environment, those clutter issues are critical because there are various kinds of clutter sources such as tall buildings, birds and interferences from telecommunication. The Rijnmond radar is such a radar facing clutter challenges as it locates in the centre of one of the biggest cities of the Netherlands, Rotterdam. Clutter filtering techniques are necessary to apply for more accurate rainfall estimates. There is one technique removing clutter in spectral-domain integrated with polarimetric radar measurements. This methodology, spectral polarimetric processing, can be reinforced by digital image processing techniques. Image segmentation, a branch of digital image processing, objects to partition an image into regions of interests. As image segmentation has been beneficial to many areas like satellite and medical image processing, this motivates to investigate the potential for contributing to mitigating clutter without removing precipitations.

This master project is part of the EU funded program ATTRACT (2019 - 2020), under the project Enhanced Urban Rain Surveillance Systems for Smart city Solutions (EU-RainS). In EU-RainS project, the X-band weather radar technology is concerned to highly developed for conversion to a societal and commercial HD weather asset. And this master project aims to accomplish the first objective, that is to remove heavy urban clutter effect on the radar data to obtain high-quality rain data.

1.2. Problem statement

There are two major challenges for this project. The first one is to search and design effective methods with the help of image segmentation techniques to remove clutter while retaining precipitation. In the following, we term clutter and noise as non-hydrometeors. As image segmentation shares some connections with spectral polarimetric processing, this area of knowledge could be the source for searching algorithms to do clutter and noise suppression. Nevertheless, most of the image segmentation studies were done to tackle images from other areas which have different features from Range-Doppler spectrograms. The methods from image segmentation should be thoroughly tested before being applied.

Another problem to be resolved is to assess the techniques applied. First, the assessment methods used from previous studies may not be suitable for testing huge amount of data. Subjective comparison of one spectrum or one PPI between before and after filtering is the primary evaluation. Some validated the filter performance by using reference data. However, both of these two methods need particular expertise and too much human intervention that is not applicable to assess the massive amount of datasets. Further, there is no reference data to use in this project. Besides, there are no common standards for assessment from previous spectral polarimetric processing studies, which makes it hard to define objective evaluation methods. That is the key issue: non-hydrometeor mitigation evaluation without ground truth. Some studies about assessment have been taken in image segmentation, but those methods still need to be validated for Range-Doppler spectrograms.

1.3. Research objectives

The main research question for this project is:

In combination with spectral polarimetry, can image processing contribute to non-hydrometeor filtering and developing filtering evaluation methods?

To answer the main research question, the following research objectives will be addressed:

1. *Which image processing techniques can be beneficial to filtering development in the context of weather radar?*

The field of digital image processing is like a tool box which contains numerous techniques. Considering the connections between image processing and spectral polarimetric filtering, image segmentation will be investigated.

2. *How to make image processing techniques applicable to Range-Doppler spectrograms?*

Due to the specificity of Range-Doppler spectrograms, most image processing methods cannot be directly applied. The solution is to visualise spectrograms and understand the features from different scopes. The way of visualization also matters.

3. *How to develop evaluation methods for filtered Range-Doppler spectrograms having no ground truth?*

In other words, what is a good result? For this purpose, clear criteria should be defined. As the filtering processing is in spectral domain and radar profile is in time domain, criteria related to these two domains should be considered. In this context, subjective and objective evaluation methods will be investigated.

4. *How well the filtering technique, which is enhanced by image processing, performs compared with other filtering methods?*

A new filtering technique will be proposed. The performance of the new developed and other previous filtering techniques will be assessed by evaluation methods for several study cases.

1.4. Research Limitations

Some related scopes of this project will not be included. First, it is the signal processing part before retrieving Range-Doppler spectrum. Image segmentation is a middle procedure; therefore, any step before acquiring the spectrograms will not be discussed. Second, it is the algorithm efficiency even though real-time processing is one of the requirements for designing filtering. The efficiency does not only depend on the complexity of algorithm architecture – other factors such as hardware utilised, programming language and data structures within the program matter too. Another critical processing, Doppler velocity de-aliasing was not considered either. Test scenarios are also limited. Only clear air echo, light rain and heavy rain cases were tested.

1.5. Thesis roadmap

Chapter 2 introduces the background information of the two main methodologies applied: spectral polarimetric processing and image segmentation. Chapter 3 gives an overview of the radar system as well as the site. Chapter 4 presents the whole data processing chain from raw data to radar ray profile. Chapter 5 explains the methodologies. Chapter 6 shows the results by applying the methods to real cases acquired by the radar. Finally, conclusions and recommendations will be drawn in Chapter 7.

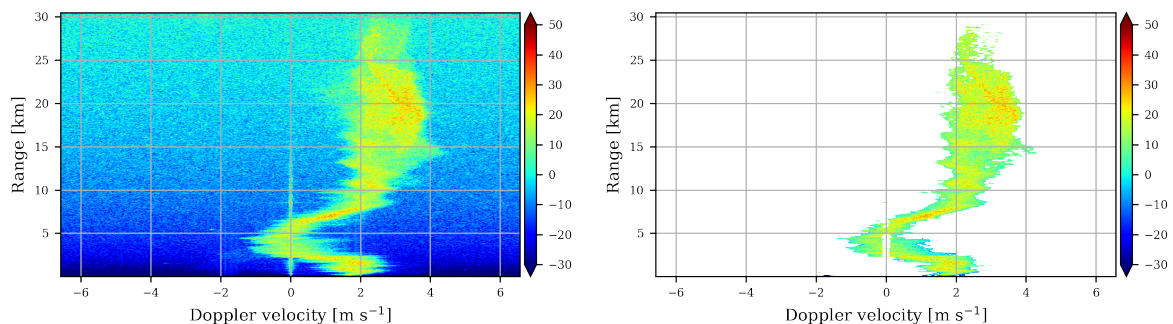
2

Literature Review

This chapter represents the background of spectral polarimetric processing and image segmentation. It objects to give an overview of theories of these two areas and find potentials to bridge them. First, an overview of spectral polarimetric processing is presented (section 2.1). Image segmentation was then introduced by its definitions and brief historical development (section 2.2). By looking at their definitions, applications and brief development, we could find the connections between them as well as further application. A conclusion is finally drawn.

2.1. Spectral polarimetric processing

Spectral polarimetric processing is the technique that is handling polarimetric radar data in the Range-Doppler domain. And a Range-Doppler spectrogram can be considered with two dimensions determined by maximum unambiguous range and Doppler velocity. With the benefits of polarization, several polarimetric Range-Doppler spectrograms can be obtained. From [1], to get polarimetric Range-Doppler spectrograms, a combined Doppler and polarization measurement is its prerequisite. Compared with statistical moment processing, Range-Doppler spectrograms retain more micro-physical information of atmospheric conditions.



(a) Raw spectrogram example

(b) Processed spectrogram example

Figure 2.1: Examples of raw and processed spectrograms

One remarkable application of spectral polarimetric processing is clutter suppression in weather radar. It is based on the different properties between clutter and hydrometeors shown in Range-Doppler spectra. Those properties include polarimetry, Doppler velocities and ranges. In [1], a thresholding-based clutter suppression filter was designed. This technique applied two spectral linear depolarization ratios to remove all Doppler-range bins larger than -5 dB. This threshold value relates to the used radar, its clutter environment and physical backscattering properties of hydrometeors. This technique can remove both ground and non-stationary clutter, as well as noise. The method showed better performance than a classical notch filter. However, some atmospheric targets with weak echo may be suppressed. Another research [2] did clutter suppression with the help of spectral copolar correlation coefficient that suggested removing range bins with values lower than 0.8.

Based on the research of [1] and [2], more clutter suppression techniques were created, such as [3], [4] and [5]. In [3], a technique was developed from double spectral linear depolarization ratio (sL_{DR}) and mathematical morphological processing. It was applied in a dual-orthogonal FMCW polarimetric radar. Its results showed this technique could significantly filter clutter, noise and artificial signals but keeping most atmospheric signals. And it is also applicable in both simultaneous transmission and reception and alternating transmission-reception modes.

[4] and [5] went further than [3] in attempting to combine spectral polarimetry with image processing techniques to enhance the spectral polarimetric processing. For both of them, the first step was also to do thresholding by sL_{DR} or spectral copolar-correlation coefficient ($s\rho_{co}$). Next, mathematical morphology was used to obtain the missing data in the precipitation spectrogram. The differences between these two methods are: [4] applied a moving window to select and remove clutter before mathematical morphological processing, while [5] removes the clutter left after mathematical morphology using object-orientated processing technique. In [5], the authors also made a comparison of the two methods. The results show that [5] has better performance than [4] for keeping weak precipitation.

2.2. Image segmentation

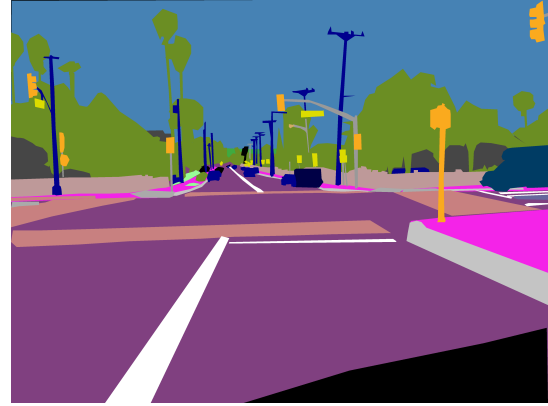
Image segmentation is a branch of computer graphics, which has been developed for decades. Inspired by some applications in this area, such as medical image, satellite image or even autonomous driving, image segmentation might also be considered to help discriminate hydrometeors and clutter.

There are several definitions for image segmentation. From the research [6], there are three definitions: generally, image segmentation is an image processing technique to separate an image into several regions. Besides, image segmentation can also be considered a method to clarify boundaries between different objects. A more technical definition of image segmentation is a process of labelling every pixel in the image to connect pixels with same labels based on some visual or semantic meaning (Fig. 2.2). [7] gives a mathematical representation of image segmentation. That is image segmentation is the process to partition the whole image into subsets which meet some requirements.

The definitions of image segmentation have not varied hugely, but the scope of the methods has changed incredibly in recent years. The development of Artificial



(a) Input Image



(b) Segmented Image

Figure 2.2: Examples of Image Segmentation [6] [8]

Intelligence (AI) has brought more innovation to image segmentation and profound benefits for further applications. Therefore, some studies ([9] and [6]) classified those algorithms into two types, before and after AI. Table.2.1 lists some examples for both of them. In the traditional scope, [10] divides image segmentation methods into three types which are matched to the three definitions from [10], namely, region-based, boundary-based, and pixel-based. In [7], the focus is given to two properties, similarity and discontinuity, leading to two methods, region-based and edge-based, respectively.

Decisions for choosing methods can sometimes be a trade-off, for example, between performance and efficiency. Intensive AI-based studies especially Deep Learning (DL) has made the entire Computer Vision (CV) into a new era including image segmentation and resulted in impressive performance compared to previous methods. Therefore, AI has made those traditional or old-school methods look like obsoletes. Besides, traditional methods are criticized because of more human intervention and expertise in relative domains([9]). The tendency of implementing DL driven methods are overwhelming, however, [11] argues that these two generation methodologies have pros and cons and suggests to optimize your choices based on the problem. Deep Learning methods have the prerequisites of high-powered GPUs([11]) and TPUs([11]), so they are more computationally expensive. Furthermore, image qualities such as high-resolution do matter for better performance by Deep Learning. And DL methods are criticized for being a black box due to its complicated layers of networks ([11]). In contrast, the traditional algorithms are valuable to overcome those constraints of Deep Learning.

Table 2.1: Traditional and Deep Learning Segmentation Methods Examples

Traditional	Deep Learning
Histogram-based Thresholding [12]	AlexNet [13]
Feature space clustering [14]	Convolutional autoencoders [15]
Region-based Segmentation [16]	Recurrent neural networks [17]
Edge detection methods [18]	Residual networks [19]

2.3. Conclusions

This chapter introduces the two fundamental techniques for this thesis project, spectral polarimetry and image segmentation. After reviewing spectral polarimetric clutter suppression and image segmentation techniques, there are some connections between them. The objectives are about partitioning input images into several objects of interests, using physical knowledge of precipitation and image processing. Image segmentation is mostly used in real-life images context, while spectral polarimetric clutter suppression focuses on Range-Doppler spectrogram to recognize clutter and atmospheric objects. From the studies above, thresholding is the fundamental image segmentation technique used in the spectral polarimetric process to do clutter suppression. More specifically, it is a histogram-based global thresholding. This method has the benefits of easy implementation and high efficiency, which suits for weather radar real-time processing. As image segmentation has been developed for decades, it is worthwhile to investigate more methods within this area to contribute to clutter suppression.

Radar Specifications and System Overview

3.1. FMCW modulation pattern

Not like most pulsed weather radar, the Rijnmond radar is an FMCW (Frequency-Modulated Continuous Wave) radar, whose frequency is modulated with time. A Saw-tooth frequency modulation is an example of linear frequency modulation versus time. The target is detected by sending transmitted signals(Tx) and accepting received signals(Rx). The characteristics of the Tx signal are determined by its start frequency (f_c), Bandwidth(B) and duration(T_c)([20, 21]).

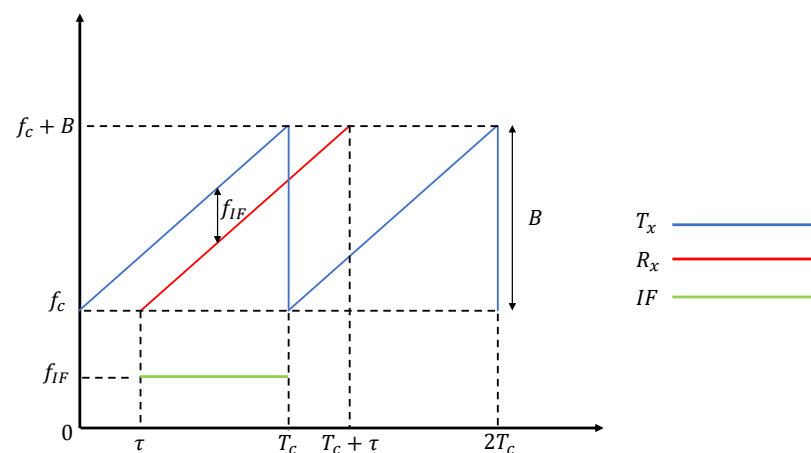


Figure 3.1: Frequency-time plot of Tx, Rx and IF signal (representation of two consecutive sweeps)

Linear modulation (Sawtooth):

$$S = \frac{B}{T_c} \quad (3.1)$$

$$f_{IF} = \frac{B}{T_c} \quad (3.2)$$

$$\tau = S \frac{2r}{c} \quad (3.3)$$

where S characterizes the slope of the linear modulation.

An object at range r is illuminated by the transmitted signals and some parts of the Tx signals are reflected and received by radar antenna after time τ . The speed of radar electromagnetic wave travelling in the air is approximately the same as light speed c . Transmitting and receiving the radar electromagnetic wave between radar antenna and the object at r is a two-way process, so the total travelling distance is $2r$. Therefore, we yield:

$$\tau = \frac{2r}{c} \quad (3.4)$$

From Nyquist sampling theorem, the maximum IF frequency that can be unambiguously measured is half of the sampling frequency f_s .

$$f_{IF_{max}} \leq \frac{f_s}{2} \quad (3.5)$$

$$S \frac{2r}{c} \leq \frac{f_s}{2} \quad (3.6)$$

$$r \leq \frac{c}{2S} \frac{f_s}{2} \quad (3.7)$$

$$r_{max} = \frac{c}{2} \frac{T_c}{B} \frac{f_s}{2} \quad (3.8)$$

The reflected Rx-signal is then mixed with the Tx-signal to produce an Intermediate Frequency (IF) signal. The frequency of the IF signal is determined by the time delay (τ) of the Rx-signal and equal to the frequency difference between Tx and Rx. Discrete Fourier Transform (DFT) is applied to resolve this frequency difference. According to DFT, minimal resolved frequency is the inverse of the sampling time, T_c . Combining with Eq.3.4, we can yield the range resolution Δr ([20, 21]).

Range resolution:

$$\Delta f_{IF} = \frac{1}{T_c} \quad (3.9)$$

$$S \frac{2\Delta r}{c} = \frac{1}{T_c} \quad (3.10)$$

$$\Delta r = \frac{c}{2ST_c} \quad (3.11)$$

$$\Delta r = \frac{c}{2B} \quad (3.12)$$

As it is a Doppler radar, it can also detect velocities relative to the radar rays. The Doppler velocity is measured by the phase change between two consecutive signals ([22]). The wavelength of the transmitted signal is λ and the scatter location is at range r . As the radar signal travels twice between the scatter and the radar antenna, the number of wavelengths is $2r/\lambda$. Besides, one wavelength corresponds to one cycle of the phase change, (2π) . The phase change (ϕ) is $4\pi r/\lambda$. The Doppler velocity (v) is defined as the range rate of change (dr) versus time (dt). Thus, we yield :

$$\phi = -\frac{4\pi}{\lambda}r + \phi_0 \quad (3.13)$$

$$\delta\phi = -\frac{4\pi}{\lambda}\delta r \quad (3.14)$$

$$\frac{\delta\phi}{\delta t} = -\frac{4\pi\delta r}{\lambda\delta t} \quad (3.15)$$

A radar can measure phase changes belonging to the interval $[-\pi, +\pi]$ and the sampling time between consecutive signals is T_c , we then yield the maximum Doppler velocity:

$$\delta\phi \leq \pi \quad (3.16)$$

$$\delta t = T_c \quad (3.17)$$

$$-\frac{4\pi}{\lambda}\delta r \leq \pi \quad (3.18)$$

$$-\frac{4\pi}{\lambda}v \leq \frac{\pi}{T_c} \quad (3.19)$$

$$|v| \leq \frac{\lambda}{4T_c} \quad (3.20)$$

$$v_{max} = \frac{\lambda}{4T_c} \quad (3.21)$$

The derivation Eq.3.13 to Eq.3.21 relates to a moving single object at the range r . In the case of hydrometeors, we have several particles with different Doppler velocities present at the range r . This is thus not sufficient to consider the phase change between two consecutive measurements in order to estimate one Doppler velocity. In that case, we consider a time series of measurements at fixed range corresponding to N_{sweeps} . The time variability of the phase during the time series is attributed to the different Doppler velocities of the hydrometeors. Using a Discrete Fourier Transform (DFT), we obtain complex Doppler spectra. Taking the modulus squared, we get Doppler power spectra, which provide the distribution of the Doppler velocities at the range r .

And the Doppler velocity resolution is determined by the number of sweeps:

$$\Delta v = \frac{2v_{max}}{N_{sweeps}} \quad (3.22)$$

Table 3.1: Manufacture Information of Rijnmond Radar

Radar Type	Polarimetric Doppler Weather Radar (FMCW)
Frequency	9.3 – 9.5 GHz (X-band)
Bandwidth	max 200 MHz
Maximum transmit power	20 W (CW)
Minimum range	200 m
Maximum range	60 km
Range resolution	Adjustable (0.75 to 150 m)
PRF	Adjustable (0.35 to 20 kHz)
Transmit polarization	Switchable from sweep to sweep between H and V
Receive polarization	Simultaneous H / V (dual-channel receiver)
Transmit power stability	$\leq \pm 0.1$ dB per second or $\leq \pm 1.0$ dB per day
Phase drift from sweep to sweep	$\leq \pm 1^\circ$ per second
Sensitivity Reflectivity	≥ 10 dBZ at 30 km range using standard weather radar data processing at 30 m range resolution
RF cable losses	1.5 dB
Radar constant	40000 (46 dB)
Antennas	
Dish diameter	1.2 m
Antenna distance	2.2 m center-to-center
Gain	38.7 dBi
HPBW	1-way 2.1 deg in azimuth and elevation (pencil beam)
Sidelobe level	first side lobe < -25 dB subsequent side lobes < -28 dB
Cross-polarization isolation	> 25 dB
Azimuth scan range	0° - 360° continuous
Elevation scan range	-2° to 90°
Angular positioning accuracy	$\pm 0.1^\circ$
Scanning speed	0 (stopped) – 6 rpm
Radome	Oratex
Receiver	
Linear dynamic range	> 70 dB
Rx gain	(at 0 dB IF attenuation) 70 dB
Noise figure	1.5 dB
IF bandwidth	30 kHz to 5 MHz (1 dB)
Signal output	Beat signal for H and V channels at 16-bit resolution

3.2. System specifications

Table.3.1 lists the manufacturer information of the Rijnmond Radar. Table.3.2 lists the specifications of Rijnmond radar. Rijnmond Radar is an FMCW radar with sawtooth modulation. The transmitted wave is at X-band with a central frequency of 9.325 GHz. It is dual-polarisation with an ATSR (Alternate Transmission Simultaneous Reception) polarimetric configuration. For PPI scanning, its default rotation speed is 6.0 degree/s, so a whole PPI measurement needs 1 min. Its sweep time is 305 microseconds, so its sampling rate or PRF is the inverse of the sweep time.

The transmitted power is 43 dBm, in linear scaling $10^{\frac{43}{10}}$ mW (19953 mW or equivalently 19.953 W). The maximum Doppler velocity equals 13.18 m/s. However, if horizontal and vertical polarizations are alternated in transmission, which is the case for the Rijnmond radar, the time between two consecutive samples with the same polarization becomes $2T_c$, which decreases the maximum Doppler velocity by a factor 2, thus 6.59 m/s. The Doppler resolution using 512 sweeps is $\frac{2 \times 6.59}{512} = 0.0257$ m/s = 2.57 cm/s.

Table 3.2: Rijnmond Radar Specifications

Transmitted power	43 dBm
Attenuation	Automatic
Transmitted polarisation	H+V
Received polarisation	H+V
Central frequency (c/λ)	9.325 GHz
Bandwidth (B)	7.5MHz
Sweep time (T_c)	610.0 micro-seconds
Range resolution (Δr)	0.02 km
Maximum range (r_{max})	30.4789 km
Doppler resolution (Δv_d)	0.0515 m/s
Maximum Doppler velocity (v_{max})	13.175953 m/s
PRF ($1/T_c$)	1.6393 kHz

Table.3.3 provides the 3 standard measurement modes: PPI, RHI and fixed. In the thesis, only the PPI mode is used.

Table 3.3: Scanning Modes

PPI	
Rotation speed	6.0 °/s
Elevation angle	Fixed
RHI	
Azimuth angle	0°
rotation speed	6.0 °/s
Elevation start angle	0°
Elevation stop angle	90°
Fixed	
Azimuth angle	10°
Elevation angle	10°

3.3. Radar location

The radar locates at a building with a height of 156 m, next to Rotterdam Central Station, Rotterdam, the Netherlands. ($51^{\circ}55'27.24''\text{N}$, $4^{\circ}28'19.04''\text{E}$). It covers almost the whole area of South Holland Province. As this province is the country's most populous, the land use here was generally planned to be urban areas. Especially its closest area is the city center of Rotterdam, the largest city of the Netherlands, high buildings result in ground clutters. In addition, the ship radio signal can contaminate its measurement, and vice versa, since Rotterdam has the largest port in Europe. Fig. 3.2 depicts the location of the radar and its coverage.



Figure 3.2: Rijnmond radar coverage area. The radius of the coverage circle is 30.48 km

Data processing and filtering

4.1. Processing pipeline

Fig.4.1 shows the overall data processing chain from raw received voltage signal to radar ray profile. The raw data is first converted into a Range-Doppler spectrogram. Spectral polarimetric processing is then conducted to remove non-hydrometeor signals. Finally, the filtered spectrogram result is summed from the whole Doppler velocity range to output a radar ray profile in the time domain.

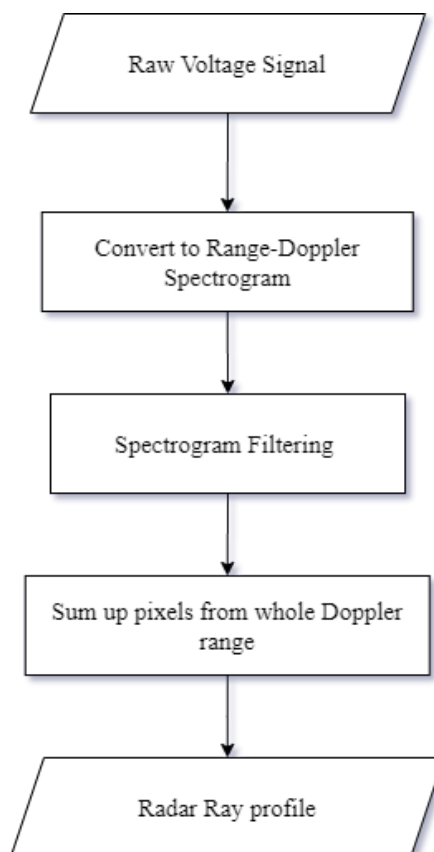


Figure 4.1: Data Processing Pipeline

4.2. Range-Doppler Spectrogram Acquisition

Before spectral polarimetric non-hydrometeor suppression, the primary step is to do a pre-processing from Level-0 data (raw voltages) to the range-Doppler spectrogram. A range-Doppler polarimetric spectrogram is a 2-D image whose intensity is one of the polarimetric parameters. The two dimensions of a spectrogram consist of the range from 0 to r_{max} and the Doppler velocity from $-v_{max}$ to v_{max} . The received signal from the input data files is represented in two components, In-phase component (I) and Quadrature component (Q). The first step is the conversion from those two components to complex signals for all polarizations, namely, HH, VH, VV and HV. Following this, the input signals are reshaped from vector forms to matrix arrays with a dimension determined by its range and Doppler resolutions. Next, a given number of samples of the input signal is removed. Then, the mean of the whole signal set is subtracted to the signal. Before the 2D-Fast Fourier Transform (FFT), a 2D Hamming window is applied to avoid leakage for FFT. Finally, the time-domain image is transformed to the frequency domain by FFT, and the zero velocity is shifted to the middle.

After the 2D-FFT, the complex Range-Doppler spectrogram S_{XY} is obtained with X and Y being the polarization in reception and transmission, respectively. From S_{XY} , the corresponding spectral reflectivity is defined as:

$$sZ_{XY}(r, v) = C_r \cdot sP_{XY}(r, v) \cdot r^2 = C_r \cdot |S_{XY}(r, v)|^2 \cdot r^2 \quad (4.1)$$

where C_r represents a calibration constant and sP_{xy} is the spectral power.

The other four spectral polarimetric parameters, $sZDR$, sL_{DR}^{HH} , sL_{DR}^{VV} , $s\rho_{co}$ can be defined based on S_{XY} and sZ_{XY} :

$$sZ_{DR}(r, v) = 10 \log_{10} \left(\frac{sZ_{HH}(r, v)}{sZ_{VV}(r, v)} \right) \quad (4.2)$$

$$sL_{DR}^{HH}(r, v) = 10 \log_{10} \left(\frac{sZ_{VH}(r, v)}{sZ_{HH}(r, v)} \right) \quad (4.3)$$

$$sL_{DR}^{VV}(r, v) = 10 \log_{10} \left(\frac{sZ_{HV}(r, v)}{sZ_{VV}(r, v)} \right) \quad (4.4)$$

$$s\rho_{co}(r, v) = \frac{|\langle S_{HH}(r, v) S_{VV}^*(r, v) \rangle|}{\sqrt{\langle |S_{HH}(r, v)|^2 \rangle \langle |S_{VV}(r, v)|^2 \rangle}} \quad (4.5)$$

where $\langle \rangle$ is the 2D averaging with consecutive Range-Doppler pixels. It is equivalent to do a cross-correlation operation with a size of $n \times n$ kernel. The $n = 3$ is chosen here to retain its highest resolution, while the size of the kernel is possible to change if $s\rho_{co}$ is used for spectral polarimetric filtering.

4.3. Description of two specific filters

This section introduces two spectral polarimetric filtering techniques. Both filters apply double sL_{DR} thresholding for spectral polarimetric filtering ([1]). Several parameters involved in the filtering techniques are empirical.

4.3.1. Default Rijnmond filtering

The Default Rijnmond filtering first uses a notch filter at 0 m/s to suppress ground clutter. Then, a noise clipping operation is done. The spectral power smaller than the averaged noise power + 3 dB, is discarded. Next, a double sL_{DR} thresholding is applied. The pixels whose sL_{DR}^{HH} and sL_{DR}^{VV} values are larger than a magnitude of -7 dB are labelled as non-hydrometeors. Subsequently, there is a speckle filter followed by the removal of narrow Doppler spectra.

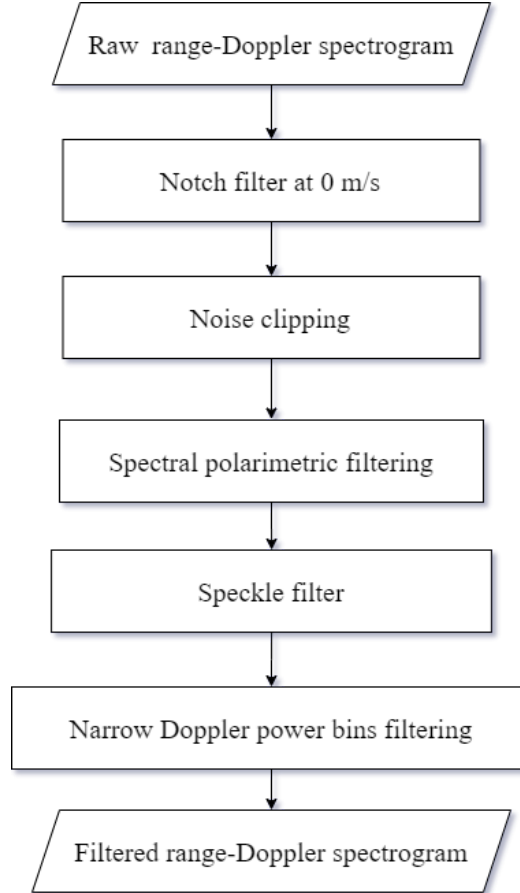


Figure 4.2: The default Rijnmond filtering algorithm

4.3.2. Object-orientated spectral polarimetric (OBSpol) filtering

In OBSpol filter ([5]), the first step is the spectral polarimetric filtering. It uses either double sL_{DR} or $s\rho_{co}$ thresholding. The double thresholding is the same as the Default Filter with a magnitude of -7 dB. The $s\rho_{co}$ is an alternative when sL_{DR} is not available. The mechanism of $s\rho_{co}$ thresholding is that those pixel values larger than 0.95 are labelled as hydrometeors. Then the mathematical morphology method is used to recover the missing precipitation pixels. The operation of the mathematical morphology is Closing by a disk-shaped structural element with a radius of 3-pixel length. The size of the structural element is empirically determined. The same size of 3-pixel length will be used in this thesis. Based on the expectation of range-Doppler continuity of precipitation spectrogram, the eight largest hydrometeor objects (or areas) are selected. They are eight in order to take into account possible Doppler aliasing. Finally,

narrow Doppler power spectra, with a Doppler spectrum width smaller than 42 cm/s (empirical threshold), are removed.

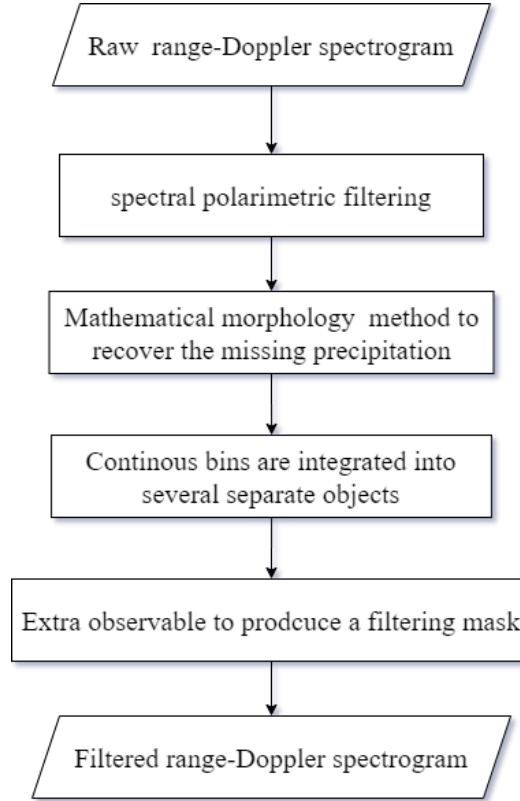


Figure 4.3: OBSpol filter algorithm

4.4. Radar ray profile and PPI

4.4.1. Radar Profile

After the filtered polarimetric spectrogram is obtained, the next step is to sum all the spectral reflectivity pixels in the same range. Therefore, a range-Doppler spectrogram ($sZ_{XY}(r, v)$) is converted to a one-dimensional array ($Z_{XY}(r)$) of which r varies from 0 to r_{max} with the step Δr .

$$Z_{XY}(r) = \int_{-v_{max}}^{v_{max}} sZ_{XY}(r, v) dv \quad (4.6)$$

where $sZ_{XY}(r, v)$ denotes the spectral reflectivity. From the reflectivity, Z_{XY} , Z_{DR} and L_{DR} can be calculated.

By calculating $\overline{C_{co}}(r)$, the co-polar correlation coefficient can be estimated:

$$\overline{C_{co}}(r) = C_r \cdot r^2 \int_{-v_{max}}^{v_{max}} S_{HH}(r, v) S_{VV}^*(r, v) dv \quad (4.7)$$

$$\rho_{co}(r) = \frac{|\overline{C_{co}}(r)|}{\sqrt{Z_{HH}(r)Z_{VV}(r)}} \quad (4.8)$$

4.4.2. PPI

Plan position indicator (PPI) is one of the most common ways to display radar profiles. The radar scans in azimuth. In this display, the north direction is considered as the origin (0°) and the directional degree or azimuth increases in a clockwise direction.

4.4.2.1. Input data and their PPI

The input data for case studies are from three scenarios, namely, clear air, light rain and heavy rain. The clear air refers to all the reflected signals from non-hydrometeors and noise. And the two rainy cases are distinguished by their rainy intensity.

Table.4.1 lists the start and end measurement time and elevation angles for the three input files. One minute measuring duration covers all directions. These three input data have the same radar specification discussed in the previous chapter except for the attenuation of the Heavy rain case and their own elevation angle. The attenuation alteration intended to resolve those radar rays with saturation problems. A radar ray suffering saturation is the radar ray whose whole range bins have incredible high reflectivity (Z_{HH}). Besides, the elevation angle of the heavy rain case are higher than the other two. Therefore, it will suffer less ground clutter.

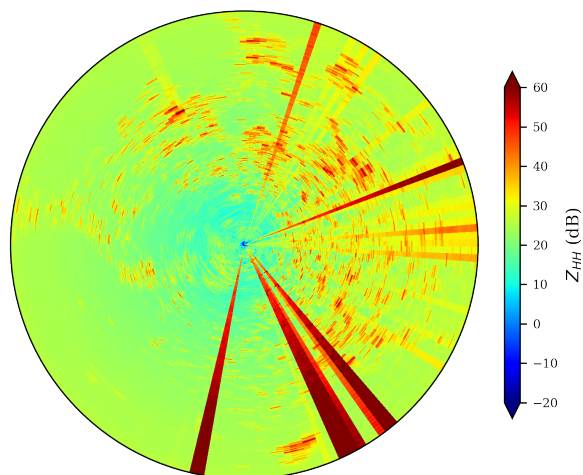
Table 4.1: Measurement time and elevation angles of the three study cases

Test case	Clear Air Echo	Light Rain	Heavy Rain
Elevation	1.22°	1.08°	2.74°
Attenuation	Fixed	Fixed	Automatic
Start time	2020-02-05, 15:45:55	2020/2/4, 13:09:20	2020/6/17, 20:23:39
End time	2020-02-05, 15:46:56	2020/2/4, 13:10:21	2020/6/17, 20:24:40

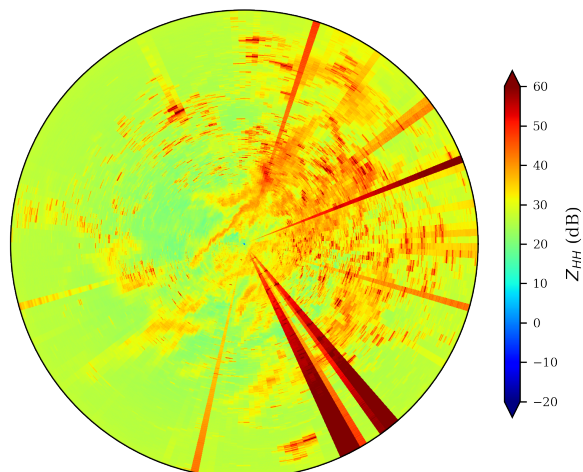
In the clear air case (Fig.4.4(a)), all echoes are from non-hydrometeors. Those received echoes can be strong (in red) and weak (dark blue). The weak echoes are mostly located close to the radar (noise). Strong non-hydrometeor echoes are those with a shape of separated red pixels in the PPI. Also, there are some radar rays with saturation issues whose whole range bins have very high Z_{HH} values.

With reference to the clear air echo PPI, the light rain (Fig.4.4(b)) looks similar to the PPI of the clear air with the addition of some hydrometeor objects covering larger areas in range and azimuth, but with weak reflectivity values.

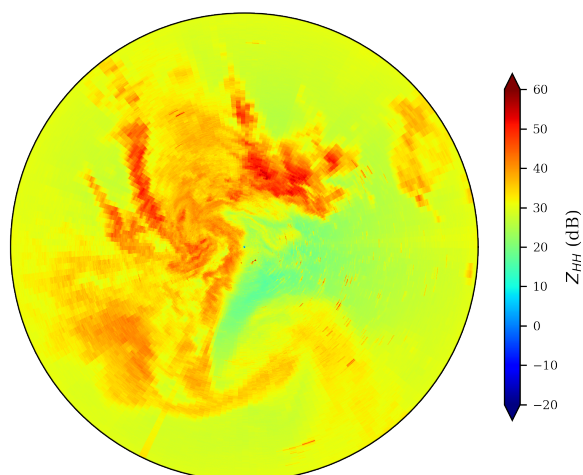
The hydrometeor area is large with high reflectivity values in the heavy rain case (Fig.4.4(c)). Because of higher elevation angles and different attenuation settings, those red separated pixels are less.



(a) The clear air PPI



(b) The light rain PPI



(c) The heavy rain PPI

Figure 4.4: Raw data PPI of the three study cases

5

Methodology

This chapter illustrates the method used in this project which consists of three main parts. The first part (section 5.1 to 5.2) explains the foundations for designing the new algorithm and evaluation methods. The foundations include the understanding and analysis of the features of polarimetric spectrograms. The second part (section 5.3) is the derivation of a new filter. It is derived from the results of previous sections. Examples of this new filter in rainy, clear air and saturation cases are given. The third part (section 5.4 to 5.6) develops the evaluation methods by defining the criteria of good results first.

5.1. Image Feature Observation and Quantification

As concluded from Chapter 2, Image Segmentation and Spectral polarimetric processing have connections in between that is to partition an input image into different parts of interest. In Computer Vision ([23]), feature description or extraction is an essential step before implementing more complicated techniques like Image Segmentation [7], as those extracted features are the foundations for discriminating against different objects. In the context of spectral polarimetric non-hydrometer or filtering, those features are those properties that can be used to distinguish hydrometeors and non-hydrometeors in a spectrogram. The first step is to observe those features from plotting polarimetric spectrograms by human perceptions. After an intuitive understanding of those feature differences between hydrometeors and non-hydrometeors, then quantify those observed features by numerical variables that can be processed, used to design algorithms and evaluations.

Before plotting and inspecting polarimetric spectrograms, the way of plotting should be considered. Each polarimetric spectrogram image can be equivalent to a grayscale image. However, due to the difference of intensity range and distribution between normal images and polarimetric spectrogram, the scaling for plotting is dissimilar. For example, a normal 8-bit image is fixed that is always bounded from 0 to 255, but the bounds of a polarimetric spectrogram are not fixed due to its correlation with physical properties of the measured medium. The scaling for plotting should be cautiously selected, as it will influence human's perception of understanding features.

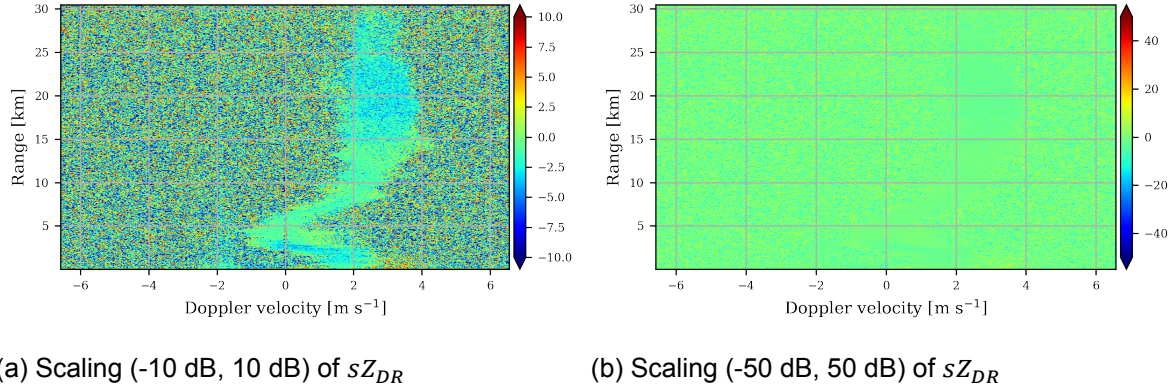


Figure 5.1: Example of scaling impact on feature observation

Fig.5.1 shows an example of the impact of scaling for understanding features in a spectrogram containing hydrometeor objects. The range is too large in Fig.5.1 (b), so the features are not distinguishable between hydrometeors and non-hydrometeors. The scalings for those spectrogram plotting are empirical and listed in Table.5.1.

Table 5.1: Scaling related to physical properties of precipitation

Polarimetric Parameters	Range for Plotting
sZ_{HH}	[-30 dBZ, 50 dBZ]
sZ_{DR}	[-10 dB, 10 dB]
sL_{DR}	[-30 dB, 5 dB]
$s\rho_{co}$	[0, 1]

After the suitable scaling are confirmed, generating and observing spectrograms are conducted. Different types of spectrograms, namely, clear air echo, rainy, and those with saturation problems should be inspected. A clear air echo is that spectrogram without any hydrometeor pixels within. The spectrogram with saturation problems exhibits large Z_{HH} values in the time domain at all ranges. The order of observing spectrograms follows the scale from the global, regional and local. The global method is used to have overviews of different types of spectrograms, for example, concerning the similarity and difference of major intensity distribution of all polarimetric parameters. Then putting more attention to the spectrograms of rainy cases to contrast the disparity between the hydrometeors and non-hydrometeors, such as intensity and roughness of each object. Finally, move to a little scale by focusing on the borders or edges between hydrometeors and non-hydrometeors. Not like the observing way as the larger scales before, the change of polarimetric parameters properties is regarded along the edges.

5.1.1. Examples of Spectrogram Observations

Figs 5.2, 5.3 and 5.4 showed three different polarimetric spectrograms in three different scenarios. Overall, all range-Doppler spectrograms have a noisy background (non-hydrometeors) while rain objects retain uniformity within. This uniformity can be identified by intensity or smoothness. For example, in Fig. 5.2, hydrometeor regions have different colours from the noisy background of sZ_{HH} , $s\rho_{co}$ and sL_{DR}^{HH} . In sZ_{DR} , the visual difference between hydrometeors and non-hydrometeors is more obvious from the smoothness that hydrometeor region has smoother texture. And the pixel areas of hydrometeors are much smaller than the that of non-hydrometeor.

Locally, the boundary between hydrometeor and non-hydrometeor regions can be observed by the change of intensity and smoothness. Intensity change can be observed in sZ_{HH} , $s\rho_{co}$ and sL_{DR}^{HH} , while the smoothness change is for sZ_{DR} . For sZ_{HH} , its noisy background has a gradient, especially in the far range areas that the boundary between hydrometeor and non-hydrometeor becomes vague. $s\rho_{co}$ and sL_{DR}^{HH} have a stable noisy background, while some part of the boundaries are also ambiguous. The boundary looks clearer in sZ_{DR} , as it has a sharp change of smoothness.

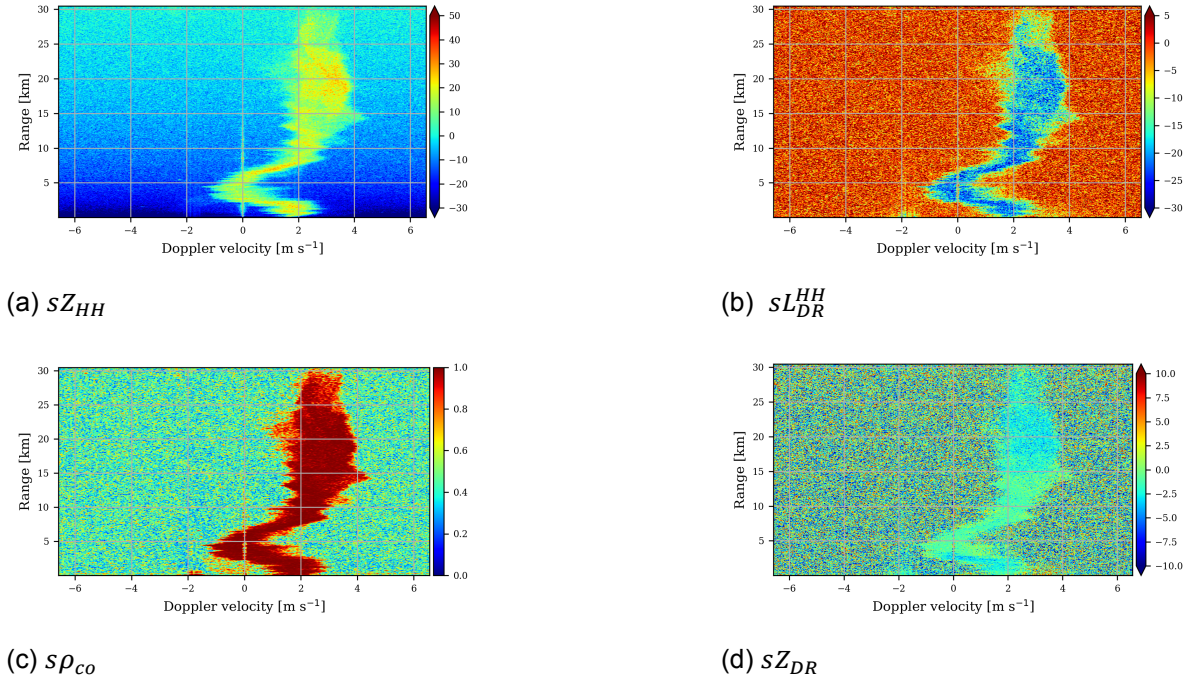


Figure 5.2: Spectrograms of heavy rain. We see mainly two objects: background noise and rain

Fig 5.4 shows one example suffering saturation problem from an either clear air or rainy case. As can be seen, the distribution of sL_{DR}^{HH} is mostly engaged in lower values. And the two double- L_{DR} based filters classifies low sL_{DR} values as hydrometeors ($sL_{DR} < 7$ dB). Therefore, sL_{DR} -based filter will fail to remove the spectrogram with saturation issue. However, the spectrogram of $s\rho_{co}$ and sZ_{DR} look similar to those from a clear air case but without ground clutter around 0 m/s (Fig.5.3(c, d)). A spectrogram can be removed once the spectrogram can be identified as a totally clear air case. This suggests that filters based on $s\rho_{co}$ or sZ_{DR} are beneficial to removing spectrograms with saturation issue should apply.

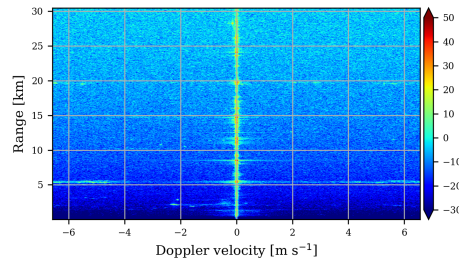
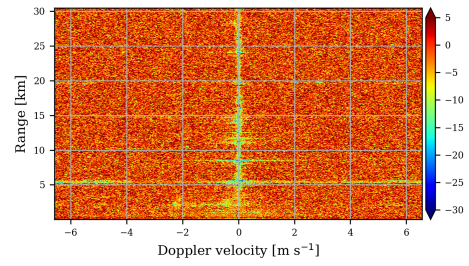
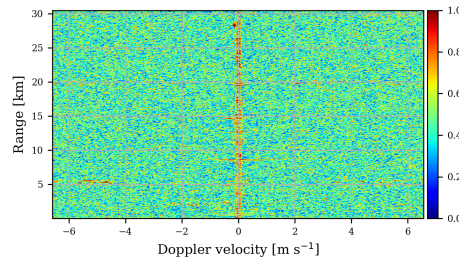
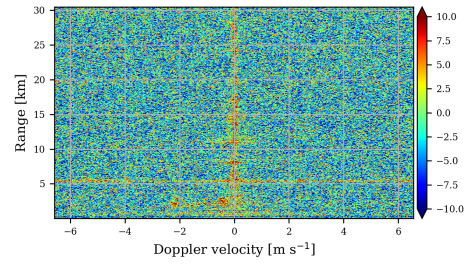
(a) sZ_{HH} (b) sL_{DR}^{HH} (c) $s\rho_{co}$ (d) sZ_{DR}

Figure 5.3: Spectrograms of clear air echo. We see two objects: background noise and ground clutter having the mean Doppler velocity 0 m/s and present at all range bins

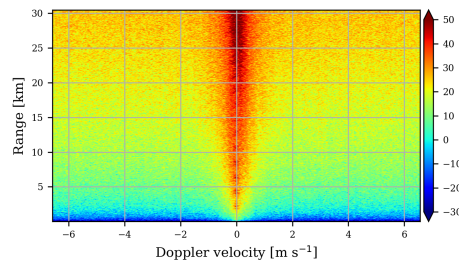
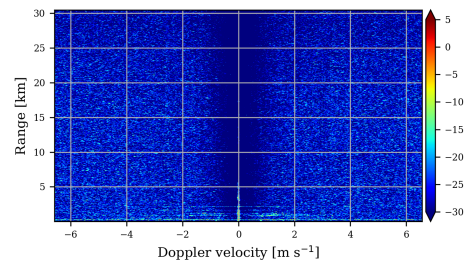
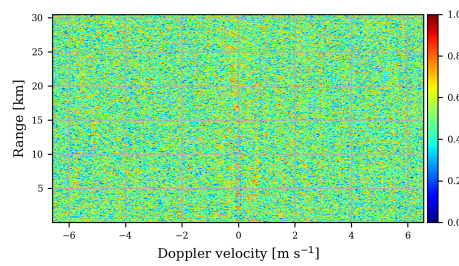
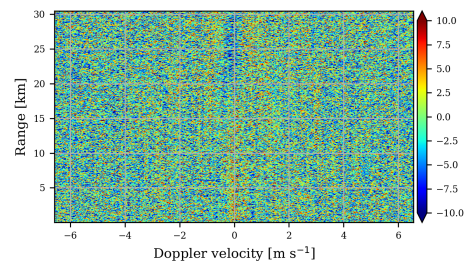
(a) sZ_{HH} (b) sL_{DR}^{HH} (c) $s\rho_{co}$ (d) sZ_{DR}

Figure 5.4: Spectrograms exhibiting the saturation problem present at some azimuth angles

5.1.2. Features Quantification

From section 5.1.1, intensity distribution, smoothness and their changes across the boundary are the main features to distinguish different types of spectrograms and regions of hydrometeor and non-hydrometeor. As those observed features will be used for designing algorithms and evaluation methods, they must be quantified first before being processed by programming. Table.5.2 lists and matches observed features to the numerical. Sections below will interpret more details the method of quantification.

Table 5.2: Quantified features

	Feature	Quantified method
Global	Intensity distribution	Histogram (section 5.1.2.1)
Regional	Intensity distribution, Smoothness	Histogram, Entropy (Eq.5.1)
Local	Smoothness gradient	Entropy gradient (Eq.5.29)

5.1.2.1. Intensity Distribution

The histogram is the easiest way to quantify the intensity distribution for the whole spectrogram, regions of hydrometeors and non-hydrometeors. As discussed before, the intensity range of each spectrogram is not bounded. The range of histogram should be cautiously selected to guarantee almost all the pixel values. Besides, the number of counts per each bin of the histogram is divided by the sum of counts so that the histogram can correspond to a probability mass function for further using. Table.5.3 lists the range and bin width of each polarimetric parameter. The bin widths are empirically determined.

Table 5.3: Histogram Range and bin widths

Polarimetric Parameters	Range	Bin width
sZ_{HH}	[-50 dBZ, 50 dBZ]	1 dBZ
sZ_{DR}	[-50 dB, 50 dB]	1 dB
sL_{DR}	[-50 dB, 50 dB]	1 dB
$s\rho_{co}$	[0, 1]	0.01

Figures below (Fig.5.4) represent the histogram examples of the four spectral polarimetric parameters from the three different cases (clear air, rainy and saturation). In common, all histograms are similar concerning the shape that is single-peaked, except for the $s\rho_{co}$ from a spectrogram containing rain, which is bimodal.

The differences between spectrograms with and without saturation issue can be seen from the three histograms that are sL_{DR}^{HH} , sL_{DR}^{VV} and sZ_{HH} . The peak centers of spectral distributions of sL_{DR}^{HH} , sL_{DR}^{VV} and sZ_{HH} from the saturation spectrogram are shifted. However, its $s\rho_{co}$ and sZ_{DR} are almost the same as those from the clear air.

Getting back to the distributions between spectrograms with and without rain, there is a hierarchy for distinguishing them. First of all, $s\rho_{co}$ is the only spectral parameter with changing its distribution pattern at the right tail. sZ_{DR} is the second distinguishable from changing its peak tip even though its shape is still bell-like. Distribution differences between the remaining parameters are subtle.

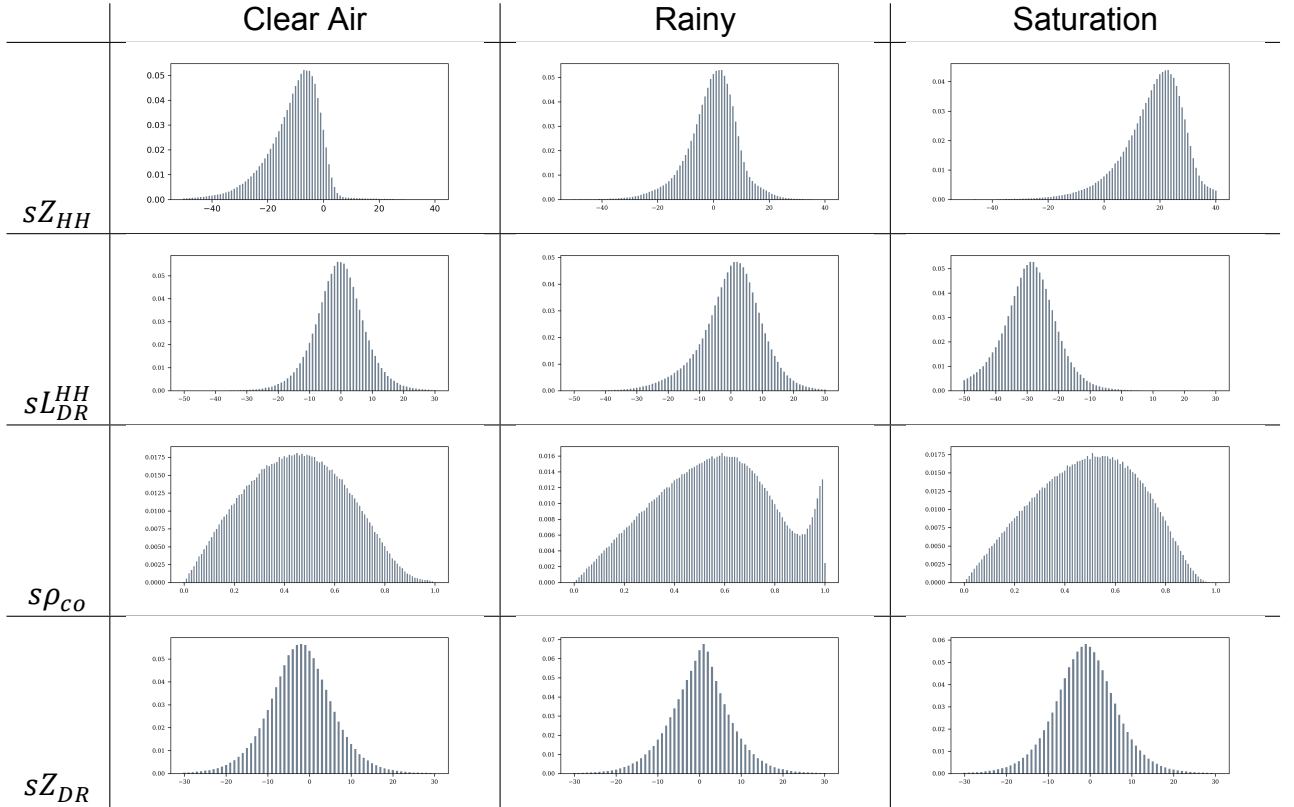


Table 5.4: Histograms of four polarimetric parameters from three different scenarios

5.1.2.2. Shannon's Entropy

Information Entropy or Shannon's Entropy ([24]) is defined as the average logarithm of the inverse of probability. There are many interpretations of information entropy depending on its application. Here, it is used as measuring the number of states of a system. A large number of states, n , leads to a high entropy, H . In the spectral polarimetric image context, event X denotes one of the polarimetric variables so $P(x_i)$ is its probability. And x_i is the i -th possible value.

$$H(X) = \sum_{i=1}^n P(x_i) \log \frac{1}{P(x_i)} \quad (5.1)$$

$$P(x_i) = Pr(X = x_i) \quad (5.2)$$

$$X \in \{sZ_{DR}, sL_{DR}^{HH}, sL_{DR}^{VV}, s\rho_{co}\} \quad (5.3)$$

Fundamental properties of Shannon's Entropy will be used here are ([25]):

1. $H(X)$ is a convex function and $H(X) \geq 0$
2. $H(X)$ reaches its maximum value, $H(X)_{max} = \log(1/n)$ when $P(x_1) = P(x_2) = \dots = P(x_n) = 1/n$. (more numbers of state)
3. $H(X)$ reaches its minimum, $H(X)_{min} = 0$, when there is only one probability exists. (single number of state)

4. Shannon entropy is extensive, i.e., for two statistical independent random variable X and Y , we have $H(X + Y) = H(X) + H(Y)$.

According to property 2 and 3, Shannon's entropy can be considered a function of the distribution of a polarimetric spectrogram. And here the probability mass function is equivalent to its histogram. Shannon's Entropy function and properties are essentially used in the filtering results assessment, see section 5.5.

5.1.2.3. Jensen–Shannon Distance

Jensen-Shannon Distance (JSD) is an entropy-based parameter and used to compare the similarity or discrepancy between distributions. The lower JSD refers to more similarity and less discrepancy between different distributions, vice versa. Jensen-Shannon Distance (JSD) is defined as the square root of Jensen-Shannon Divergence ([26]):

$$JSD(P\|Q) = \sqrt{\frac{1}{2}D_{KL}(P\|M) + \frac{1}{2}D_{KL}(Q\|M)} \quad (5.4)$$

$$D_{KL}(P\|M) = \sum_{x \in \mathcal{X}} P(x) \log \left(\frac{P(x)}{M(x)} \right) \quad (5.5)$$

$$D_{KL}(Q\|M) = \sum_{x \in \mathcal{X}} Q(x) \log \left(\frac{Q(x)}{M(x)} \right) \quad (5.6)$$

$$M = \frac{1}{2}(P + Q) \quad (5.7)$$

where P and Q are two different probability mass functions and M is the pointwise mean of P and Q . D_{KL} is defined as Kullback–Leibler Divergence. Therefore, P and Q can be normalised if they do not sum to 1.

The Jensen–Shannon divergence is bounded by 1 for two probability distributions, given that one uses the base 2 logarithm ([27]).

$$0 \leq JSD(P\|Q) \leq 1 \quad (5.8)$$

Jensen-Shannon Distance is also a significant variable. It is used to determine the most sensitive spectrogram globally (section.5.2.1) when the radar antenna receives hydrometeor signals with respect to the clear air spectrogram. The most sensitive spectral parameter is then selected for spectral filtering (section.5.3) and contamination assessment (see section.5.5.2.1).

5.2. Histogram based Thresholding Enhancement

Global thresholding is an effective and fast method to do segmentation. It is highly correlated to the intensity distribution of the image, for example, histogram. The key question for filtering non-hydrometeors while keeping hydrometeors is to find which polarimetric spectrogram(s) to be chosen and the magnitude of the threshold values.

5.2.1. Spectral polarimetric parameter sensitivity analysis

When comparing histograms of polarimetric spectrograms of rainy cases and clear air, we note differences between them of each polarimetric spectrogram (see Fig.5.4). For example, there is a spike in the right tail of a $s\rho_{co}$ histogram in the rainy case while the clear air $s\rho_{co}$ does not show such a pattern. However, different spectral polarimetric parameters have different sensitivities to the occurrence of hydrometeors in the spectrogram, so the histogram patterns are different. The most sensitive one is supposed to be selected for the thresholding.

After observing spectrogram images with and without hydrometeor objects, we could find that those non-hydrometeor pixels are similar to those within clear air spectrogram. And clear air polarimetric spectrograms are also similar to each other even though they are acquired from different radar rays.

A sensitivity test is conducted to find the most sensitive polarimetric spectrogram to the presence of hydrometeors. As defined before, Jensen-Shannon Distance (JSD) is the variable to use to assess the difference between two distributions. JSD of four spectral polarimetric parameter distributions (sL_{DR}^{HH} , sL_{DR}^{VV} , sZ_{DR} , $s\rho_{co}$) from rainy cases and those four from clear air cases are used to indicate the distribution difference. The spectral polarimetric distributions of clear air ($\overline{P_{cla}(X)}$) are obtained from the averaged distributions of a full PPI.

$$\overline{P_{cla}(X)} = \frac{1}{n} \sum_{\alpha=0}^{\alpha=2\pi} P_{cla}(X, \alpha) \quad (5.9)$$

$$X \in \{sZ_{DR}, sL_{DR}^{HH}, sL_{DR}^{VV}, s\rho_{co}\} \quad (5.10)$$

$$\alpha \in [0, 2\pi] \quad (5.11)$$

Where $P_{cla}(X, \alpha)$ is the distribution of one spectral polarimetric parameter X at location α . And α is the azimuth angle with reference to 0 at the north. n is the total number of radar rays per PPI.

Fig.5.5 shows the results of JSD between rainy polarimetric spectrogram histograms and the averaged clear air one. Overall, $s\rho_{co}$ has the highest sensitivity and no distinct sensitivity difference among the other three is found. With respect to the PPI of heavy rain (Fig.5.6), the JSD of $s\rho_{co}$ has also a positive correlation with the rain areas. For example, when the radar ray is around 100° , both JSD and reflectivity decrease dramatically. Also, they peak at around 210° .

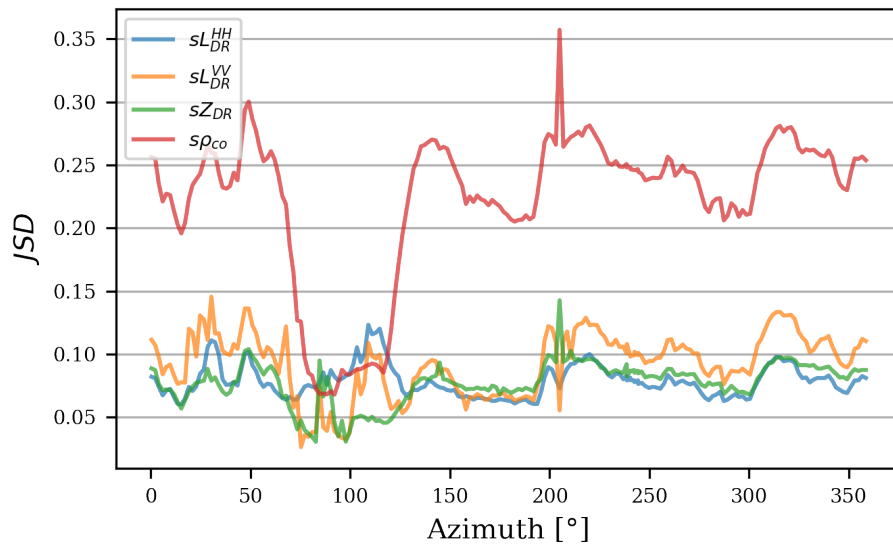


Figure 5.5: JSD of spectral polarimetric parameter distributions between rainy spectrograms and the averaged clear air

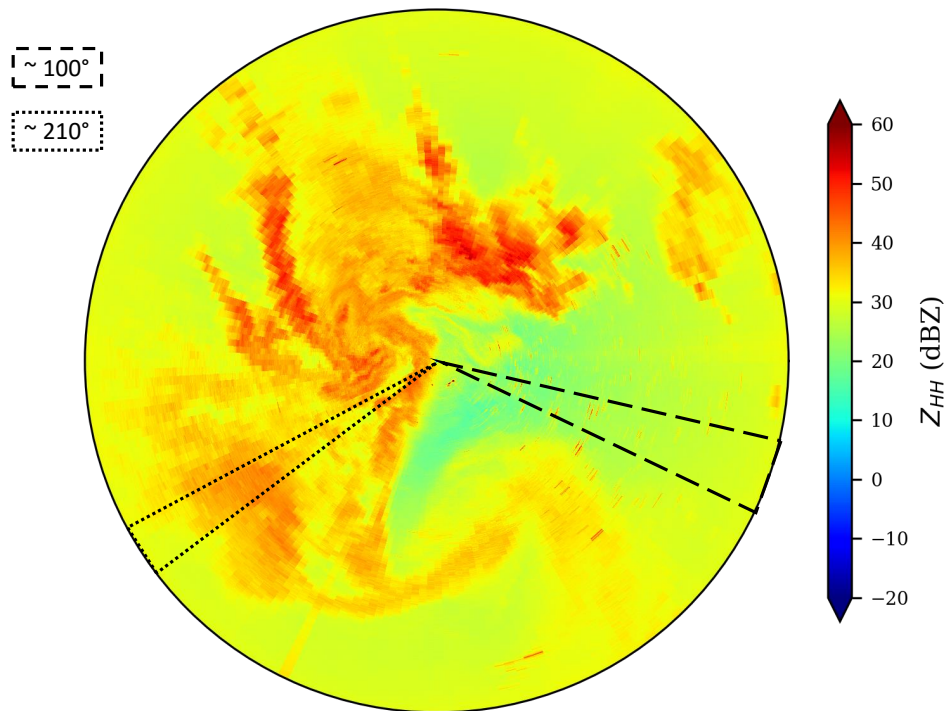


Figure 5.6: Locations at 100° and 210° of the heavy rain PPI

5.2.2. Threshold magnitude determination

The threshold magnitude should be determined by the own property of each spectrogram. In other words, the threshold magnitude is not a fixed value for all spectrograms. We assume it is due to the randomness of the event of hydrometeor and hydrometeor compositions in a spectrogram.

The occurrence of hydrometeors in one spectrogram is arbitrary. Any segmentation process including thresholding is done only if the spectrogram contains hydrometeor pixels within. If a spectrogram does not have any hydrometeor pixel or it is from pure clear air, the threshold values should be the limit to discard the total spectrogram.

Not only is the occurrence of hydrometeors arbitrary, the composition of hydrometeors can significantly vary. Different hydrometeor types have also distinct magnitudes of the same polarimetric parameter. As concluded from [28], time-domain polarimetric magnitude ranges of various types of hydrometeors are shown in Fig.5.7. Therefore, the thresholding to discriminate hydrometeors and non-hydrometeors should also be composition dependent. For example, if rain is the only composition in a spectrogram, the threshold of $s\rho_{co}$ would be around 0.95. However, if this spectrogram also includes wet snow, a $s\rho_{co} = 0.95$ would be too high. Or if it is a clear air echo spectrogram, $s\rho_{co} = 0.95$ is so low that part of non-hydrometeor signals will be classified as hydrometeors.

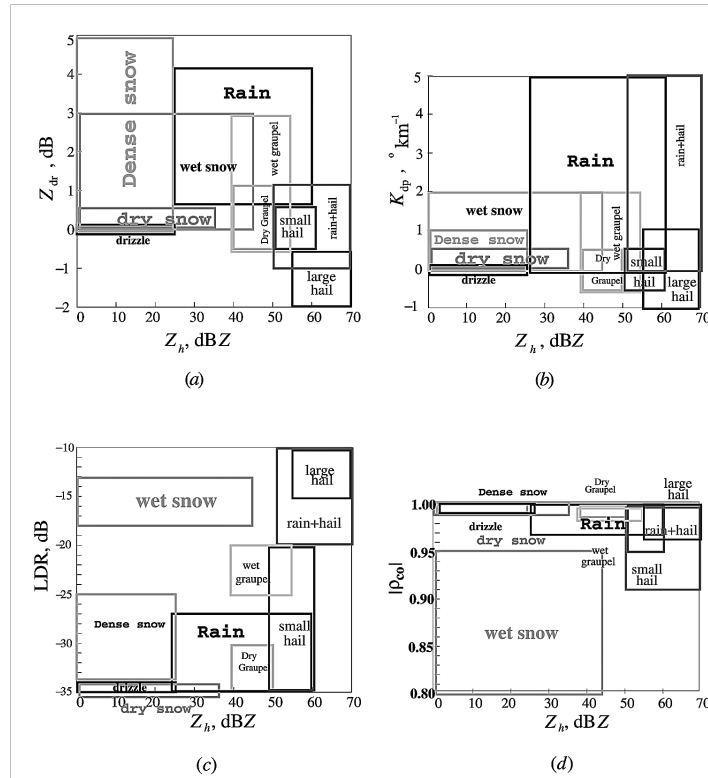


Figure 5.7: Hydrometeor classification decision boundary in Z_{HH} versus Z_{DR} (a), K_{DP} (b), L_{DR} (c), and ρ_{co} (d) space, respectively. Reference: Polarimetric Doppler weather radar: Principles and applications by V. N. Bringi and V. Chandrasekar, chap. 7, page 517. ([28])

5.3. A Jensen–Shannon Distance-based thresholding filter

A new filtering technique with reference of clear air echo spectrogram is proposed in this master thesis. Fig.5.8 shows all the steps of this filtering method. The first step consists of spectral polarimetric filtering using an optimized $s\rho_{co}$ threshold value per spectrogram. Then a zero doppler filter is implemented to filter the ground clutter. Next, missing precipitation pixels are reconstructed by mathematical morphology processing. Finally, when the percentage of hydrometeor Doppler bins is very low, full filtering is applied and the range bin has no data (NaN value). The following sections dive into details of each step, and examples of this filtering technique implemented in rainy, clear air echos and saturation spectrograms are given.

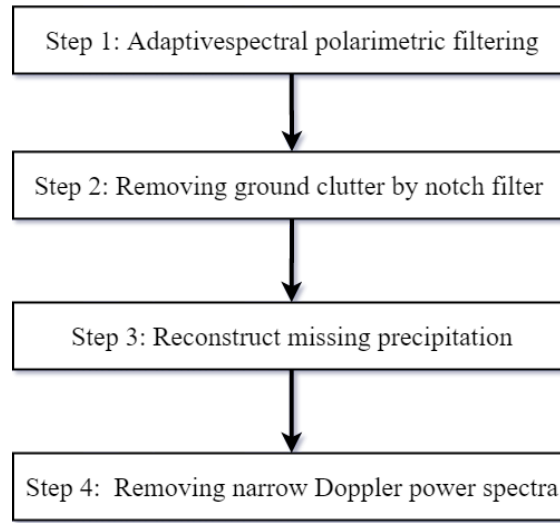


Figure 5.8: The JSD-based filtering algorithm

5.3.1. Step 1: Adaptive spectral polarimetric thresholding

The first step is to remove most of the non-hydrometeors by giving a suitable threshold value. As concluded from Section 5.2, $s\rho_{co}$ is chosen to be the spectrogram for thresholding and the magnitude should be composition dependent. A 2D-averaging with 5 consecutive Range-Doppler bins (kernel 5×5) is used to estimate $s\rho_{co}$. After the threshold value is determined, a binary mask ($M^{s\rho_{co}}(r, v)$) is created:

$$M^{s\rho_{co}}(r, v) = \begin{cases} 1, & \text{if } s\rho_{co} > T_{opt} \\ 0, & \text{otherwise} \end{cases} \quad (5.12)$$

where T_{opt} refers to the threshold values of $s\rho_{co}$, and **1** is the label for hydrometeors and **0** is for non-hydrometeors. T_{opt} is determined for each spectrogram.

The critical part of this step is to determine the magnitude of $s\rho_{co}$ for thresholding. The optimized threshold magnitude, T_{opt} is determined by the $s\rho_{co}$ distribution similarity and disparity between an input spectrogram and the averaged distribution of clear air echo spectrogram ($\overline{P_{cla}(s\rho_{co})}$). It is based on the spectrogram observation results (see section 5.1, Fig.5.3 and Fig.5.4) where the non-hydrometeor part of

a spectrogram is similar to the spectrogram of clear air echoes. Thus, the $s\rho_{co}$ distribution of the non-hydrometeors is similar to the distribution of a clear air spectrogram. In addition, there is a great disparity of polarimetric parameters distribution between hydrometeor objects and the clear air. Those similarities and disparities are measured by Jensen-Shannon Distance (JSD).

$$JSD_U = JSD(P(s\rho_{co})_{M^{s\rho_{co}=0}} || \overline{P_{cla}(s\rho_{co})}) \quad (5.13)$$

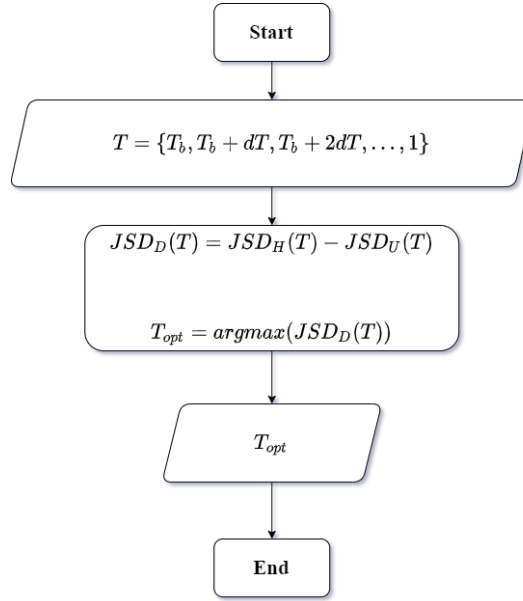
$$JSD_H = JSD(P(s\rho_{co})_{M^{s\rho_{co}=1}} || \overline{P_{cla}(s\rho_{co})}) \quad (5.14)$$

$$JSD_D = JSD_H - JSD_U \quad (5.15)$$

where JSD_U denotes the JSD between the $s\rho_{co}$ distribution of the non-hydrometeors part of an input spectrogram ($P(s\rho_{co})_{M^{s\rho_{co}=0}}$) and the averaged $s\rho_{co}$ distribution of clear air echoes spectrograms ($\overline{P_{cla}(s\rho_{co})}$). And JSD_H is the JSD between $s\rho_{co}$ distribution of the hydrometeors part ($P(s\rho_{co})_{M^{s\rho_{co}=1}}$) and $\overline{P_{cla}(s\rho_{co})}$.

As the binary mask is a function of T and JSD is a function of the mask, the JSD corresponds to a function of T . We expect the extracted non-hydrometeor $s\rho_{co}$ distribution to be more similar to the clear air echo one and more discrepancy between the distributions of the extracted hydrometeor part and the clear air echo. In other words, it prefers a smaller JSD_U but larger JSD_H . Therefore, the optimized threshold value (T_{opt}) is selected when the difference (JSD_D) between the disparity and similarity reaches the maximum.

Fig.5.9 shows the algorithm to find the suitable threshold value, T_{opt} . According to the algorithm, the input is a suitable range of $s\rho_{co}$ threshold values involving hydrometeors from Fig.5.7. Thus, the default lower bound of this range, T_b is set to be 0.8 and the maximum is 1. The increment, dT , is a variable affecting the tolerance of the result as well as the operation time, so it depends on the user's demand. As has been explained previously, the three JSD variables (JSD_H , JSD_U and JSD_D) are functions of the T and we expect T_{opt} is chosen when JSD_D reaches the maximum. Therefore, $\text{argmax}(JSD_D(T))$ denotes the value of T at which JSD_D is maximum. Examples of this algorithm implementation on rainy and clear spectrograms are shown in sections 5.3.5.1 and 5.3.5.2.

Figure 5.9: The optimised $s\rho_{co}$ threshold value determination process.

5.3.2. Step 2: Removing ground clutter by notch filter

After thresholding, it may still remain some non-hydrometeor signals which have similar polarimetric signatures compared to hydrometeor's ones. At some ranges, there is still ground clutter present. Then a notch filter around 0 m/s is accordingly implemented:

$$M^G(r, v) = \begin{cases} 1, & \text{otherwise} \\ 0, & \text{if } |v| < v_G \end{cases} \quad (5.16)$$

where v denotes the Doppler velocity of the whole spectrogram and v_G is the Doppler velocity limit for ground clutter. As ground clutter has narrow Doppler velocity spectra close to 0 m/s, v_G should be close to 0 m/s. Here, the value of the notch filter from Default Rijnmond radar filtering is applied, namely, $v_G = 0.2$ m/s.

Then, a mask is created based on the steps 1 - 2:

$$M^{step\ 2} = M^{s\rho_{co}}(r, v) \circ M^G(r, v) \quad (5.17)$$

where $\circ$ denotes the Hadamard product or element-wise multiplication.

5.3.3. Step 3: Reconstruct missing precipitation

Spectral polarimetric thresholding may remove some hydrometeor pixels with low $s\rho_{co}$ magnitude. We assume that hydrometeor pixels are continuous within the spectrogram. Therefore, filling holes inside is necessary.

$$M^{step\ 3} = M^{step\ 2} \oplus q \quad (5.18)$$

where $\oplus$ denotes the dilation operation¹ from mathematical morphology processing. And q is a square-shaped structural element with size of 1×1 .

¹The dilation operation is done by the algorithm `scipy.ndimage.morphology.binar_fill_holes`. For details, see [29].

5.3.4. Step 4: Removing narrow Doppler power spectra

The final step is to remove those narrow Doppler power spectra. The width of Doppler power spectra is determined by the number of pixel in the Doppler domain ($L(r)$). When the number of spectra pixels is lower than a ratio of total spectra pixels at a certain range, the whole range of spectra will be set to NaN. Again, the value of this ratio is the same as that used in the Default Rijnmond radar filtering, $R_{nd} = 2\%$.

$$L(r) = \int_{-v_{max}}^{v_{max}} M^{step\ 3}(r, v) dv \quad (5.19)$$

$$M^{nd}(r, v) = \begin{cases} 1, & \text{if } \frac{L(r)}{N_{Doppler}} > R_{nd} \\ 0, & \text{otherwise} \end{cases} \quad (5.20)$$

$$M^{step\ 4}(r, v) = M^{step\ 3} \circ M^{nd} \quad (5.21)$$

where $L(r)$ denotes the number of pixels labelled by 1. $M^{nd}(r, v)$ is the label mask with narrow Doppler power spectra removed.

5.3.5. Examples of implementing JSD-based Filtering

The following figures, 5.10 - 5.16, illustrate the implementation of JSD-based filtering by representing binary masks produced by each step. Spectrograms of rain, clear air and saturation are discussed.

5.3.5.1. Rainy Case

As can be seen in Fig.5.10, both JSD_H and JSD_U increase monotonically with $s\rho_{co}$ threshold value (T). A higher T contributes to adding more potential hydrometeor pixels to the non-hydrometeors group. Meanwhile, the hydrometeor pixels are tending to be purely hydrometeors. Therefore, larger T results in decrease of similarity between the non-hydrometeor part of this input spectrogram and the clear air echo spectrogram of reference (larger JSD_U). And the hydrometeor group has more substantial disparity with the reference clear air echo spectrogram.

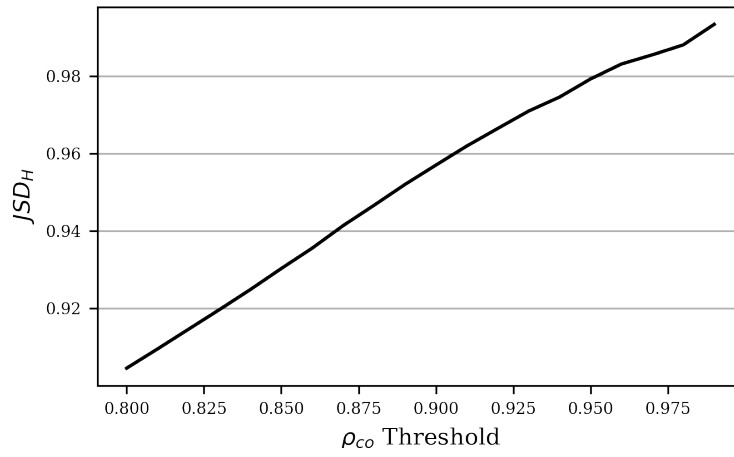
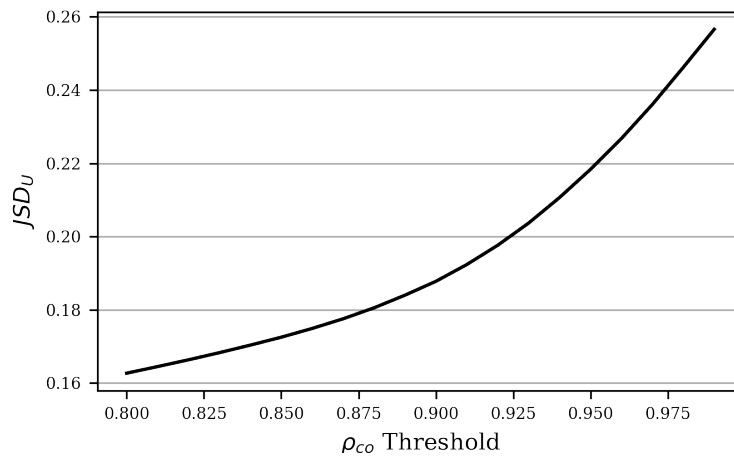
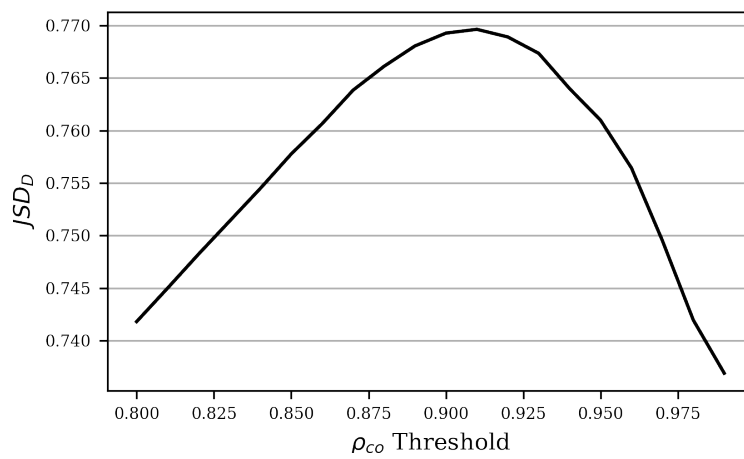
(a) JSD_H Profile(b) JSD_U Profile(c) JSD_D Profile

Figure 5.10: JSD_H , JSD_U and JSD_D profiles versus $s\rho_{co}$ for T_{opt} determination of the rainy spectrogram. In this case, $T_{opt} = 0.91$

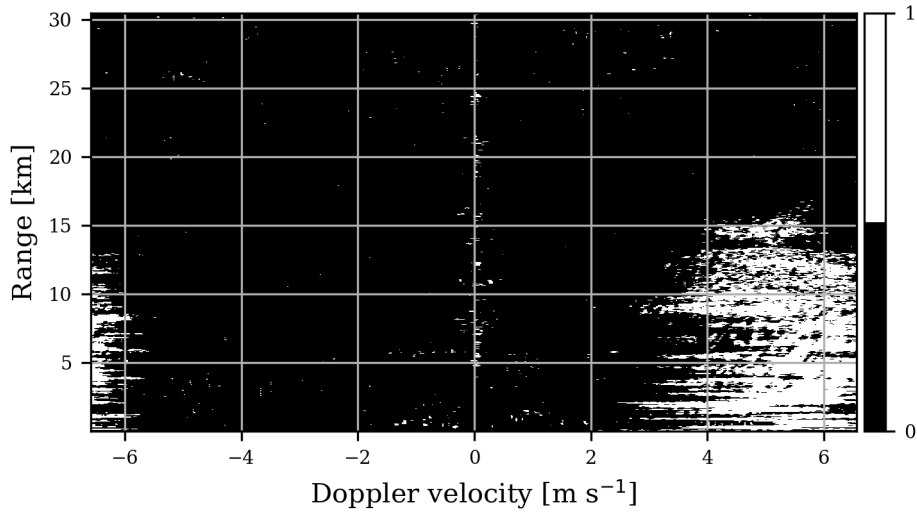


Figure 5.11: Light rain case spectrogram. $M^{SP_{co}}$ after T_{opt} selection. The hydrometeors and non-hydrometeors are labelled by 1 (white) and 0 (black), respectively.

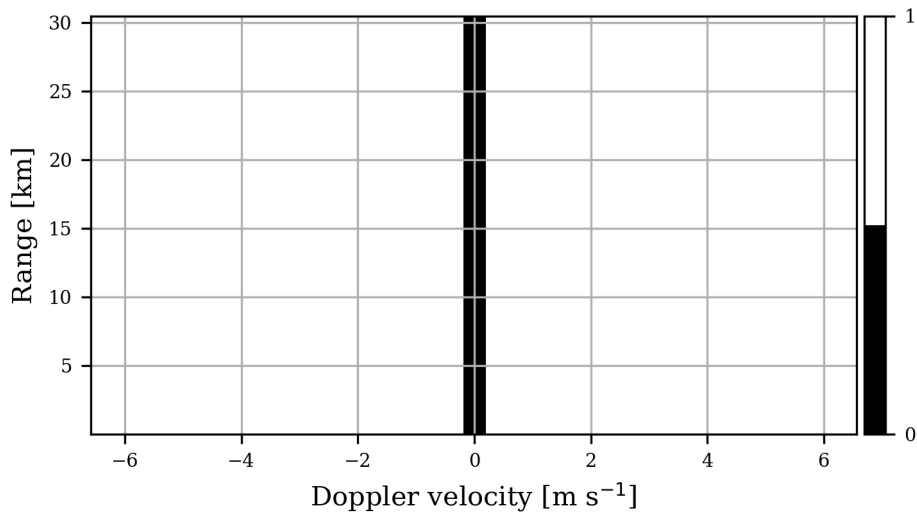


Figure 5.12: $M^G(r, v)$ with given $|v_G| = 0.2$ m/s.

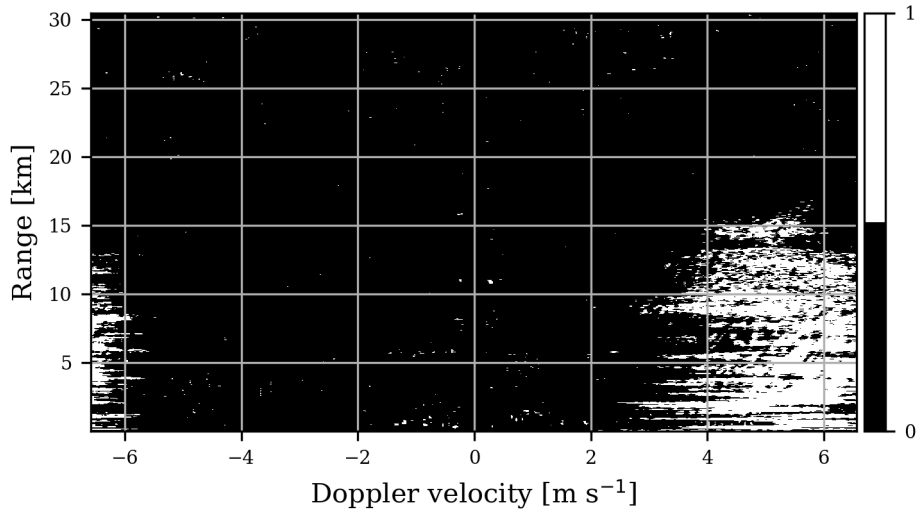


Figure 5.13: $M^{step\ 2}$ of the light rainy spectrogram. All pixels are labelled as 0 whose Doppler velocities are within $|v_G| = 0.2$ m/s.

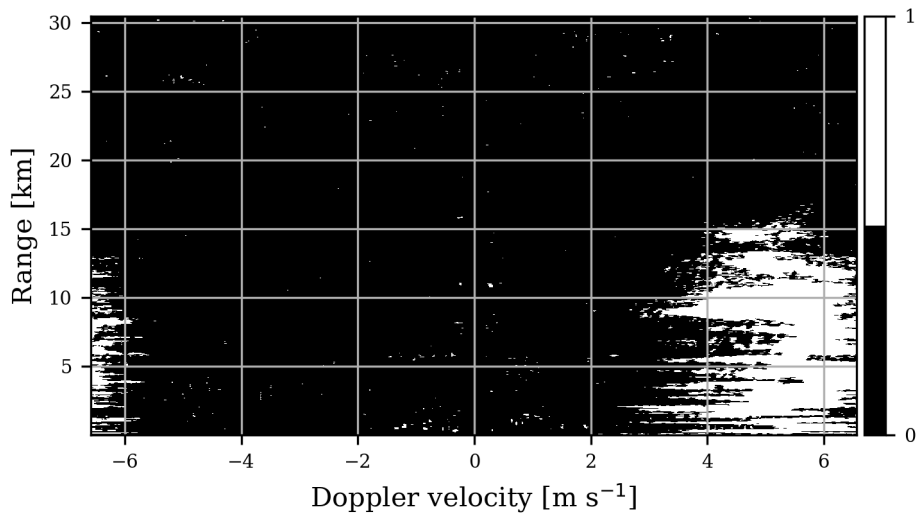


Figure 5.14: $M^{step\ 3}$ of the light rain spectrogram. In comparison with $M^{step\ 2}$, all holes inside white pixels are filled while the boundary are the same.

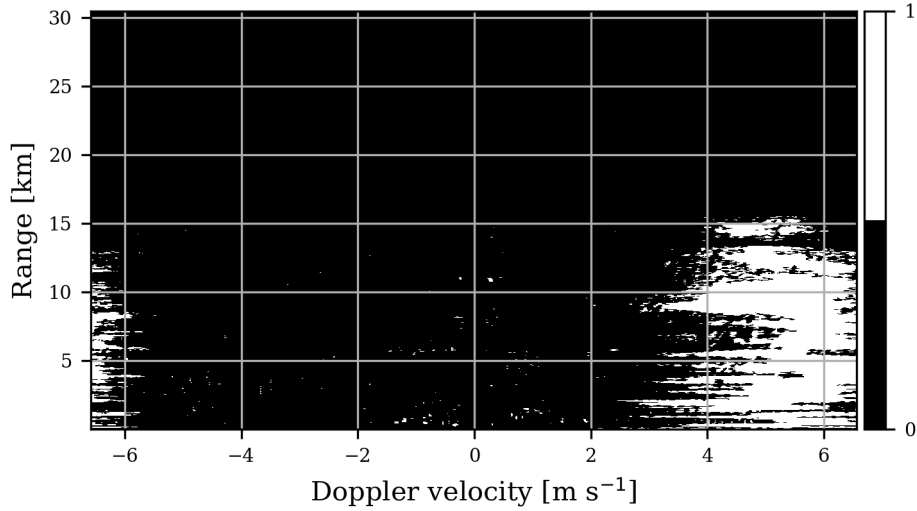


Figure 5.15: $M^{step\ 4}$ of the light rain spectrogram. Narrow Doppler spectra are removed which is equivalent to that white narrow Doppler pixels are turned to black.

5.3.5.2. Clear air and saturation Cases

In the clear air and saturation cases, their JSD_H , JSD_U and JSD_D profiles are in the same condition that matches our results of feature observation in section.5.1.

Fig. 5.16 gives examples of JSD_H , JSD_U and JSD_D profiles from a clear air or saturation spectrogram. Since the category of an input spectrogram can never be known before implementing any filtering, any input spectrogram is treated as a spectrogram with hydrometeors. As T increases, the JSD_H profile has the same tendency as a normal rainy spectrogram. However, the similarity between the extracted non-hydrometeors part and a clear air spectrogram is closer (very low values of JSD_U). And JSD_U reaches its minimum when the whole spectrogram is labelled as non-hydrometeors ($T = 1$). At the same time, JSD_D has been its maximum. Therefore, a clear air or saturation spectrogram is totally filtered by this algorithm.

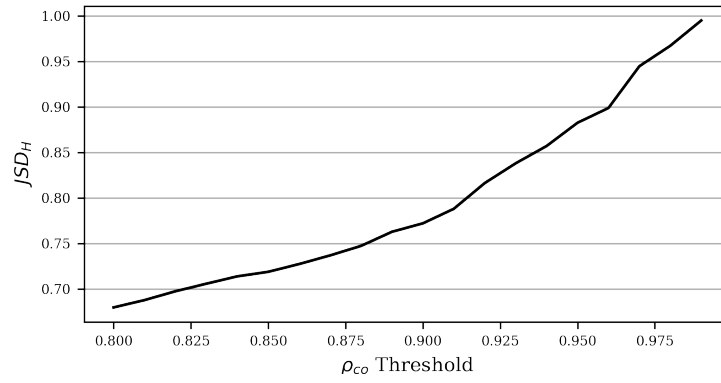
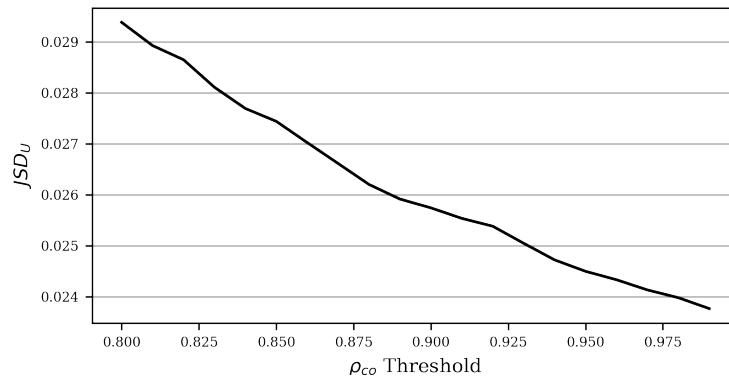
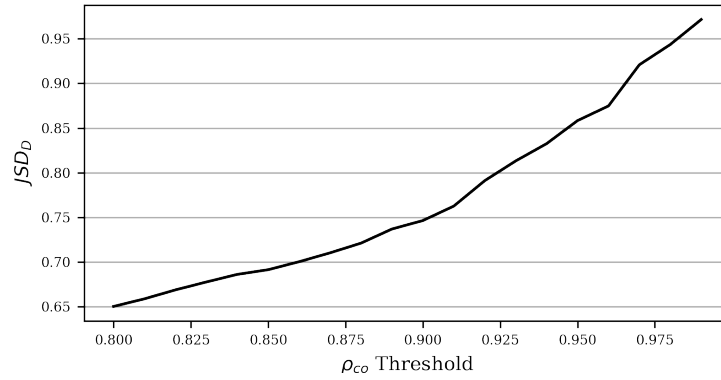
(a) JSD_H Profile(b) JSD_U Profile(c) JSD_D Profile

Figure 5.16: JSD_H , JSD_U and JSD_D profiles versus $s\rho_{co}$ for T_{opt} determination of the clear air or saturation spectrogram

5.4. Criteria for good segmentation

The number of algorithms for image segmentation is endless, we never know which one is more suitable until we define the criteria for a good segmentation. Criteria are the standards for good filtering and foundations for assessment. Besides, they also contribute to optimizing parameter options for some current filters. Without these standards, it is hard to determine whether a filter performs well or not since designing and testing different filters are those tasks continuously done through all of this project.

5.4.1. Criteria in spectral domain

[30] defined four criteria for image segmentation. These criteria give general guidelines to define the standards for the application of non-hydrometeors mitigation. Based on the assumption of Range-Doppler continuity of precipitation and the speciality of polarimetric spectrogram, those criteria should be adjusted:

1. Segmented hydrometeor and non-hydrometeors regions should be uniform and homogeneous concerning their spectral polarimetric features, for example, intensity.
2. Boundaries between hydrometeors and non-hydrometeors should be not ragged and be spatially accurate
3. Hydrometeor region interiors should be without holes and with a spatial connection

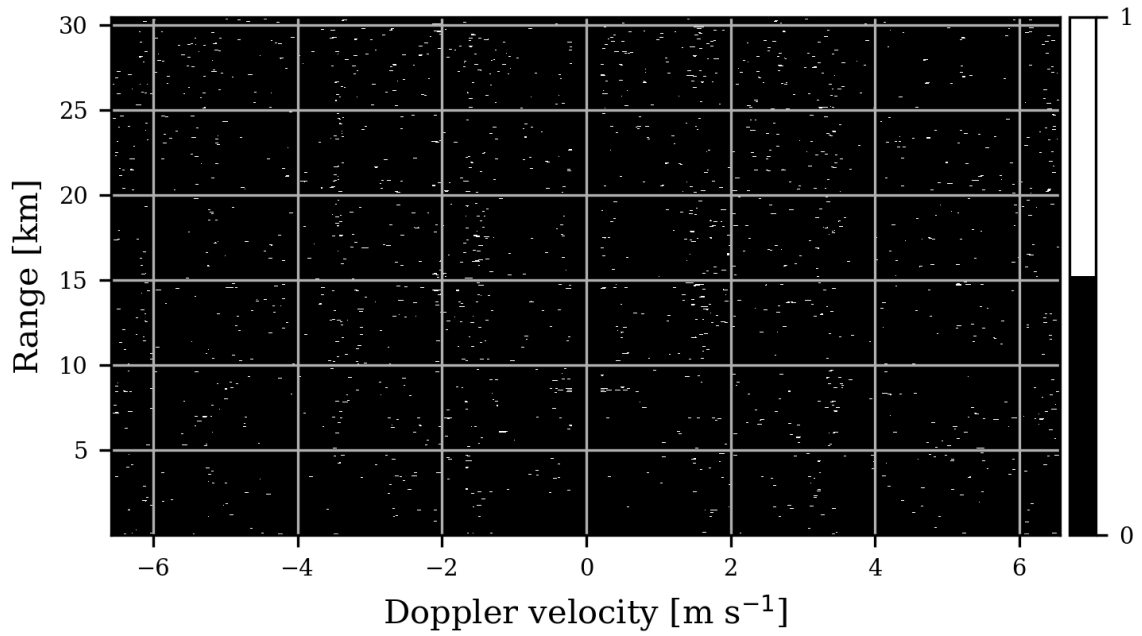
5.4.2. Criteria in time domain

The above criteria are all for spectrogram images. However, the final representation of weather radar is in time-domain instead of spectral domain. The time-domain reflectivities (Z_{HH}) at each range bin are resulting from summation on spectral pixels (Eq.4.6). This implicitly shows that if any pixel exists at one range can result in one echo bin in the time domain. Fig.5.17 shows one example of mask with high filtering result of clear air echo cases. Nonetheless, the whole range in the time domain may keep unwanted signals.

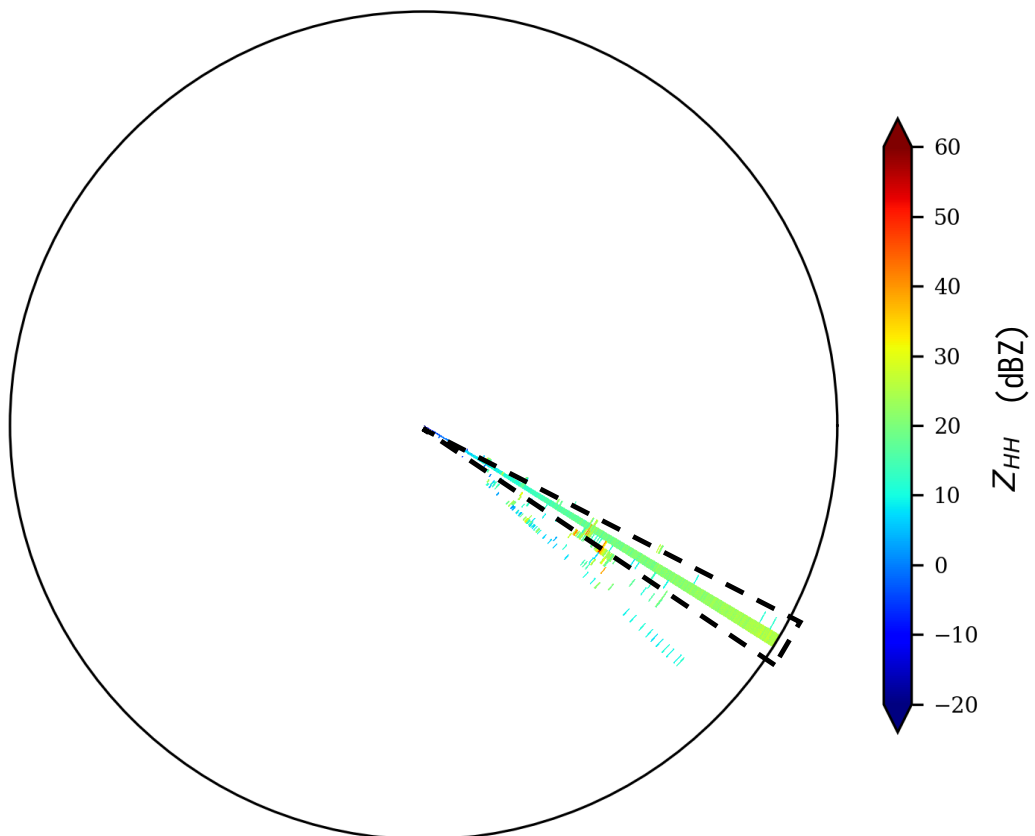
And with the assumption that rain echo bins retain continuity no matter in spectral or time domain, the existence of those single unconnected echo bins are accordingly considered bad filtering. From the conclusions of [28], Fig. 5.7, those Z_{HH} values out of 0 dBZ and 70 dBZ are potentially classified non-hydrometeors. The low ratio of those echo bins in one range is an indicator to show good filtering too. Examples of the single unconnected, the weak and the strong echo bins are shown in Fig. 5.18.

Therefore, three criteria in the time domain are:

1. All echo bins should be removed in the time domain when they relate to clear air echoes.
2. Non-hydrometeor echoes should be fewer.
3. Non-connected echo bins should be fewer.



(a) The filtering mask with 96.15% pixels removed



(b) The radar ray (dashed) with high filtering results in spectral domain

Figure 5.17: The example of a clear air echo radar ray with full range echoes while its filtering mask has a high removal ratio.

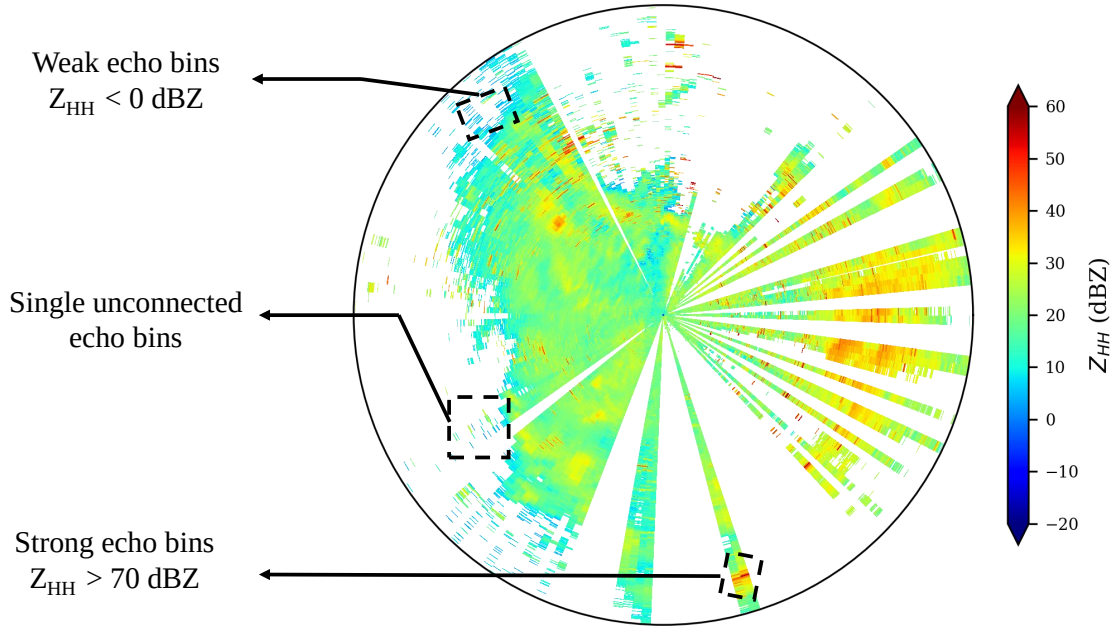


Figure 5.18: Examples of three types of unwanted echoes

5.4.3. Schematic of good and bad segmentation results

This section explains the good and bad segmentation results in a visual method. As each Range-Doppler spectrogram is a high resolution image (1525×512), subtle details like those in the boundary between hydrometeor and non-hydrometeor are hard to show. Some down scale schematic images are used for better interpretations (see section 5.4.3.2). Before interpreting the segmentation results, some image units should be defined first in section 5.4.3.1. Those image units are also used for objective evaluation methods derivation in section 5.5.2.

5.4.3.1. Image Units

1. **Labelled region**: those groups of pixels have the same label. For example, in Fig.5.19 (a), there are two different labels to group all the pixels. One is labelled by yellow and the other one by blue. If the number of labels is two and one region is much larger than the other, the larger one can be considered as background. For example, the blue region corresponds to the background.
2. **Boundary**: those curves located at the interface between different objects. (Fig.5.19 [b]) As its location is inter-pixelated, it cannot be detected directly. The boundary-line then developed to contribute to reveal its existence.
3. **Boundary-line**: the boundary-lines are the pixel series with one pixel unit width along the boundary. Since the boundary divides two objects, boundary lines have two forms with respect to the labelled object.
 - (a) **Inner-line**: With respect to the yellow region, the inner line is defined as the boundary-line locating inside the yellow region.
 - (b) **Outer-line**²: The outer line is the pixel series in the adjacent region.

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	32	33	34	35
36	37	38	39	40	41	42
43	44	45	46	47	48	49

(a) Two regions by 2 different labels

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	32	33	34	35
36	37	38	39	40	41	42
43	44	45	46	47	48	49

(b) Boundary (red) between the two regions

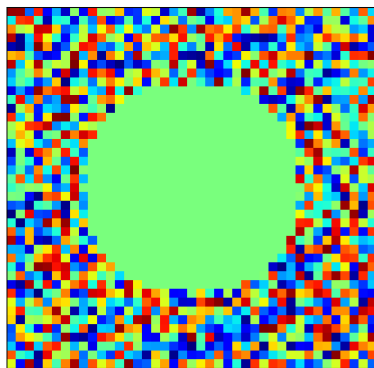
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	32	33	34	35
36	37	38	39	40	41	42
43	44	45	46	47	48	49

(c) Inner line respected to Yellow region

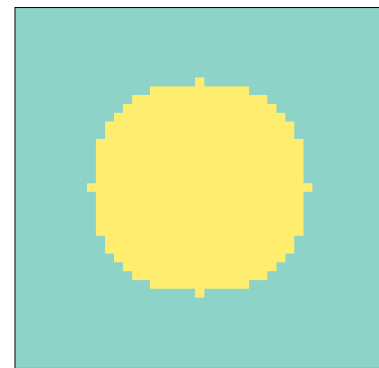
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	32	33	34	35
36	37	38	39	40	41	42
43	44	45	46	47	48	49

(d) Outer line respected to Yellow region

Figure 5.19: Basic image units



(a) Modelled image



(b) The mask with correct labels

Figure 5.20: Modelled image with a smooth object in the centre and noisy background and its correct label mask

²The algorithm of obtaining Inner-line and Outer-line from a label mask is **skimage.segmentation.find_boundaries**. For details, see [31].

5.4.3.2. Under- and Over-segmentation

As for bad segmentation, simply there are two types, namely, under-segmentation and over-segmentation. Under-segmentation refers to a segmented object containing a number of pixels which should belong to other objects. In an over-segmented object, all pixels within are labelled correctly respected to its origin, but part of its pixels are grouped to other objects.

The following schematic diagrams show good, under-, and over-segmentation by its labels, segmented results, and boundary-line profiles. Fig.5.20(a) shows a modelled image with two different regions. By human vision perceptions, it is easy to recognise that one region is the green at the centre surrounding by another noisy region from their different intensity distributions and smoothness properties. Fig.5.20(b) is the mask labelling the two regions correctly. However, an image segmentation algorithm may mis-classify part of the regions resulting in under- or over-segmentation. Fig.5.21 show two mis-classification examples with their label masks.

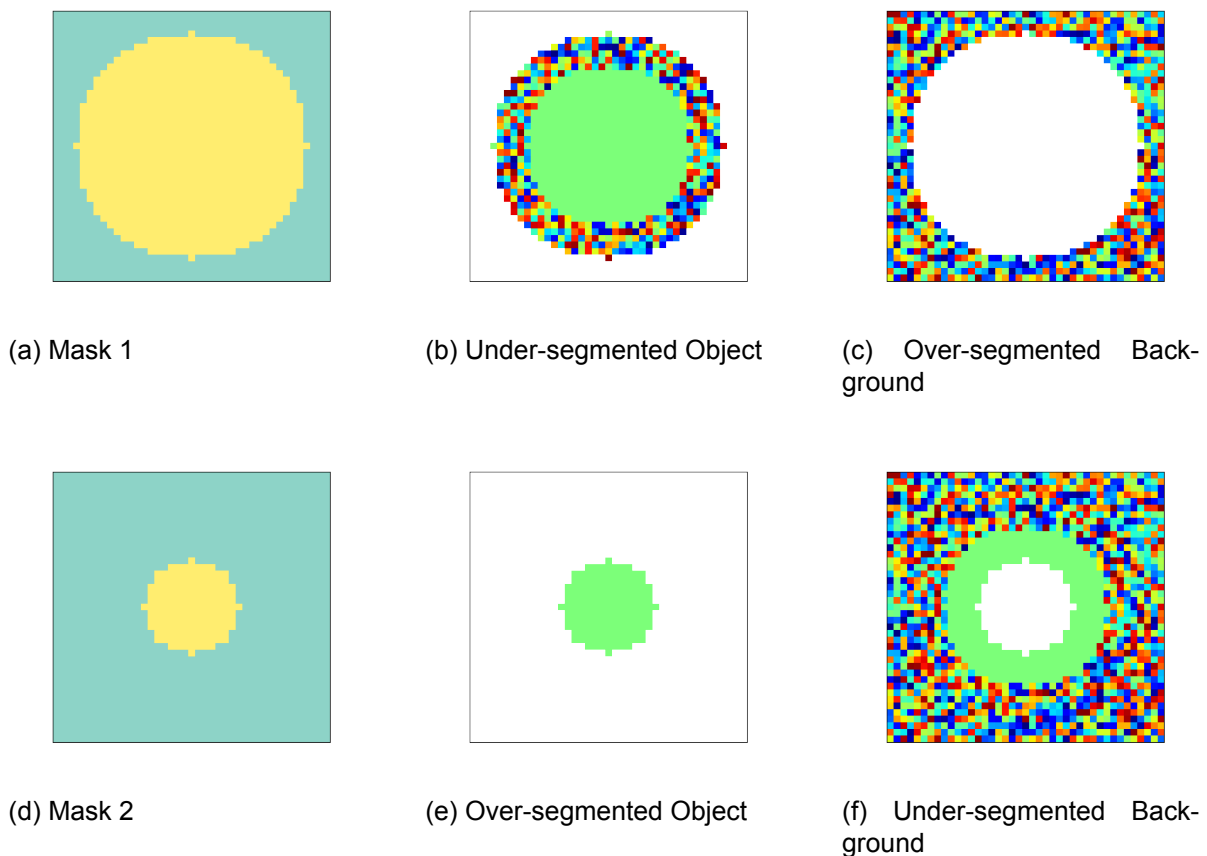
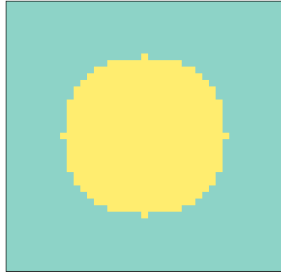


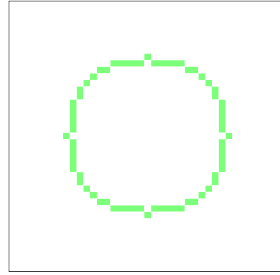
Figure 5.21: Example results of under- and over-segmentation by their label masks and segmented objects.

As can be seen (Fig.5.21), the areas of mis-classified parts are smaller than the segmented regions. If the mix-classified parts are much smaller, those mis-classified parts will be averaged by the whole segmented region. This averaged effect makes it harder to assess the segmented results either for human perceptions or for computer.

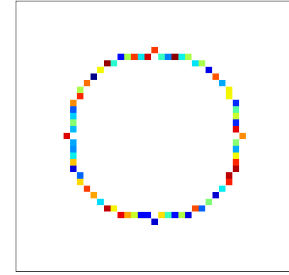
To avoid this averaged effect, one solution is to focus on the boundary. As shown by Fig.5.22, only the correctly segmentation results have an obvious difference between the inner and outer-line profiles. In other words, when the inner and outer-line profiles are the same, it refers to that the boundary is in the same region. This method will be applied in evaluation (see section 5.5.2.2).



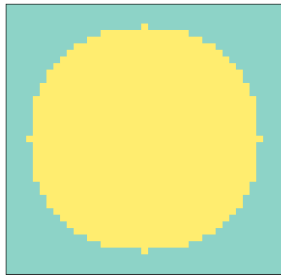
(a) The correct mask



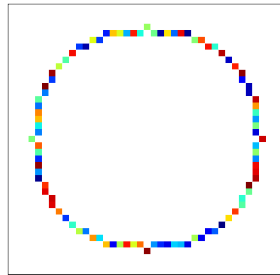
(b) The correct inner-line profile



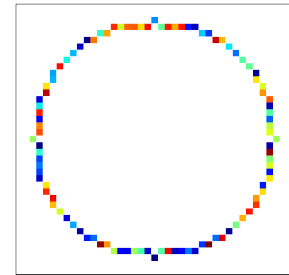
(c) The correct outer-line profile



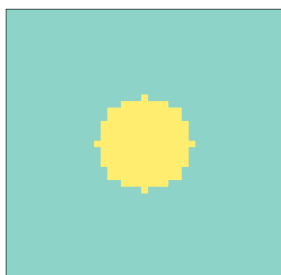
(d) Mask 1



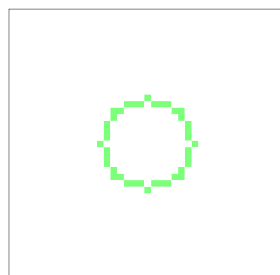
(e) Mask 1 inner-line profile



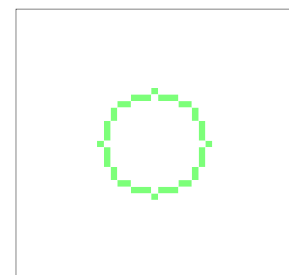
(f) Mask 1 outer-line profile



(g) Mask 2



(h) Mask 2 inner-line profile



(i) Mask 2 outer-line profile

Figure 5.22: Three segmentation results masks and their inner-line and outer-line profiles

5.5. Evaluation

Evaluation is to assess the performance of any algorithm based on the criteria defined in the section. To assess the filtering technique, equivalently the segmented spectrogram, subjective and objective evaluations are used based on the criteria. The difference between them is from whether using human evaluators or not.

5.5.1. Subjective evaluation

Subjective methods are the most widely and easiest to do the assessment, but its application is limited. It is the way that judging the segmented results by human evaluators, for example, looking at the results by human vision directly. Also, related expertise in weather radar polarimetric processing is needed to confirm a correct judgement. Due to the necessary human intervention, this method is time-consuming when the data-set is large. And it is not feasible to be used in a real-time assessment.

5.5.2. Objective evaluation

Objective evaluation methods are those without human evaluators ([10]). From previous studies ([5], [4]), the mainly used objective evaluation method is the supervised evaluation, which compares the segmented results with a given ground-truth data. While in the study cases of this project, no spectrogram with true precipitation pixels is given. And manually generating those reference data is impossible. Unsupervised methods are accordingly necessary to develop. Instead of relying on the reference, those methods evaluate a segmented image by a set of characteristics as desired by humans ([10]). Several unsupervised methods matching the criteria in section 5.4 are introduced in the following parts. In the following sections, each objective evaluation method will be matched with each criterion. S and T refers to Spectral and Time domain respectively, and the number behind denotes the order. For example, S-1 means the No.1 criterion in Spectral domain.

5.5.2.1. Uniformity measurement (Criterion S-1)

According to Criterion 1, the segmented areas should be uniform with their own characteristics such as the intensity distribution of spectral polarimetric parameters. More uniform within an object is equal to less contaminated inside. The entropy-based uniformity measurement was proposed by [32], as it has also validated a good performance in [10]. The region Shannon's entropy (H_v) is used to assess the uniformity within each region and defined:

$$H_v(R, X) = \sum_{i=1}^n P(x_i) \log \frac{1}{P(x_i)} \quad (5.22)$$

where $H_v(R, X)$ denotes the region entropy on region R and X represents the assessed spectral polarimetric parameter. n is the number of probability mass function bins $P(X)$.

As there are two regions, hydrometeor and non-hydrometeors, the expected region entropy H_r ([32]) is defined by each region entropy H_v multiplying the weight proportional to the pixel numbers of that area S_j :

$$H_r = \sum_{j=1}^m \left(\frac{S_j}{S_I} \right) H_v(R_j) \quad (5.23)$$

$$R_j \in \{R_C, R_H\} \quad (5.24)$$

$$S_j \in \{S_C, S_H\} \quad (5.25)$$

$$S_I = N_{range} \times N_{Doppler} \quad (5.26)$$

where R represents the label of the segmented regions and S is the area or pixel numbers of the region. The subscripts C and H of R_j and S_j denote non-hydrometeors and hydrometeor, respectively. S_I is the total area of the whole Range-Doppler spectrogram which corresponds to the total pixel number.

With the properties of 2 and 3 of Shannon's Entropy, a low entropy value corresponds to a limited number of states in the segmented region. Consequently, the region shows uniformity for the spectral polarimetric variable considered. A larger entropy value refers to partial regions contaminated with pixels from the other one. As concluded from section 5.2.1, $s\rho_{co}$ has the most sensitive reaction on distribution variations when there is a rainy object in the spectrogram. Therefore, $s\rho_{co}$ is the spectral parameter to be chosen for the uniformity measurement.

However, uniformity measurement by the expected region entropy has also limitations. As lower entropy results are considered better, this assessment may prefer over-segmented objects. Besides, the areas of non-hydrometeors objects are much larger than the rain objects in some light rain spectrograms, so the regional entropy from non-hydrometeors might dominant the expected region entropy results that shades the impact from the hydrometeors part.

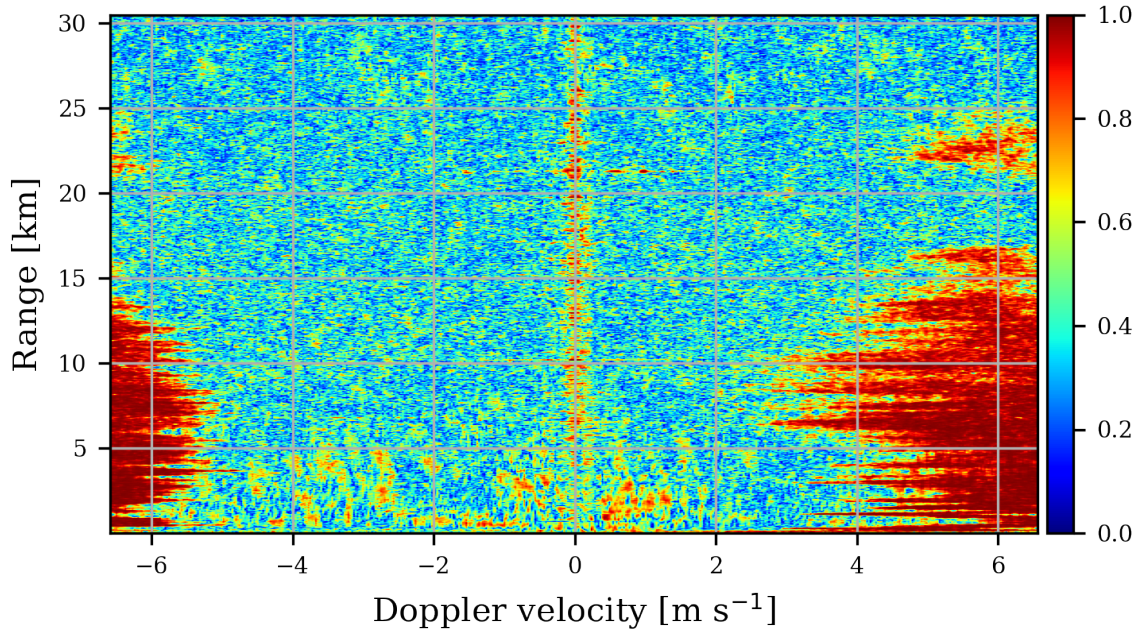


Figure 5.23: The example spectrogram to be segmented

Figures 5.23 - 5.26 explain the connections between segmented spectrogram, distribution, region entropy (H_v) and the expected region entropy (H_r) by giving two segmented results. Fig.5.23 is the spectrogram used to segment. Two label masks (Fig.5.24) are resulted from two different methods.

Fig.5.25 provides the histograms and spectrograms of extracted hydrometeors and non-hydrometeors by mask A. For the extracted hydrometeors by mask A (Fig.5.25 (b)), it contains part of non-hydrometeor pixels which has lower $s\rho_{co}$ magnitudes or some ground clutter pixels around 0 m/s. Thus, its $s\rho_{co}$ distribution has a wide left tail and larger $H_v(R_H, s\rho_{co})$.

In comparison, the segmented results by mask A have higher contamination level than that of mask B. For segmented hydrometeor region, the extracted hydrometeor by mask A (Fig.5.25 (b)) includes pixels with low $s\rho_{co}$ values along the boundary of hydrometeor regions and around 0 m/s, which account for the left tail distribution of the histogram of hydrometeor (Fig.5.25 [d]). And the segmented result by mask A is purer without any contamination of hydrometeor pixels. However, in the case of mask B (Fig.5.26), the segmented hydrometeor regions are without low $s\rho_{co}$ so its histogram is more engaged in the high $s\rho_{co}$ value part. In the segmented non-hydrometeor region, there are part of pixels with higher $s\rho_{co}$ values, which are potentially hydrometeors. Therefore, the segmented hydrometeor by mask A suffers under-segmented and its non-hydrometeor is over-segmented. For the result of mask B, it has the contrary condition that it has an over-segmented hydrometeor and under-segmented non-hydrometeor regions. Their contamination levels are measured by H_r , and their H_r results agrees the observed results above.

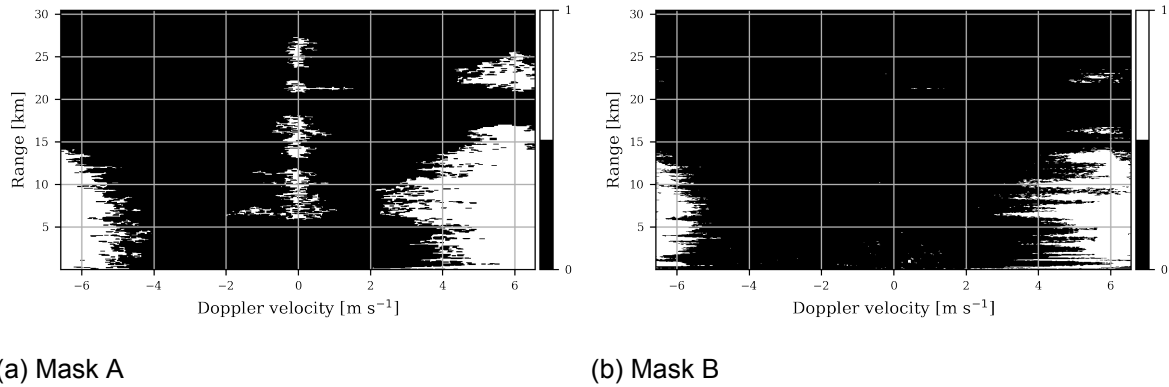
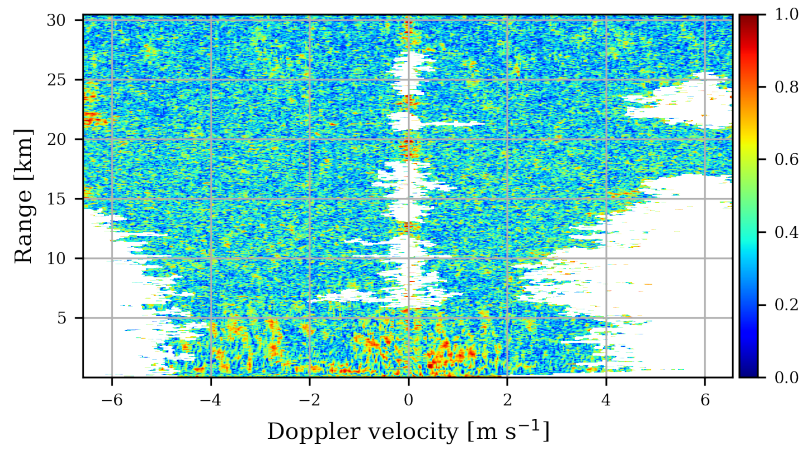
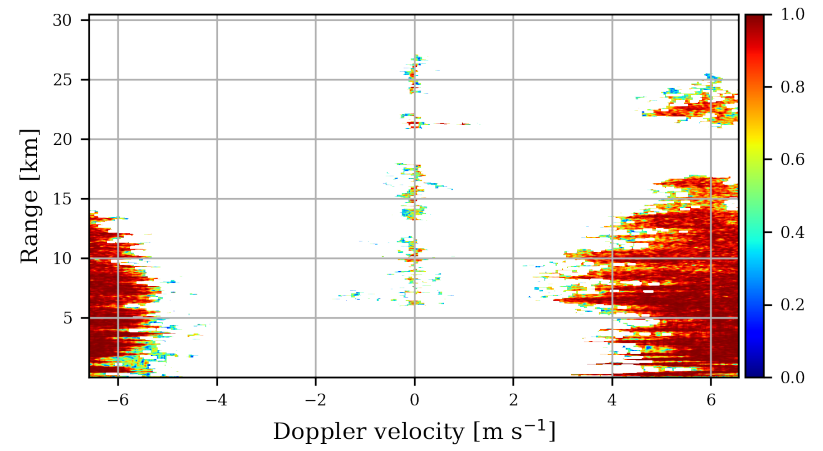


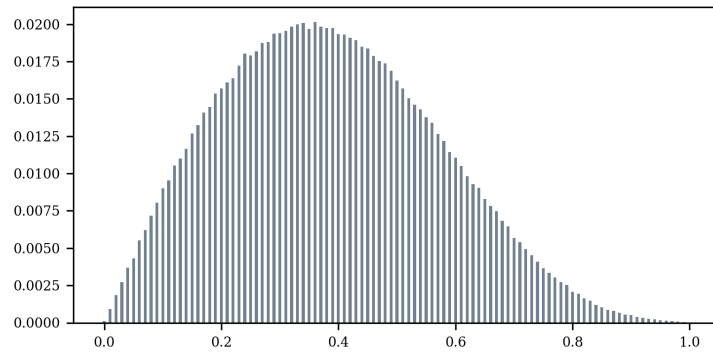
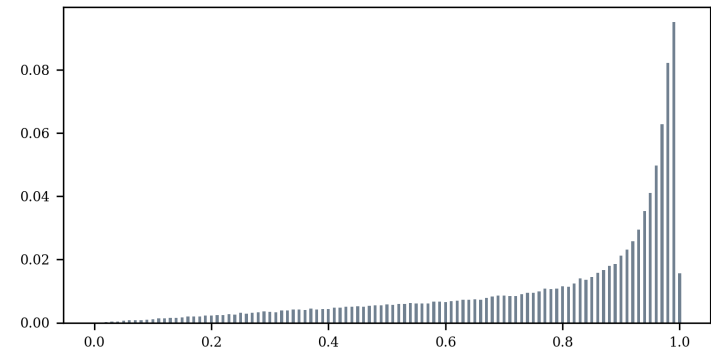
Figure 5.24: Two masks labelling hydrometeors and non-hydrometeors for the spectrogram of Fig.5.23

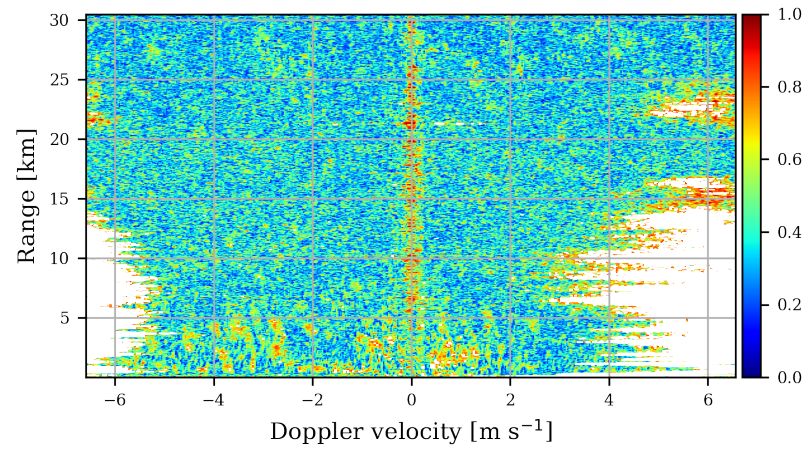


(a) Extracted non-hydrometeor by mask A

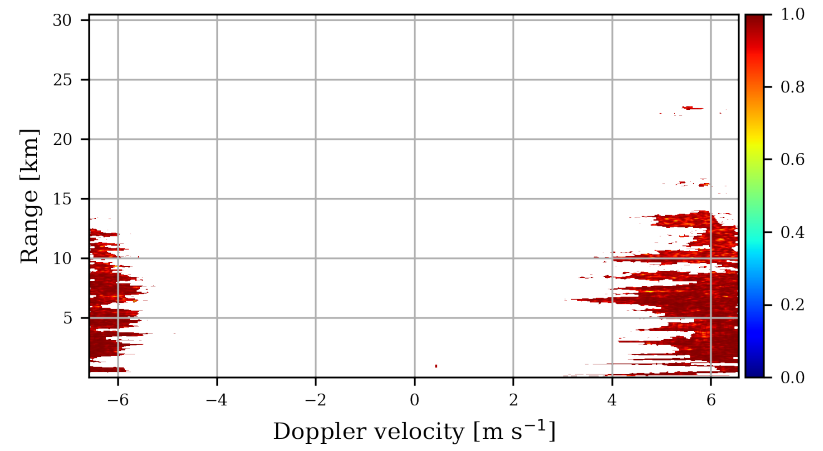


(b) Extracted hydrometeor by mask A

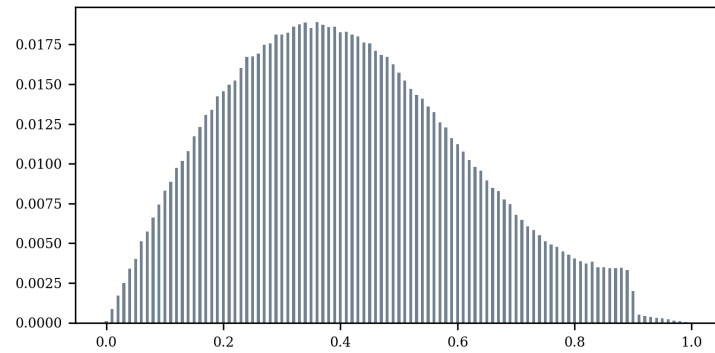
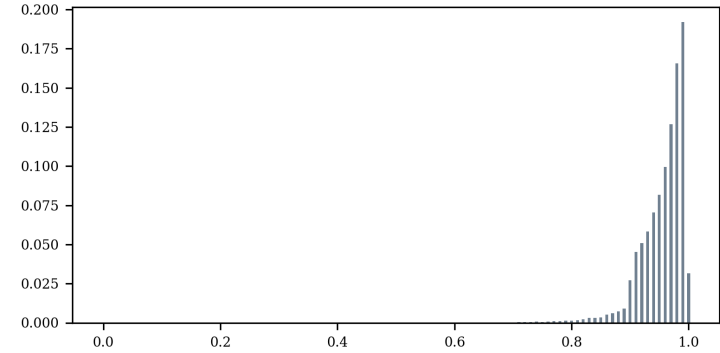
(c) $s\rho_{co}$ distribution of non-hydrometeor ($H_v(R_C, s\rho_{co}) = 6.192$)(d) $s\rho_{co}$ distribution of hydrometeor ($H_v(R_H, s\rho_{co}) = 5.727$)Figure 5.25: Segmented spectrogram by mask A with $H_r = 6.108$



(a) Extracted non-hydrometeor by mask B



(b) Extracted hydrometeor by mask B

(c) $s\rho_{co}$ distribution of non-hydrometeor ($H_v(R_C, s\rho_{co}) = 6.298$)(d) $s\rho_{co}$ distribution of hydrometeor ($H_v(R_H, s\rho_{co}) = 3.539$)Figure 5.26: Segmented spectrogram by mask B with $H_r = 6.051$

5.5.2.2. Boundary Contrast (Criterion S-2)

Concerning the drawbacks of the expected region entropy and the 2nd criterion, a local assessment based on Shannon's entropy was derived. The first step is to find the spectrogram which has a sharp boundary between the rain objects and non-hydrometeors. The boundary to distinguish rain and non-hydrometeors is due to a sharp change of smoothness – the rain object is smooth while the non-hydrometeor area at the rain boundaries is generally noisy. According to the spectrogram observation results, sZ_{DR} is the suitable polarimetric spectrogram. Remember the interpretation of Shannon's entropy for measuring the number of states, so the noisy part will have a larger entropy.

However, that boundary is something novel as its location is inter-pixel, so it is impossible to be measured. The solution here is to calculate the entropy gradient through the boundary between the two objects. A high entropy gradient corresponds to a sudden change. Concerning the rain object, we define the inner line as the pixel series with one-unit pixel width along the boundary where located within rain objects. The outer line is the pixel series on the non-hydrometeor object side (see illustrative examples in section 5.4.3.1). . And the expected entropy of each line is defined:

$$H_v(L_j, sZ_{DR}) = \sum_{i=1}^n P(x_i) \log \frac{1}{P(x_i)} \quad (5.27)$$

$$L_j \in \{L_{in}, L_{out}\} \quad (5.28)$$

where L is the pixel number of each line. The entropy gradient is the difference between.

$$H_d = H_v(L_{out}) - H_v(L_{in}) \quad (5.29)$$

There are two cases when the entropy gradient is lower, namely, over-segmented and under-segmented. When the boundary locates whether, within the non-hydrometeors or the rain object, the entropy difference through the segmented boundary is small so lower entropy gradient.

Fig.5.27 shows the schematic of boundary contrast in a down scale way: the upper-left is a 9×5 image which is part of the sZ_{DR} on the left. The lower-left is the label mask of the upper-left image including labelling its inner-line and outer-line regions. We can see those pixels from outer-line region have larger variations while the inner-line pixels are more uniform (smoother).

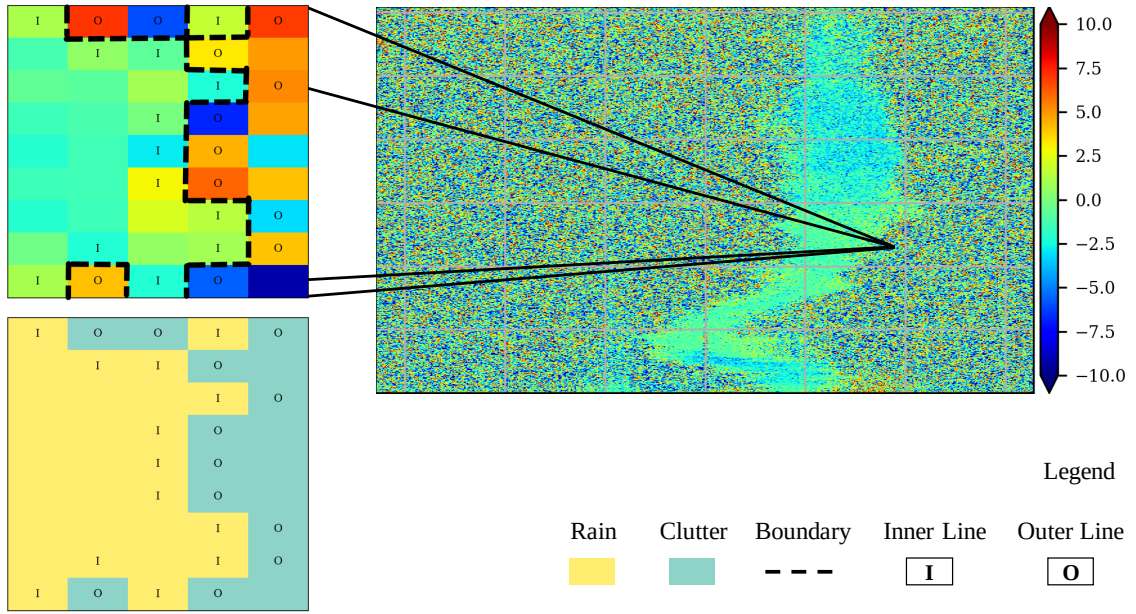


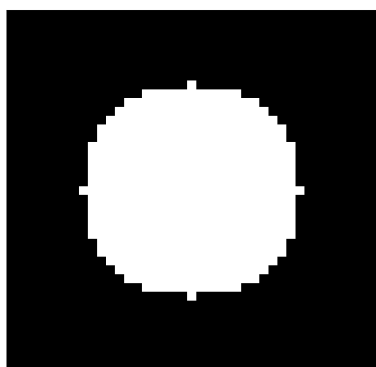
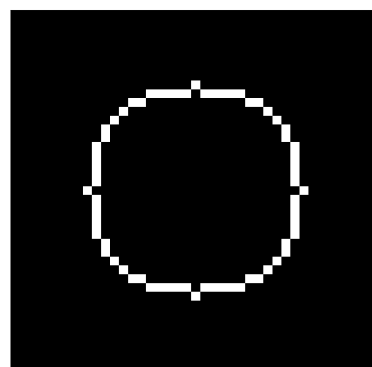
Figure 5.27: Schematic of Boundary Contrast

5.5.2.3. Connectivity measurement (Criterion S-3)

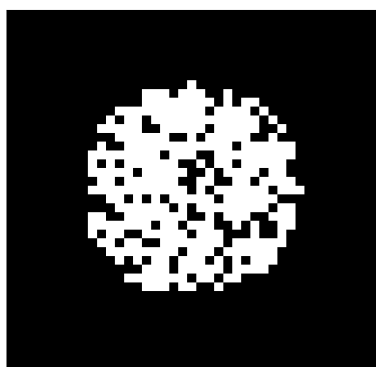
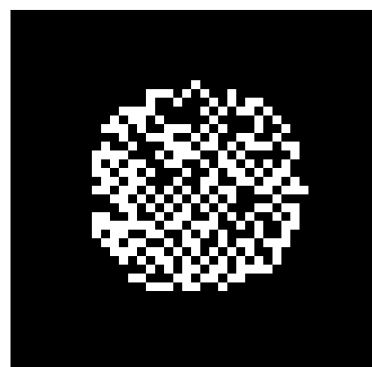
Connectivity is defined by the ratio between the inner-line pixel numbers and the total pixel numbers of hydrometeor objects. The inner-line pixel is the pixel which if any of its neighbours has a different label. If the ratio is closer to 1, it implies the hydrometeor objects losing their spatial connectivity – more labelled hydrometeor pixels are isolated with neighbouring pixels. Intuitively, low connectivity label masks are granular while the high connectivity filter is a combination of blocks.

$$C = \frac{S_L}{S_H} \quad (5.30)$$

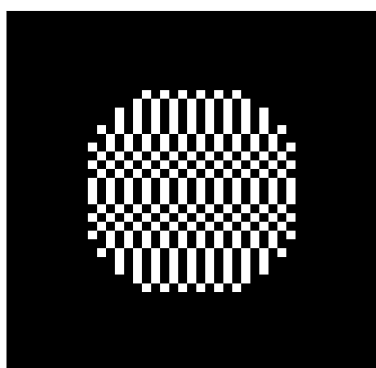
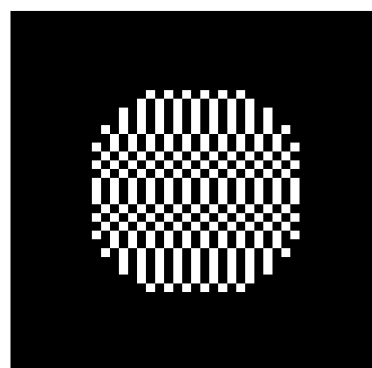
Fig.5.28 shows three masks with different connectivity and their own inner-line objects. All white or black pixels are connected in the first label mask (a) without any hole inside, so its C is lowest. In the second mask (b), there are some holes within the object so its C is larger. In the third mask, when it has more holes inside, the inner-line object is the same as its own. Its C reaches its maximum, 1. Therefore, the value of C is ranged from 0 to 1. When C is closer to 1, the label mask is more intend to be granular like Mask 3 of Fig. 5.28.

(a) Mask 1 ($c = 0.146$)

(b) Inner-line of Mask 1

(c) Mask 2 ($c = 0.703$)

(d) Inner-line of Mask 2

(e) Mask 3 ($c = 1$)

(f) Inner-line of Mask 3

Figure 5.28: Schematic of connectivity measurement

5.5.2.4. Removal percentage (Criterion T-1)

A removal percentage (R_{rm}) is defined as the ratio between pixel numbers retained and the total range numbers. It is the key factor to non-hydrometeor filtering assessment. According to the first criterion in time domain, a high removal percentage suggests a strong non-hydrometeor removing ability, since all the signals in the clear air are non-hydrometeors.

$$R_{rm} = \frac{N_e}{N_{range}} \quad (5.31)$$

where N_e denotes the number of echo bins in a radar ray and N_{range} is the total number of radar resolution volumes in the whole range.

5.5.2.5. Weak and strong echoes ratio (Criterion T-2)

The ratio indicating the amount of weak and strong echoes in the whole radar beam is defined,

$$R_{ws} = \frac{N_{ws}}{N_{range}} \quad (5.32)$$

where N_{ws} denotes the number of echo bins whose Z_{HH} is lower than 0 dBZ or larger than 70 dBZ.

5.5.2.6. Single unconnected echo bins ratio (Criterion T-3)

According to the third criterion, more numbers of those single unconnected echo bins are considered as bad filtering results. Here, we define the single unconnected echo bins ratio (R_{su}):

$$R_{su} = \frac{N_{su}}{N_{range}} \quad (5.33)$$

where N_{su} denotes the number of single unconnected pixels.

5.6. Compare and assess different filtering methods

Subjective and objective methods were both used to assess different filtering techniques performance. Different scenarios, namely, clear air echo, light rainy, and heavy rainy cases introduced in chapter 4 were tested. For the evaluation result of each case and filter, one full PPI consisting of 196 rays is considered. The subjective method consists of plotting and visually assessing the filtered PPI. Objective methods include those unsupervised methods defined before in section 5.5. The use of assessing variables should depend on the test scenarios. For example, considering clear air, the removal percentage is the appropriate assessing variable to test the filtering performance.

Because of the lack of ground truth data, objective evaluation variables in spectral domain (H_r , H_d , C) can only give a relative assessment. The optimised results of these three spectral domain assessment variables can only be obtained in the ground truth. However, all time-domain evaluation (R_{rm} , R_{ws} , R_{su}) can show the exact measurement as we expect the value of 0 is the best.

One PPI has 196 rays, so every study case will have 196 test results by each filter for all assessment variables. The distribution of the results will be analysed. The distribution of each assessment variable will be visualised by the boxplot. When analyzing the boxplot, the median, interquartile range, potential outliers are considered. The median (Q_2 or 50th percentile) location can be equivalent as the average level of one evaluation. Interquartile range (IQR) is defined as:

$$IQR = Q_3 - Q_1 \quad (5.34)$$

where Q_3 and Q_1 denote 75th percentile and 25th percentile. Outliers (q_{out}) are those values which are in the condition:

$$q_{out} \in \{q | q > Q_3 + 1.5 \times IQR\} \quad (5.35)$$

$$q_{out} \in \{q | q < Q_1 - 1.5 \times IQR\} \quad (5.36)$$

where q denotes any value of one assessment variable among the 196.

Both IQR and number of potential outliers characterise the performance of the different filtering techniques.

For further assessing the techniques, the complexity of the filter algorithm should also be considered. The complexity may directly influence the applicability and fault tolerance. A filter with poor fault tolerance means this filter could perform well in some known input data but poor in some unknown tested scenarios, which corresponds to *Over-fitting* in machine learning. The parameters especially those empirical parameters might lead to *Over-fitting* problems. To assess the complexity of the three filters, their key parameters from each step are listed.

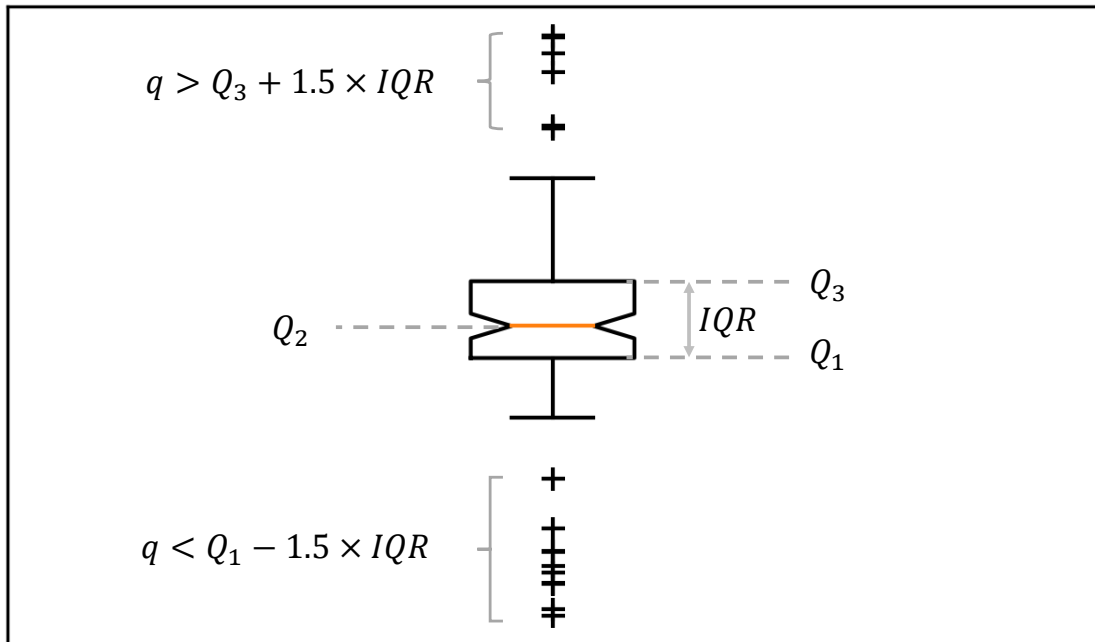


Figure 5.29: Example boxplot showing Q_1 , Q_2 , Q_3 , IQR and q_{out}

6

Results

This chapter discusses the results of the three study cases introduced in Chapter 4, which are obtained from three different filtering techniques, namely, the Default Rijnmond filter, JSD-based and OBSpol filter. For the filtered results, subjective (Section 5.5.1) and objective methods (5.5.2) are both used to assess the three filtering techniques' performances according to the criteria defined at Section 5.4. Displaying and then discussing the filtered PPI is the subjective evaluation. PPIs filtered by the three filters and the corresponding raw PPI are plotted for comparison. All the metrics defined in Section 5.5.2 are applied for objective evaluation.

6.1. Subjective evaluation results

This section will present and discuss the PPI results from the three filtering techniques. The subjective evaluation method is used according to the criteria for time-domain radar profiles in section 5.4. Besides, some radar rays suffer from saturation so the ability of removing those saturation radar rays should also be considered.

6.1.1. Clear air case

Fig.6.1 shows the raw and filtered PPIs of the clear air study case. For filters performance in clear air case, we only consider the amount of radar pixels remaining. The filtered result by JSD-based method has the best performance, as all radar pixels are removed. Then the default Rijnmond radar filter has relatively good results. However, it cannot remove those radar rays with saturation problems, whose whole radar rays have incredible strong Z_{HH} . And it keeps some weak echoes like those radar rays from north to east. The OBSpol filter keeps most of the radar pixels especially those radar rays having non-hydrometeors echoes with high Z_{HH} . It also failed to remove those radar rays with saturation issues.

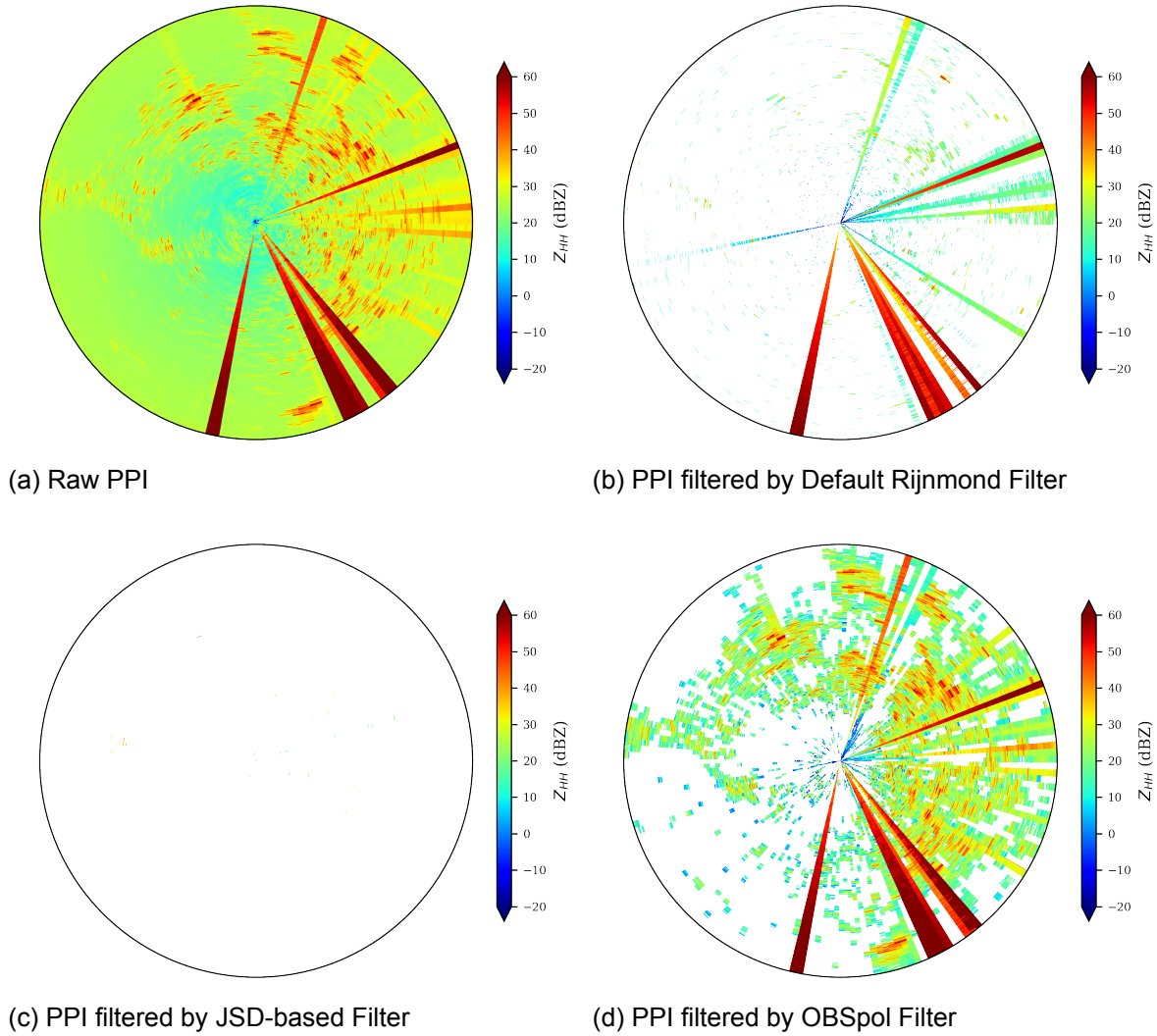


Figure 6.1: PPIs of the clear air case

6.1.2. The rainy cases

Fig.6.2 and Fig.6.3 show the raw data and filtered results in the two rainy cases. The former is for the light rain; the latter is the heavy rain. Overall, filter OBSpol retains the most radar bins including some weak echoes and ground clutter. The default filter can suppress most of the ground clutter but there are residual non-hydrometeor echo bins in far ranges left. Filter JSD can remove the ground clutter and leads to less residual echo bins. Besides, filter JSD is the only one succeeding in filtering saturation radar rays in the light rainy case.

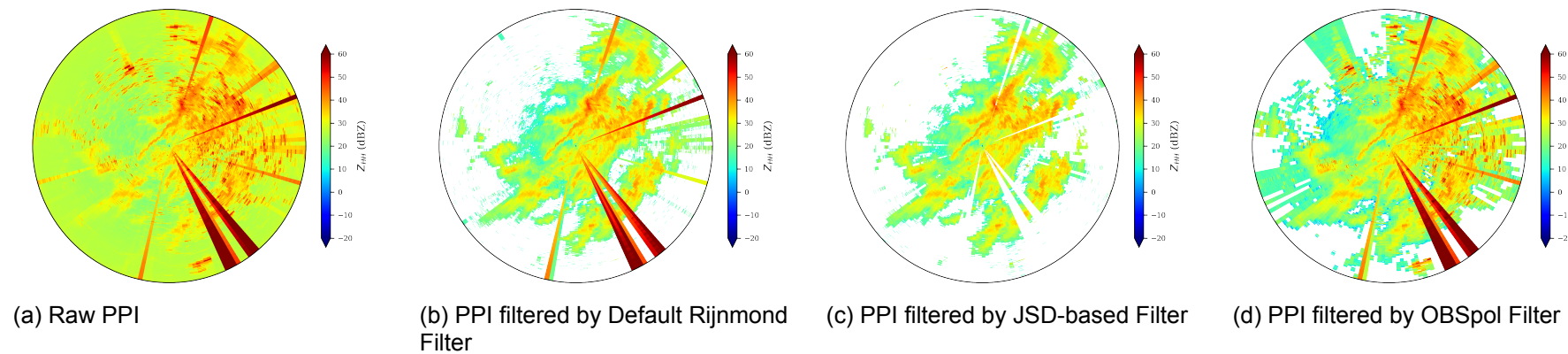


Figure 6.2: PPIs of the light rain case

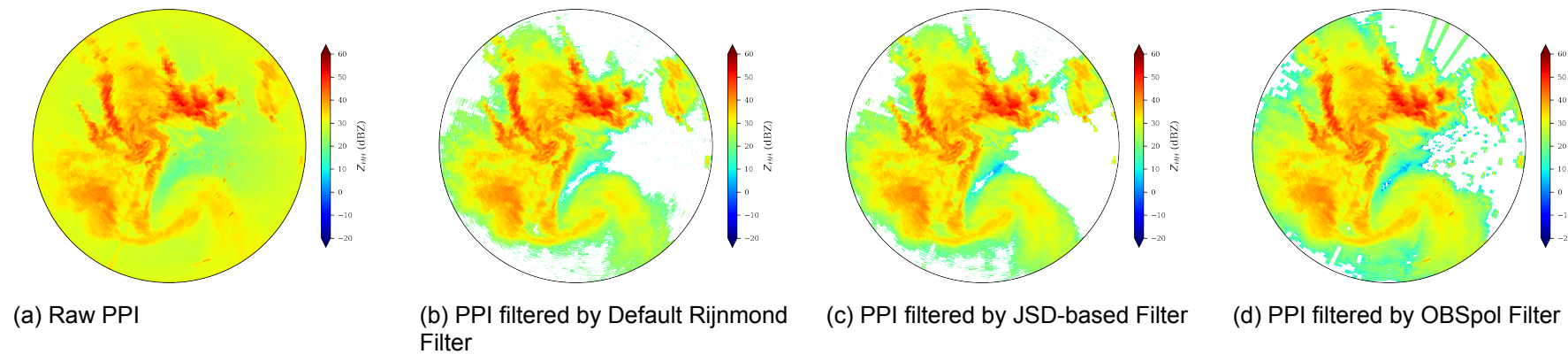


Figure 6.3: PPIs of the heavy rain case

6.2. Objective Evaluation

6.2.1. The clear air case evaluation

For clear air case evaluation, we only consider the removal percentage (R_{rm}) in time domain. Fig.6.4 shows the removal percentage from three filters. For the box median, Filter JSD and Filter Default locate close to 100%, which shows these two filters have good performance in filtering clear air echoes compared to filter OBSpol. However, filter Default has larger IQR and more outliers. Therefore, filter JSD works the best in filtering clear air echoes.

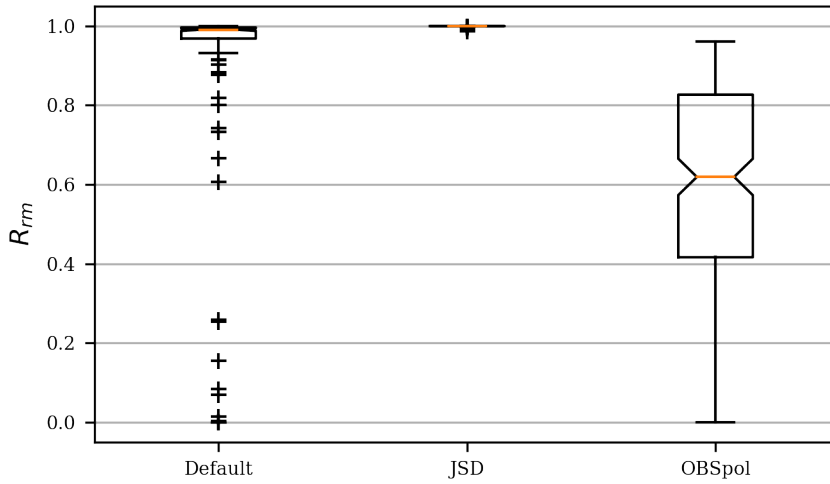


Figure 6.4: Removal percentage results of the three filtering techniques

6.2.2. The two rainy cases evaluation

For the two rainy cases evaluation, all evaluation methods in spectral and time domain will be applied except for the removal percentage.

6.2.2.1. Spectral domain evaluation results

Fig.6.5 shows the expected regional entropy (H_r) distribution of three filters. Filter JSD has the lowest median, while Filter Default has the largest expected regional entropy in both rainy cases. According to section 5.5.2.1, a lower H_r value suggests less contamination either in the hydrometeor part or in the non-hydrometeor region. Therefore, filter JSD works better in regional uniformity than the other two filters.

The magnitude of H_r results from the three filters are comparable, though their relative levels are obvious. This is due to the H_r is dominated by $H_v(R_C)$ of each spectrogram. See the example of segmented results by the two masks in Fig.5.25 and Fig.5.26. The $H_v(R_H)$ of mask A is much larger than the $H_v(R_H)$ of mask B, as the segmented hydrometeor region by mask A is more contaminated than that by mask B. Their own non-hydrometeor regions also show different level of uniformity, while their $H_v(R_H)$ are comparable.

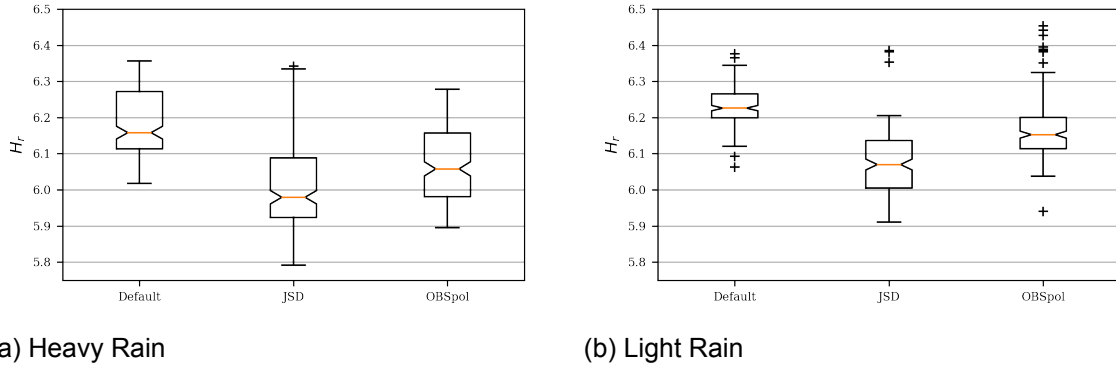


Figure 6.5: The expected region entropy H_r for the two rainy cases

Comparing the results between the heavy and the light rain, Filter Default and Filter JSD have more results with H_r values. On the contrary, the entire box of Filter OBSpol was shifted lower that is its H_r results distribution locates more in the lower H_r region. This is the impact of notch filter implementation. In most of the heavy rain spectrograms, the hydrometeors have a significant number of pixels around 0 m/s. After applying the notch filter filtering, which filters ground clutter around 0 m/s, hydrometeors corresponding to the Doppler velocity interval $[-v_G, +v_G]$ are mis-classified as non-hydrometeors. Therefore, these mis-classified non-hydrometeor pixels increase the $H_v(R_C)$, which leads to larger H_r values as well. Fig.6.6 shows an example of the mis-classification by notch filter, at range from 2.5 to 7 km, part of the hydrometeors was classified as non-hydrometeors.

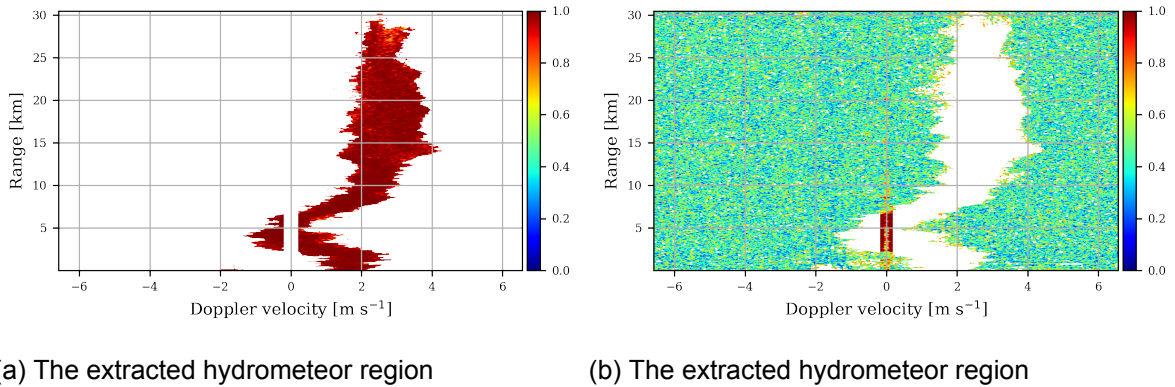
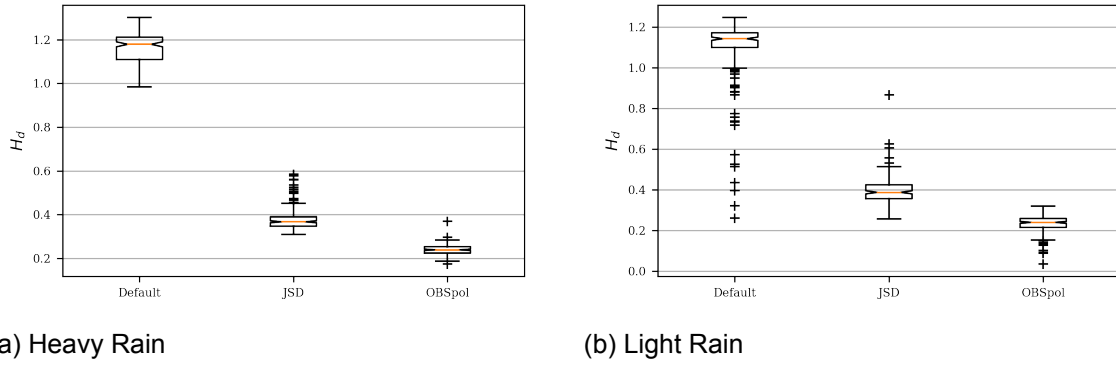
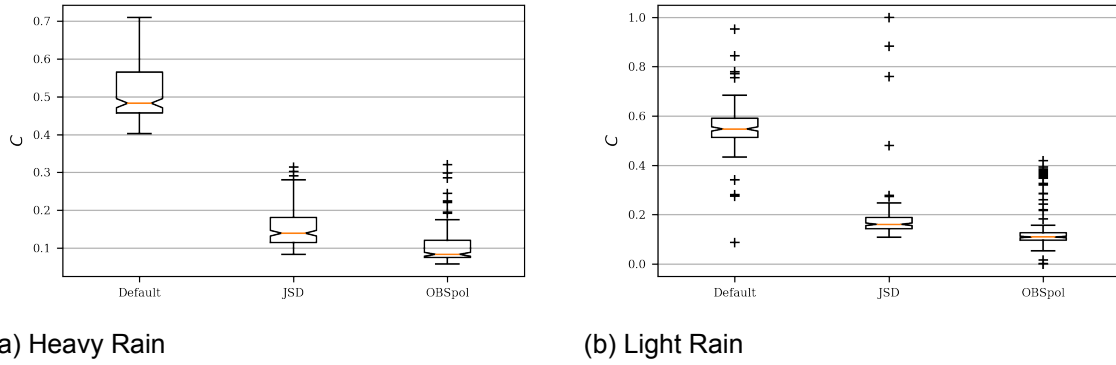


Figure 6.6: Example of mis-classification due to notch filter

For the boundary contrast (H_d) and spectral connectivity (C) assessment (see Fig.6.7 and Fig.6.8), Filter Default takes the highest values and Filter OBSpol has the lowest. Filter JSD performs neutrally. As suggested by section 5.5.2, higher H_d reveals the boundary position between hydrometeors and non-hydrometeors might be more spatially correct. And higher C values refer to poorer connectivity.

Figure 6.7: The boundary contrast measurement H_d for the two rainy casesFigure 6.8: Spectral connectivity measurement C for the two rainy cases

6.2.2.2. Time domain evaluation results

In the time domain evaluation, R_{ws} has no much difference between these three filtering techniques. Only several numbers of radar rays have high R_{ws} in Filter Default and Filter OBSpol.

Fig.6.9 shows that the weak and strong echoes ratio, R_{ws} , are similar for the three filters. And every filter has a certain value of R_{ws} in the whole PPI, which is mainly due to those weak echoes in near range. Only several rays have high R_{ws} values using Filter Default and Filter OBSpol especially in the light rain case. Therefore, the number of outliers (markers with + sine) of each box is assessed. Filter JSD has the least number of outliers in both rainy cases. Filter OBSpol has the most outliers in the heavy rain case, while in the light rain case, Filter Default and Filter OBSpol have comparable numbers of outliers.

The performance differences become larger for the time domain connectivity, R_{su} , in Fig.6.10. Filter Default has more single unconnected echo bins, but Filter OBSpol works the best in retaining this. R_{su} is highly correlated to the spectral domain connectivity measurement, C .

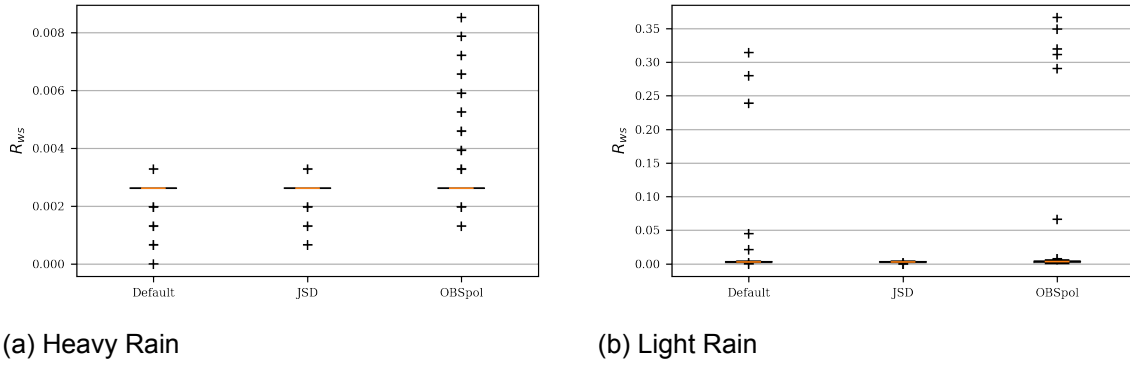


Figure 6.9: Weak and strong echoes ratio assessment results for the two rainy cases

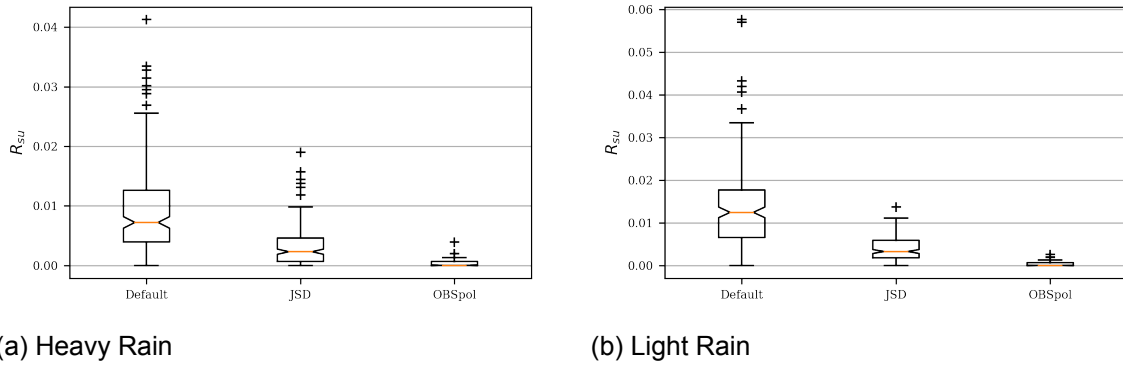


Figure 6.10: Time domain connectivity assessment for the two rainy cases

6.3. Filter performance: summary and discussions

In this section, we summarise and discuss the three filters' performances from the subjective and objective evaluation results given in the above sections.

Table. 6.1 summarises the performances of the three tested filtering techniques, and the number of star signs (*) qualitatively characterizes the performance per evaluation parameter. A larger number of star signs indicates a better performance. Filter JSD has the largest number of star signs. Filter Default and Filter OBSpol have equivalent total numbers of star signs, but these two filters have their own strengths. And the objective evaluation results match the subjective evaluation.

The parameters used from the three filtering techniques are listed in Table.6.2. The number of parameters used from each filter is comparable. Filter Default and OBSpol both use sL_{DR} for the thresholding, while Filter JSD uses sp_{co} . sL_{DR} cannot be retrieved by some transmission and reception mode, for example, STSR, which means sp_{co} is more accessible. Noise clipping is conducted in the Filter Default which will lower the operation speed of the filtering.

Table 6.1: Filter performance summary

	Default	JSD	OBSpol
Clear Air Remove	★ ★	★ ★ ★	★
Uniformity	★	★ ★ ★	★
Boundary	★ ★ ★	★ ★	★
Connectivity (Spectral)	★	★ ★	★ ★ ★
WS Echo Remove	★ ★	★ ★ ★	★
Connectivity (Time)	★	★ ★	★ ★ ★
Saturation Remove	/	★	/

Filter Default performs relatively well in removing clear air and weak-strong echoes. It can filter ground clutter in the rainy cases by its notch filter. However, it fails to remove the saturation radar rays and it retains residual non-hydrometeor pixels either in the clear air case or in some far ranges of the rainy cases. According to its connectivity evaluation results (relatively high C), it may lose some precipitation pixels that could underestimate the rainfall rate. According to section 5.5.2, a high C mask has a granular pattern so its inner-line object also includes the part inside of the region (see Fig.5.28 b and c). This means the H_d results do not show the result of the boundary between the hydrometeor and the non-hydrometeor. Therefore, its high H_d result cannot be considered as it has the most spatially correct boundary. The label mask with high C created by Filter Default is determined by its algorithm side. The two core steps of Filter Default, the double sL_{DR} thresholding and noise clipping, both generate label masks with granular pattern.

Filter OBSpol performs the best in retaining continuity of hydrometeor echoes areas both in spectral and time domain (low C and high R_{su}) but it fails to filter most signals in clear air case. This performance property is resulted from its step 2 and step 3. These two steps are beneficial to the connectivity performance. However, these two filtering steps performance highly depend on its serious prerequisites (see Table.6.2). For step 2, the *Closing* operation, an improper structural element may result in mis-classification; for example, a too large size structural element classifies more non-hydrometeors to the hydrometeors. The boundary contrast evaluation result of Filter OBSpol, which it has the lowest overall H_d values, may suggest that the *Closing* operation expands the hydrometeors region to the non-hydrometeor part. Besides, when the spectral polarimetric thresholding fails to remove some clutter, like the ground clutter, the *Closing* operation will expand the areas of the clutter. Therefore, it shows high R_{rm} in the clear air and H_r results in the rainy cases.

Filter OBSpol performs well in the case when the ground clutter and hydrometeors are overlapped, since it is without applying the notch filter. When the ground clutter and hydrometeors are overlapped, the use of notch filter will mis-classify the part of hydrometeors at around 0 m/s (see section 6.2.2.1). Filter OBSpol retains that part of hydrometeors, though it will suffer contamination of ground clutter when most of the hydrometeors have large Doppler velocity like the light rain case.

Filter JSD performs better than Filter Default and Filter OBSpol both in clear air and rainy cases. In the clear air case, Filter JSD removes the most echoes. We can infer that Filter JSD is strongly capable to remove the non-hydrometeors. In the rainy cases, it filters most of the non-hydrometeor signals including ground clutter

with less residuals. What is the most remarkable is that Filter JSD is the only one succeeding removing the saturation radar rays. This is because Filter JSD is the only filter determining the spectral polarimetric threshold value ray by ray, while the others apply a fixed threshold value for all rays of the PPIs. Besides, Filter JSD has taken the advantages of the other two filters. It applies the notch filter from Filter Default to remove ground clutter. And the mathematical morphology processing of Filter OBSpol is also used to keep the continuity of the precipitation. However, unlikely to the *Closing* used in Filter OBSpol, we only apply *Dilation* to fill the holes inside the precipitation part excluding processing on the boundary.

The connectivity (C) and uniformity (H_r) in the heavy rain case suggest what Filter JSD should be refined. Its connectivity is due to the impact of some residual narrow Doppler spectra. For example, as can be seen in the mask of Fig.5.15, there are still some white particles between -2 to 2 m/s of the near ranges. This implies the restriction of the step 4 of Filter JSD that it is hard to remove narrow Doppler spectra locating the same range as hydrometeors. And as concluded from section 6.2.2.1, a fixed width of notch filter will result in mis-classification issues.

Table 6.2: Filters parameters overview

Filter Tech- niques	Rijnmond Default Filter	OBSpol Filter	JSD based Filter
Thresholding condition for hydrometeor	sL_{DR}^{HH} and $sL_{DR}^{VV} < -7\text{dB}$	sL_{DR}^{HH} and $sL_{DR}^{VV} < -7\text{dB}$	$s\rho_{co} > T_{opt}$
Other empirical parameters	Noise clipping Threshold = 3dB Narrow Doppler bins ratio = 2 % Notch filter width = 0.2 m/s	Structal element radius = 3 Object numbers = 8 Non-stationary clutter Doppler width = 42 cm/s	Consecutive numbers for $s\rho_{co}=5$ Narrow Doppler bins ratio = 2 % Notch filter width = 0.2 m/s
Prerequisite	Noise power measurement	/	Clear air echo data base

Conclusions and recommendations

7.1. Conclusions

This research project aims to design methods for extracting hydrometeors in Range-Doppler spectrogram and evaluating the results with the help of image processing. The objectives have been achieved. Therefore, we can conclude that image processing techniques can contribute to non-hydrometeor filtering and developing evaluation methods to assess the filtered results.

The conclusions are discussed in four parts:

1. *Which image processing techniques can be beneficial to filtering development in the context of weather radar?*

Image segmentation is the image processing technique that can make significant contributions to spectral polarimetric filtering development because of a strong connection between them. These two types of techniques consist of partitioning input images into several objects of interests. One image segmentation technique, the histogram-based global thresholding, has already been applied in spectral polarimetric filtering. This method has the benefits of easy implementation and high efficiency, which suits for weather radar real-time processing.

2. *How to make image processing techniques applicable to Range-Doppler spectrograms?*

A framework was proposed to better apply image processing to Range-Doppler spectrogram, since most current image segmentation methods are designed for real-life images, which have different attributes compared with Range-Doppler spectrograms. This framework includes understanding and quantifying the features of polarimetric spectrograms from global, regional and local scales. Based on the feature properties, a new global thresholding method was developed.

3. *How to develop evaluation methods for filtered Range-Doppler spectrograms having no ground truth?*

The first thing to do before developing evaluation methods is to clarify the criteria of good results. As the filtering processing is in spectral domain and radar profile is presented in time domain, criteria of these two domains should be both considered. Criteria in spectral domain are inspired from those for image segmentation

with considering the speciality of polarimetric spectrogram. And those spectral domain criteria match the regional, edged properties of hydrometeor and non-hydrometeor. In time domain, the profile of each radar ray is to be assessed by concerning the magnitude of reflectivity and the continuity of the radar ray. After the criteria have been confirmed, they can be directly used to evaluate filtered results subjectively. However, due to the limitations of subjective evaluation, objective evaluation methods were proposed based on the criteria.

4. *How well the filtering technique, which is enhanced by image processing, performs compared with other filtering methods?*

Three polarimetric filtering techniques were assessed in different cases by the evaluation methods developed of this project. By subjective evaluation, the JSD-based filter can remove all the echoes in the clear air case while the other two retains part of echoes. In the rainy cases, it retains the hydrometeor and with less residual pixels. And it is the only filter succeeding in removing those radar rays with saturation problems. For objective evaluation results, it also has the best quality which corresponds to subjective assessment. From the algorithm side, this new filtering technique is easy to implement.

7.2. Contribution

The contributions from this project can be summarised:

1. Successful design of an effective filtering technique, which has good performances compared with other two current filters.
2. Evaluation methods for filtering assessment without ground truth data are developed and those evaluation methods results are consistent with subjective evaluation.
3. Building of connections between Digital Image Processing, Computer Vision and image Understanding and spectral polarimetric processing. Therefore, more relative studies can be conducted in the future.

7.3. Recommendations

To continue, refine or do further research based on this project, some recommendations are presented:

1. Some key parameters may have significant impact
 - (a) Signal-to-noise ratio (SNR) discussion
SNR plays an important role to characterise Range-Doppler spectrograms. However, this project lacks the discussion for this variable. It is expected to involve SNR this variable in the future.
 - (b) Histogram bin size
Histogram plays an essential role in this research, for example, for the similarity test in section.5.2.1 and any histogram-based segmentation techniques. Bin size is one of the critical factors affecting histogram distribution

patterns. Thus, the histogram bin size of each polarimetric spectrogram will affect the research results.

(c) Window size for sp_{co}

Another factor is the window size for sp_{co} calculation. The variable, sp_{co} , is obtained from 2D convolution of S_{HH} and S_{VV} for which a window size or kernel has to be defined. The spectrogram of the resulting sp_{co} like the related histogram depend on the window size. As can be seen from Fig.7.1 and 7.2, when the window size increases, the histograms of sp_{co} values show a more distinct separation between the hydrometeor and non-hydrometeor distributions. This is beneficial to classify the two different objects by histogram-based segmentation algorithm. Further, a larger window size contributes to filling the interiors within the hydrometeor areas. However, the spectrogram resolution in range and Doppler decreases. A lower resolution spectrogram dismisses small scale atmospheric details. Therefore, the window size is a trade-off between removing non-hydrometeor objects and retaining small scale details.

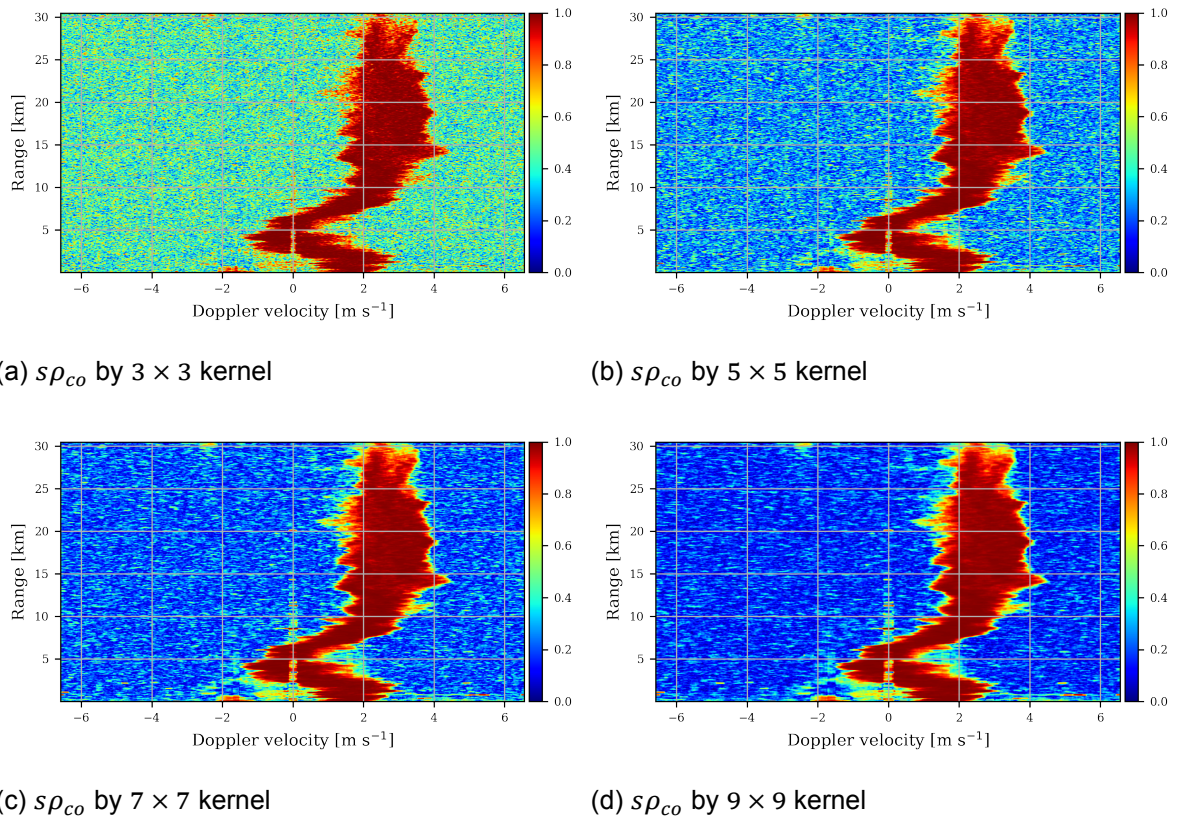
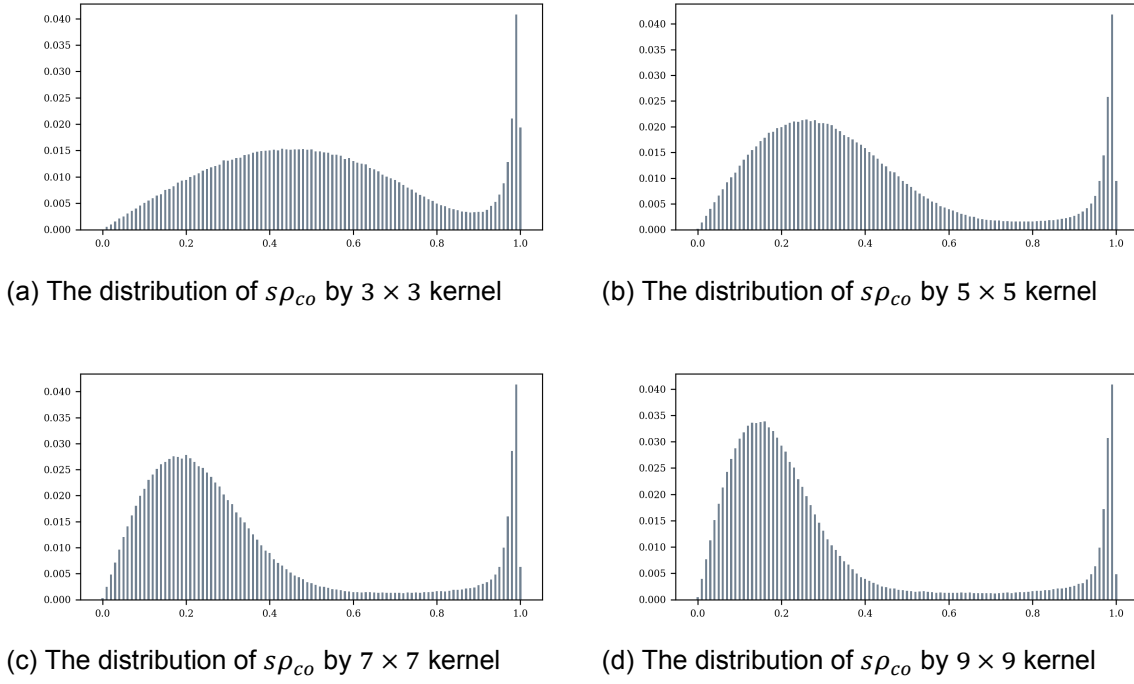


Figure 7.1: Impact of the kernel size to calculate sp_{co}

Figure 7.2: Distributions of $s\rho_{co}$ by different kernel sizes

2. Filter JSD improvement

Though the JSD-based filter performs well, it remains some drawbacks. Discussing those drawbacks gives directions to refine this filtering technique.

(a) Extraction of other hydrometeors, which are not rain

Hydrometeors like wet snow may have values ranging from 0.8 to 0.95 (see Fig.5.7). These $s\rho_{co}$ lower values hydrometeors may challenge every single-valued global thresholding techniques including the JSD-based filter.

(b) Ground clutter and narrow Doppler power spectra removal

Ground clutter and narrow Doppler power spectra removal could also be improved. By default, the notch filter width and the removal percentage of narrow Doppler power spectra are both empirical numbers, which makes the filter less robust. Considering more potentially useful features may help to design algorithms. As we go back to the PPI of the clear air echo and light rainy cases, the ground clutter locates at fixed geolocations. Therefore, the azimuth angle and range are the two factors beneficial for ground clutter suppression. And for narrow Doppler power spectra, those having the same ranges as hydrometeors are hard to remove (see Fig.5.15). Features like pixel areas may contribute to remove them.

3. Evaluation methods validation

The evaluation methods are investigated in chapter 5 without ground truth reference data, the evaluation methods can only assess segmented results relatively between each other. The actual meaning of the evaluation results is not always well understood, in particular the entropy values. Simulation work could help to

gain in understanding. The same for the boundary contrast by inputting spectrogram with known hydrometeor and non-hydrometeor pixels.

4. Investigating more from Computer Vision

As human vision is constantly used during the research of polarimetric processing, Computer Vision has a great impact on this research. The feature understanding part in chapter 5 is inspired from [23]. In Computer Vision, there has already been mature theorem and frameworks for this part. It may contribute to not only understanding and describing image features, but also simulating spectrograms from image synthesising. Investigating those theorem and frameworks will be beneficial to understand spectrogram.

5. More advanced image segmentation methods

Those advanced image segmentation methods with Deep Learning in Chapter 2 should also be considered to apply, since those AI methods have never been used in spectral polarimetric processing. With the help of those advanced methods, more types of hydrometeors and non-hydrometeors would be expected to classify in the spectrogram. For example, all those types of hydrometeors in Fig.5.7 and non-hydrometeors like birds and ground clutter can be identified in the spectrogram. Besides investigating the JSD-based thresholding, the author attempted to apply some more advanced methods such as K-means clustering and super-pixel methods from [7], but no expected results were obtained. Global thresholding still plays an important role, because it is suitable for real-time processing. However, as concluded from [33], global thresholding has poor accuracy compared to other image segmentation algorithms. If the hardware of weather radar processors permits in the future, it is worthwhile to investigate more advanced methods to obtain better results.

6. Filters further application on different radars

To assess more about the robustness of the filter, one method is to apply it to different radars. Those radars can be FMCW radars with different specifications or pulsed radars.

7. Rainfall rate assessment

Rainfall rate estimation is one of the major application of weather radar. However, this project does not involve that part. The rainfall rate can be retrieved by polarimetric parameters and the process is complicated. It is expected to do the assessment related to rainfall rate in the future.

Bibliography

- [1] Christine Unal. Spectral polarimetric radar clutter suppression to enhance atmospheric echoes. *Journal of Atmospheric and Oceanic Technology*, 26(9):1781–1797, sep 2009. doi: 10.1175/2009jtecha1170.1.
- [2] D. Moisseev;C. Unal;H. Russchenberg;L. Ligthart. Doppler polarimetric ground clutter identification and suppression for atmospheric radars based on co-polar correlation.
- [3] Mi He, Yongjian Nian, Xuesong Wang, Shunping Xiao, and Yongzhen Li. Improved clutter suppression algorithm for atmospheric target detection using polarimetric doppler radar.
- [4] Jiapeng Yin, Christine M. H. Unal, and Herman W. J. Russchenberg. Narrow-band clutter mitigation in spectral polarimetric weather radar. *IEEE Transactions on Geoscience and Remote Sensing*, 55(8):4655–4667, aug 2017. doi: 10.1109/tgrs.2017.2696263.
- [5] Jiapeng Yin, Christine Unal, and Herman Russchenberg. Object-orientated filter design in spectral domain for polarimetric weather radar. *IEEE Transactions on Geoscience and Remote Sensing*, 57(5):2725–2740, may 2019. doi: 10.1109/tgrs.2018.2876632.
- [6] Swarnendu Ghosh, Nibaran Das, Ishita Das, and Ujjwal Maulik. Understanding deep learning techniques for image segmentation. *ACM Computing Surveys*, 52(4):1–35, sep 2019. doi: 10.1145/3329784.
- [7] Rafael Gonzalez. *Digital image processing*. Pearson, New York, NY, 2018. ISBN 9780133356724.
- [8] Gerhard Neuhold, Tobias Ollmann, Samuel Rota Buló, and Peter Kotschieder. The mapillary vistas dataset for semantic understanding of street scenes. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4990–4999, 2017.
- [9] Image Segmentation in Deep Learning: Methods and Applications. URL <https://missinglink.ai/guides/computer-vision/image-segmentation-deep-learning-methods-applications/>.
- [10] Hui Zhang, Jason E. Fritts, and Sally A. Goldman. Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, 110(2):260–280, may 2008. doi: 10.1016/j.cviu.2007.08.003.

- [11] Niall O'Mahony, Sean Campbell, Anderson Carvalho, Suman Harapanahalli, Gustavo Velasco Hernandez, Lenka Krpalkova, Daniel Riordan, and Joseph Walsh. Deep learning vs. traditional computer vision. In *Advances in Intelligent Systems and Computing*, pages 128–144. Springer International Publishing, apr 2019. doi: 10.1007/978-3-030-17795-9_10.
- [12] Orlando José Tobias and Rui Seara. Image segmentation by histogram thresholding using fuzzy sets. *IEEE transactions on Image Processing*, 11(12):1457–1465, 2002.
- [13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [14] Dorin Comaniciu and Peter Meer. Robust analysis of feature spaces: color image segmentation. In *Proceedings of IEEE computer society conference on computer vision and pattern recognition*, pages 750–755. IEEE, 1997.
- [15] Jonathan Masci, Ueli Meier, Dan Cireşan, and Jürgen Schmidhuber. Stacked convolutional auto-encoders for hierarchical feature extraction. In *International conference on artificial neural networks*, pages 52–59. Springer, 2011.
- [16] Jordi Freixenet, Xavier Muñoz, David Raba, Joan Martí, and Xavier Cufí. Yet another survey on image segmentation: Region and boundary information integration. In *European conference on computer vision*, pages 408–422. Springer, 2002.
- [17] Larry Medsker and Lakhmi C Jain. *Recurrent neural networks: design and applications*. CRC press, 1999.
- [18] N Senthilkumaran and Reghunadhan Rajesh. Image segmentation-a survey of soft computing approaches. In *2009 International Conference on Advances in Recent Technologies in Communication and Computing*, pages 844–846. IEEE, 2009.
- [19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [20] Sandeep Rao. Introduction to mmwave sensing: Fmcw radars.
- [21] Jordi Ventura. *Design of a high resolution X-band Doppler polarimetric weather radar*. s.n, S.I, 2009. ISBN 9789090247595.
- [22] Kyosuke Hamazu Shoichiro Fukao. *Radar for Meteorological and Atmospheric Observations*. Springer-Verlag GmbH, 2013. ISBN 9784431543343. URL https://www.ebook.de/de/product/21992052/shoichiro_fukao_kyosuke_hamazu_radar_for_meteorological_and_atmospheric_observations.html.

- [23] Scott Krig. *Computer Vision Metrics*. Springer-Verlag GmbH, 2016. ISBN 9783319337623. URL https://www.ebook.de/de/product/27955495/scott_krig_computer_vision_metrics.html.
- [24] C. E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28(4):656–715, oct 1949. doi: 10.1002/j.1538-7305.1949.tb00928.x.
- [25] Jiaquan Chen, Binglei Guan, Hailun Wang, Xuguang Zhang, Yinggan Tang, and Wenzhao Hu. Image thresholding segmentation based on two dimensional histogram using gray level and local entropy information. *IEEE Access*, 6:5269–5275, 2018. doi: 10.1109/access.2017.2757528.
- [26] D.M. Endres and J.E. Schindelin. A new metric for probability distributions. *IEEE Transactions on Information Theory*, 49(7):1858–1860, jul 2003. doi: 10.1109/tit.2003.813506.
- [27] J. Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991. doi: 10.1109/18.61115.
- [28] V. Chandrasekar V. N. Bringi. *Polarimetric Doppler Weather Radar*. Cambridge University Press, 2011. ISBN 0521623847. URL https://www.ebook.de/de/product/3406159/v_n_bringi_v_chandrasekar_polarimetric_doppler_weather_radar.html.
- [29] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. URL <http://www.scipy.org/>.
- [30] R. H. HARALICK. Survey-image segmentation techniques, computer vision. *Graphics and Image Processing*, 12:60–73, 1980. URL <https://ci.nii.ac.jp/naid/10003789064/en/>.
- [31] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors. scikit-image: image processing in Python. *PeerJ*, 2:e453, 6 2014. ISSN 2167-8359. doi: 10.7717/peerj.453. URL <https://doi.org/10.7717/peerj.453>.
- [32] Hui Zhang, Jason E Fritts, and Sally A Goldman. Entropy-based objective evaluation method for image segmentation. In *Storage and Retrieval Methods and Applications for Multimedia 2004*, volume 5307, pages 38–49. International Society for Optics and Photonics, 2003.
- [33] AM Khan and S Ravi. Image segmentation methods: A comparative study. 2013.