CLASSIFICATION OF MICRO-DOPPLER SIGNATURES WITH THE USE OF ORTHOGONAL MOMENT-BASED FEATURES

CLASSIFICATION OF MICRO-DOPPLER SIGNATURES WITH THE USE OF ORTHOGONAL MOMENT-BASED FEATURES

MASTER THESIS

To obtain the degree of Master of Science in Electrical Engineering at Delft University of Technology, to be defended publicly on Monday, July 3, 2023 at 09:30 AM

by

Bob VAN NIFTERIK

This thesis has been approved by

Supervisor: Dr. Francesco Fioranelli

Thesis Committee:

Dr. Francesco Fioranelli	Technische Universiteit Delft
Dr.ir. R.C. Hendriks	Technische Universiteit Delft

Keywords: Micro-Doppler, Orthogonal moments, Classification

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

CONTENTS

Abstract vii

Preface ix

1 Introduction 1
- 1.1 Classification pipeline . . . 2
- 1.2 Contributions of this thesis . . . 3
- 1.3 Structure of the thesis . . . 3

2 Literature review on Orthogonal moment-based classification for micro-doppler signatures 5
- 2.1 Radar data and data representation . . . 6
  - 2.1.1 Data and target analysis . . . 6
  - 2.1.2 State of the art pre-processing techniques . . . 8
- 2.2 Orthogonal moments for classification . . . 9
  - 2.2.1 Discrete moments . . . 11
  - 2.2.2 Continuous circular moments . . . 14
  - 2.2.3 Continuous cartesian moments . . . 16
  - 2.2.4 Summary on orthogonal moments for classification . . . 17
  - 2.2.5 State of the art techniques for micro-Doppler signature classification . . . 18
- 2.3 Summary and research gaps . . . 18

3 Orthogonal moments 25
- 3.1 Moments as feature . . . 25
- 3.2 Continuous circular moments . . . 26
- 3.3 Continuous moments defined on a square grid . . . 33
- 3.4 Discrete moments . . . 37
- 3.5 To summarize . . . 55
- 3.6 Contribution of code . . . 56

4 Micro-Doppler signature data 57
- 4.1 Data generation . . . 57
  - 4.1.1 Micro-Doppler signatures . . . 57
- 4.2 Datasets used for the analysis . . . 58
  - 4.2.1 Dop-NET Dataset . . . 58
  - 4.2.2 Glasgow dataset . . . 60
  - 4.2.3 Cross-frequency set - XeThru . . . 62
  - 4.2.4 Topology overview . . . 63
- 4.3 CVD as pre-processing method . . . 63
- 4.4 Noise introduction . . . 64

5 Classification Processing Step **65**
5.1 Feature vector generation . . . 65
5.1.1 Complex moments . . . 65
5.1.2 Feature vector normalization . . . 67
5.2 Euclidean distance-based classifiers . . . 67
5.2.1 K-nearest neighbor (KNN) . . . 67
5.2.2 Support vector machine (SVM) . . . 67
5.3 Feature selection . . . 68
5.3.1 Univariate feature ranking for regression using F-tests . . . 69
5.3.2 Univariate feature ranking for classification using chi-square tests . . . 69
5.4 Evaluation metric and pipeline . . . 69
5.4.1 Proposed pipeline . . . 70

6 Testing & Results **71**
6.0.1 Note on the presentation of the results . . . 71
6.1 General classification results . . . 71
6.1.1 Influence of the different data sets . . . 72
6.1.2 Comparison to neural networks . . . 73
6.1.3 Assessment of moments defined on a disk . . . 74
6.1.4 Proposed feature vector for complex moments . . . 74
6.1.5 Assessment of continuous moments defined on a square grid . . . 75
6.1.6 Assessment of discrete moments defined on a square grid . . . 75
6.1.7 Effect of variation of moment parameters . . . 76
6.1.8 Overview evaluation of required moment order . . . 78
6.2 Classifier influence . . . 79
6.3 Assessment of performance of the CVD . . . 81
6.3.1 Assessment of the influence of noise introduction . . . 82
6.4 Classification results with feature selection . . . 85
6.5 Summary of results . . . 88

7 Conclusion and Future work **89**
7.1 Summary and Conclusion . . . 89
7.2 Future work . . . 91

A Recursive Gegenbauer moment calculation **103**

B General results **105**
B.1 Influence of the moment orders . . . 105
B.2 Influence of the CVD . . . 106
B.3 Influence of the classifier . . . 108

ABSTRACT

Radar micro-Doppler signatures are powerful indicators of target movements and activities, enabling the extraction of valuable information about various objects' internal and external dynamics. Consequently, classifying these signatures has become crucial in numerous applications, ranging from target recognition in surveillance, to biomedical sensing and interaction with smart sensors.

In this thesis, an evaluation of classification performances for a wide variety of orthogonal moments, when applied to micro-Doppler classification problems, is presented. A pipeline is proposed to evaluate all moments commonly used in image processing, but not routinely employed in radar-based classification. The evaluation results are compared with other state-of-the-art classification approaches, such as using micro-Doppler signatures directly as the input of Convolutional Neural Networks. The influence of noise in the data on the classification performance is also shown.

The classification results demonstrate the different moments' capabilities with a variety of publicly available datasets containing human micro-Doppler signatures, resulting in a very well performing classification pipeline for this type of classification problem, and novel insights into the potential of these moments for radar classification problems.

PREFACE

With great satisfaction, I present this preface as it signifies the completion of my final master thesis, symbolizing a significant milestone in my academic journey. I would like to express my sincere appreciation to the individuals who have played significant roles in its completion. I am profoundly grateful to my supervisor, Francesco Fioranelli, for his guidance, expertise, and support throughout this research endeavor. I would also like to extend my thanks to the professors who have contributed to my academic growth during my study. Their knowledge and inspiration have greatly influenced the ideas presented in this thesis.
Special thanks go to Jurgen for your companionship in our shared office and Ze-sheng, Marien, Patrick and Florens for our enjoyable lunches. Furthermore, I would like to express my gratitude to the housemates of the JVB and my current housemates for creating a supportive environment and providing companionship during the writing of this thesis. Lastly, I would like to express my deep appreciation to my family for their encouragement.

I offer my sincere thanks to all those mentioned above and to anyone else who has contributed in any way. Your support has been pivotal in the successful completion of this final master thesis.

Bob van Nifterik
Delft, June 2023

1

INTRODUCTION

Radar signal processing has experienced a paradigm shift in the analysis of complex radar returns since the introduction of micro-Doppler signatures [1]. Micro-Doppler signatures have emerged as powerful indicators of target motion and activities, enabling the extraction of valuable information about various objects' internal and external dynamics. Consequently, classifying micro-Doppler signatures has become crucial in numerous applications, ranging from target recognition in surveillance, to biomedical sensing [2][3].

The classification of micro-Doppler signatures has emerged as a thriving field of research, with numerous methods being proposed to address this task. In recent years, machine learning algorithms have gained prominence in this domain, with neural networks emerging as the dominant approach [3]–[5]. Other methods rely on handcrafted features with a more conventional classification algorithm, for example, K-nearest neighbor (KNN) and support vector machines (SVM)[6][7]. These methods provide better insight into the explainability of the classification process. In addition, they are, on average, less computationally expensive than neural networks.

One of the earlier proposed handcrafted feature methods is to extract the pseudo-Zernike moments of the radar spectrogram, which is used as a feature[8]. The pseudo-Zernike moments are part of a larger family of orthogonal moments. This family is most commonly used in image processing and classification, where it has shown the ability to have good compression and classification potential. More generally, orthogonal moments utilize different basis functions that meet the orthogonality criterion with varying properties. These differences have an impact on the classification behaviour.

While research exists on comparing and testing various moments for classification in image processing [9], *there is a lack of comprehensive research on the suitability of orthogonal moments for micro-Doppler classification.* Currently, only a few moments have been proposed for this purpose, where no evidence is given as to why other moments would not perform as well. Furthermore, the classification performances of a processing pipeline with orthogonal moments can vary when other processing blocks are changed, such as the data set, the data format and the classifier. Therefore it is hard to say with

1

the current knowledge in the state of the art how orthogonal moments contribute to the classification process without assessing those dependencies.
In this thesis, a pipeline is proposed and experimentally verified to evaluate the classification performances of a wide variety of orthogonal moments applied to radar micro-Doppler signatures.

1.1. CLASSIFICATION PIPELINE

This thesis aims to evaluate the most commonly researched orthogonal moments in image-based classification when applied to radar micro-Doppler classification problems. A large portion of the novel contributions of the thesis originates from the literature study in Chapter 2. This evidenced that rather limited research has been performed on this topic, essentially showing the wide variety of moments proposed in the image processing literature, but their very limited investigation when considering radar-based classification problems. To fill this gap, a classification pipeline is formulated and experimentally verified as the backbone of the research presented in this thesis. Additional details on the several constituent elements are discussed more thoroughly in Chapter 5, but the general structure of the pipeline is presented as follows.

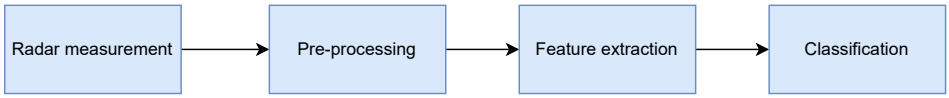


Figure 1.1: General structure of the investigated classification pipeline

The several components of the pipeline are:

- **Radar measurement:** Dataset used for the classification problem at hand, in this thesis specifically focusing on experimental data in the context of human micro-Doppler signatures.
- **Pre-processing:** The type of processing of the data performed before extracting the moments, i.e., how the radar data are processed and formatted to be used by the moment extraction algorithms.
- **Feature extraction:** The extraction of class-defying features, in this thesis specifically by calculating different moments. Table 2.3 displays all the orthogonal moments used in this study.
- **Classification:** The type of classification algorithm used, for example, K-NN and SVM among others, where the extracted features are used for classification.

Different options for the above components of the classification pipeline will be investigated, with a specific emphasis on assessing classification performances when different orthogonal moments are used as features. These performances will eventually be tested against the classification performances of other state-of-the-art algorithms, e.g., convolutional neural networks that use micro-Doppler signatures directly as input.

1.2. CONTRIBUTIONS OF THIS THESIS

The main contributions of this thesis are summarized below:

- An assessment of classification performances for a *wide variety of orthogonal moments used as features* for experimental micro-Doppler datasets is presented. To the best of the author's knowledge, this is the first comprehensive investigation of the suitability of diverse orthogonal moments for radar-based classification. These results are compared with other common classification algorithms.
- The influence of different variants for the key components of the proposed classification pipeline is thoroughly investigated. These variants include: three experimental *datasets* collected with different models of radars and participants; different orthogonal *moments* as well as different orders and parameters in the definition of their orthogonal basis functions; different types of *classifiers* using the moments as features.
- The effect of adding *noise* in radar data on classification performance is evaluated and compared with other common classification algorithms, demonstrating that the proposed moment-based approach is more robust to noise than convolutional neural networks.
- All studied orthogonal moments have been provided with a rigorous mathematical implementation and cleaned *MATLAB code*, publicly shared in `https://doi.org/10.4121/efd059aa-f2c2-4c54-be85-4e600876f444` [10]. Furthermore, the main results of this thesis are currently being summarized for *publication* in the IEEE Transactions on Radar Systems, or comparable journal.

1.3. STRUCTURE OF THE THESIS

The thesis addresses all elements of the base pipeline presented in Figure 1.1. Chapter 2 presents a comprehensive review of the literature on the classification of micro-Doppler signatures using orthogonal moments. The review highlights the current state of the art and identifies research gaps. Next, Chapter 3 presents the moments evaluated in this work and provides a mathematical explanation of their calculation to high orders. Afterwards, Chapter 4 introduces and evaluates various datasets and pre-processing techniques, covering the 'Radar measurement' and 'Pre-processing' components blocks as outlined in the previously proposed block diagram. In Chapter 5, several classification algorithms are proposed, which are suitable for testing the classification based on the moments. In addition, the construction of feature vectors is discussed. Finally, Chapter 6 presents the results, with different selected studies that highlight the various effects of the variations made in the pipeline. Finally, Chapter 7 concludes the thesis.

2

LITERATURE REVIEW ON ORTHOGONAL MOMENT-BASED CLASSIFICATION FOR MICRO-DOPPLER SIGNATURES

This chapter provides a review of the literature on previous work in micro-Doppler signature classification and feature extraction using orthogonal moments. This includes an analysis of the data, data preprocessing, and feature extraction using orthogonal moments. It is shown that only a very limited analysis exists in the literature for the suitability of orthogonal moments for micro-Doppler based classification.

Detecting and classifying micro-Doppler signatures from radar data has become a growing area of study since its introduction in [1]. Since then, several methods have been put forth, with a significant portion of the contributions being based on machine learning-related approaches, including neural networks, SVM, and k-NN algorithms. The input of these algorithms typically consists of carefully selected features. For this, a wide variety of approaches are proposed to optimize the the distinctiveness of the extracted features. One such approach is proposed in [11] where a type of orthogonal moments is suggested as a feature for micro-Doppler signature classification. Orthogonal moments (OM) were a brand-new feature for micro-Doppler radar data and demonstrated the potential of using orthogonal moments in radar classification. However, the usage of OM for classification is not new. This approach was previously widely used in image and pattern recognition. Clemente's study in [11] demonstrated that one specific OM provides accurate classification results. Although more moments were used for micro-Doppler data after the aforementioned study by Clemente et al., comprehensive research encompassing all relevant OMs in micro-Doppler classification is missing.

2.1. RADAR DATA AND DATA REPRESENTATION

In micro-Doppler signature classification, different targets are of interest. These targets will have different micro-dynamic behaviour and, therefore, different signatures and properties of interest for a classification algorithm. Different types of radar and radar settings are preferred in measuring the different targets. A simple classification pipeline is presented in Figure 2.1. The blocks illustrate the stages in the process, from measurement to classification. The first block: Radar measurement data, is discussed in section 2.1.1 containing the analysis of targets, applications and data topology. Furthermore, publicly available data sets will be evaluated and, lastly, in section 2.1.2, a review of the state-of-the-art prepossessing techniques will be presented.

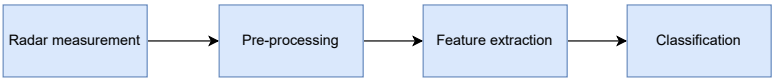


Figure 2.1: General structure of the investigated classification pipeline

2.1.1. DATA AND TARGET ANALYSIS

One of the most researched targets in micro-Doppler signature recognition is human motion. Where in [12], human walking models are created to simulate the micro-Doppler response of a walking man. In several types of research on micro-Doppler of human motion, [13][14] the harmonics created by the non-rigid body, e.g. arms, legs, torso, etc. define the class of the activity. This information is mainly spread around the bulk motion of the object. In the case of gait analysis [14], the peaks of the multibody limbs specify the class. The length of the recording for gait analysis, compared to other human activity data, i.e. hand gestures [15] or activity detection [16], is longer. This gives the impression that when dealing with human gait walking analysis or human gestures, the signature's length will depend on the target class. In gait analysis, as done in [14], the whole captured time window information is present. In contrast to a semifinite data sequence in a captured hand gesture, only a part of the captured window contains Doppler frequencies related to their class.

Rotor blades, e.g. helicopter blades, are studied in [17] [18]. The Doppler data can be seen as a superposition of each blade's Doppler response. The rotor blade commonly has a higher Doppler velocity than human motion signatures. The harmonics produced by the rotating blades are visible in the Doppler spectrogram. The clear presence of this rotational frequency makes techniques based on extracting the repetition frequency of Doppler frequencies in time a good feature for class detection [19]. In [20], a review is given in different target signatures. The most common targets are discussed, summarizing the research done and the most common types of radar/settings used for this type of target.

Another field where a clear characteristic harmonic defines the target class is the detection of spinning ballistic missiles. As studied in [21], the data is simulated as an explicit continuous harmonic related to the Doppler frequency of a spinning object.

For classification using orthogonal moments, the topology of the data, i.e. how the data spatially look on an image, is important. The moments represent the image from orders of the basis function, where more errors are introduced in some regions of the image than in other regions. The location of the errors depends on the type of moment used. Therefore, to be able to give a good evaluation of the OMs for classification, the data topology is analyzed. The radar data topology from the different targets can be divided into two categories:

- **Whole time frame**: The signature stretches the entire captured time frame. Most continuous harmonic targets fall into this category; an example would be the rotor blades of a drone.
- **Partial time frame**: The signature is present only in part of the captured time frame. All temporal activities fall into this category. An example would be the signature of a person falling.

Further, the micro-Doppler data' presence of harmonics in the spectrogram will be important in the prepossessing step. Especially when using techniques relying on the cadence of an object. This is further discussed in Section 2.1.2. For evaluation, the data is divided into three categories;

- **Clearly Present**: The harmonics of the microdynamics are visible and distinctive for the signature class, e.g. the rotor blade of a drone.
- **Present**: The microdynamic harmonics are visible but hard to distinguish. The harmonics, or superpositions of harmonics, are still distinctive for the signature class, e.g. the human gait signature.
- **Not visible**: The harmonics of the microdynamics are visible but not distinctive for the signature class, e.g. human activity signature.

As discussed in this section, the analyzed data from previous research are summarized topology-wise in table 2.1.

Object type	**Harmonic micro dynamics**	**Data Spread**	**Reference**
Human Walking	Present	Whole time frame	[12], [13],[14]
Human activity	Not visible	partial time frame	[13]
Hand gesture	Not visible	partial time frame	[15]
Rotor	Clearly Present	Whole time frame	[17], [18]
Ballistic missiles	Clearly present	Whole time frame	[21]

Table 2.1: Summary of micro-Doppler data topology

The most relevant publicly available data sets are shown in table 2.2. Most of the relevant available data sets focus on the classification of human behavior. Additionally, all references that link to the research in which the data set is presented include research

Dataset	Reference	Target	Radar	Remark	Classes	Samples	Institute
Dop-NET	[22]	Human hand gesture	CW FMCW	Single activity each sample	4	3052	Delft University of Technology
Radar signatures of human activities	[2]	Human Activity	FMCW	Single activity each sample	10	1754	University of Glasgow
Continuous Activities of Daily Living		Human Activity	UWB * 5	Multiple radars, Multiple activities on one sample	9	-	Delft University of Technology
ASL-Sequential-Dataset	[23]	Human Activity	77 GHz FMCW MIMO	Multiple activaties each sample	5	-	The university of Alabama
OPERAnet	[24]	Human Activity	UWB	More signals are present in data set e.g. PWR, SDR	10	-	UK Engineering and Physical Sciences Research Council (EPSRC)
Cross-frequency	[25]	Human Activity	77 GHz FMCW, 24 GHz FMCW, XeThru UWB	Same targets for three different radars	11	-	Delft University of Technology, University of Alabama

Table 2.2: Summary of the analyzed datasets with experimental micro-Doppler data

with classification results. These results could be used as a benchmark for new proposed classification techniques.

TO SUMMARIZE

There is a wide variety of micro-Doppler signatures. The topology of the signatures differs significantly, depending on the type of target. The summary of the different topologies of target signatures is given in Table 2.1. For publicly available data, most of the data sets are based on human movements and activities.

2.1.2. STATE OF THE ART PRE-PROCESSING TECHNIQUES

The output of radar is not a Doppler-time representation of the data. Raw data first need to be processed in the desired format. This step is referred to as the prepossessing step in the classification pipeline. The output of the prepossessing step defines how the data is shaped before feature extraction. Therefore, the effect of the feature will depend on how the data is processed. The Doppler time spectrogram is used in the general case of Doppler classification. To obtain the Doppler representative of the data, the STFT has to be taken over the range in the range time representation, referred to as the spectrogram [26]. The spectrogram is mainly used as a first tool to evaluate micro-Doppler data. Extensive research exists where the features are directly extracted from the spectrogram, e.g. the separation of different harmonic components[27][28][29]. The spectrogram is an intermediate stage of the full preprocessing pipeline and will be followed by another mapping, e.g., the Cadence velocity diagram.

CADENCE VELOCITY DIAGRAM

The Cadence Velocity Diagram(CVD) maps the radar data's spectrogram by calculating the Fourier transform along its time axis.[30] This returns the frequency at which the different micro-doppler frequencies occur in time. The representation technique has been shown to be useful in human activity recognition[30][31]. It relies on the characteristic harmonics of the microdynamic motions that create the micro Doppler response, e.g. recognition of the frequency of moving body parts in human activity. In addition to human recognition in [19], the CVD was used to classify different kinds of drones. Here, the CVD is projected onto the frequency axes, resulting in the Cadence Frequency Spectrum

(CFS). The main peaks of the CFS have proven to be a good feature [11] [32]. The CVD can be calculated as a step before mapping onto an orthogonal basis function to create moments as a feature. This mapping changes the data's topology and emphasizes the target's cadence. Table 2.7 gives a summary of the most relevant classification research in micro-doppler. From this evaluation, it can be seen that CVD is the most commonly used type of mapping used as a tool for preprocessing. The lack of other representations when extracting features using OM can be an opportunity for future work.

OTHER PROMISING STATE-OF-THE-ART TECHNIQUES

Next to the CVD, several other promising techniques have been proposed. The downside of implementing the lesser-used techniques will be that fewer results will be comparable as a benchmark for the classification of the proposed feature extraction algorithm. Some lesser-used algorithms that showed promising results/characteristics are listed below.

- In [27] the short-time fractional Fourier transform is used to map the spectrogram. The STfrFT has the advantage that for fast-varying signals, the STfrFT will perform better resolution compared to FTs than STFTs [33].

- Joint time-frequency analysis approaches such as the Hilbert Huang transform where the HHT has proven to be more effective for non-linear and non-stationary signals[34].

- In [35], the inverse Radon Transform is performed on the spectrogram. Where the IRT mapping produces another topology of the data compared to other transformations related to Fourier, e.g. Wavelet, that mainly focus on the change in resolution, this change in data topology could be a good variety for the data representation as input for a moment extraction algorithm. Also, it has been shown to provide good results under the influence of noise [36].

TO SUMMARIZE

A spectrum of techniques exist that have proven effective; the CVD is the most widely used step in formatting radar data as a pre-processing tool. However, the benefit of the CVD can be questioned when dealing with non-periodic signals. The CVD mapping is most prevalent in current state-of-the-art research regarding micro-Doppler signature classification, where the analysis of the different types of mappings and pre-processing techniques would be a field of future research.

2.2. ORTHOGONAL MOMENTS FOR CLASSIFICATION

Orthogonal moments have proven to be a suitable type of feature of micro-Doppler signatures as in the classification pipeline in Figure 2.1. This section evaluates the research performed in the literature on OMs in micro-Doppler classification. The evaluation includes an analysis of the general classification in pattern and image recognition. This illustrates the research gap between the orthogonal moment-based classification of microdoppler signatures compared to the field of pattern and image classification.

ORTHOGONAL MOMENTS

Several moments have been proposed after Hu proposed moment invariants for image recognition in [37]. The moments of interest are the moments of which their basis function, satisfies the orthogonality condition[1]. The property of orthogonality in the basis functions means that the inner products of two different basis functions from the same set of basis functions are perpendicular, i.e., the two functions are uncorrelated and uncorrelated with all the other functions from the basis set. This uncorrelated property means no redundancy possible at the moment set [38]. Furthermore, orthogonal moments are used because of their robustness against noise, rotation, scaling and translation [39].

A division can be made between two categories of moment functions; Continuous and discrete. Continuous functions need to be discretized to map a digital image. This introduces corresponding computational errors, which makes them, in theory, less suitable for high-precision image processing [38].

Moment type	**Abbreviation**	**Reference**
Legendre moments	LM	[40]
Gaussian-Hermite moments	GHM	[41]
Continuous Chebyshev moments	CHM	[42]
Gegenbauer moments	GM	[43]
Zernike moments	ZM	[44]
Pseudo-Zernike moments	PZM	[40]
Chebyshev-Fourier moments	CHFM	[45]
Pseudo Jacobi-fourier moments	PJFM	[46]
Jacobi-Fourier moments	JFM	[47]
Exponent-Fourier moments	EFM	[48]
Bessel-Fourier moments	BFM	[49]
Radial harmonic Fourier moments	RHFM	[50]
Orthogonal Fourier-Mellin moments	OFMM	[51]
Krawtchouk moments	KM	[52]
Hahn moments	HM	[53]
Racah moments	RM	[54]
Dual Hahn moments	DHM	[54]
Discrete Chebyshev moments	CHDM	[42]
Meixner moments	MM	[55]
Charlier moments	CM	[55]
Polar complex exponential moments	PCEM	[56]

Table 2.3: Moment types investigated in this thesis and their abbreviations

There are two main classes of coordinate systems on which basis functions are defined; Cartesian and circular. The circular basis has the advantage that functions represented in polar coordinates perform better in rotational invariance[38]. When dealing with data

[1] $\langle V_{nm}, V_{n'm'} \rangle = \iint_D V_{nm}(x,y) V^*_{n'm'}(x,y)\,dx\,dy = \delta_{nn'}\delta_{mm'}$

defined on a Cartesian coordinate system, like most digital images, the data will first be mapped onto a unit disk, or there will be a loss of information at the corners of the image. This mapping will introduce computational errors, but techniques are proposed for error reduction[38]. Most relevant research is done on moments defined on a polar coordinate system because of the rotational invariant property. One could question the advantage of rotational invariance when dealing with radar data because of the fixed location of the axes. Table 2.3 lists the moments used for this research. The different families of the moments discussed can be also seen in figure 2.2.

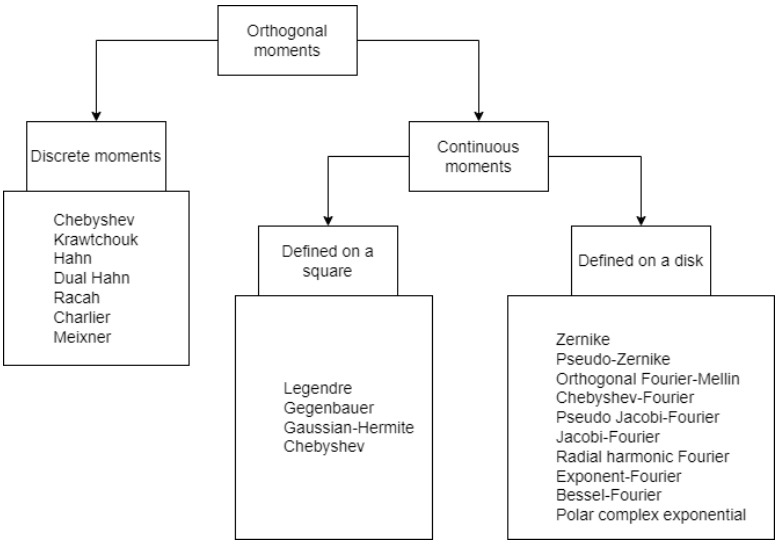


Figure 2.2: Schematic view on the different classes of orthogonal moments, presented in detail in Chapter 3

2.2.1. DISCRETE MOMENTS

The first discussed group of moments will be the discrete set of moments. The general advantages of discrete moments are:

- No discretization step is needed when calculating the moments. A higher theoretical reconstruction of the image is possible without the discretization errors. Therefore, this family of moments is more suitable for high-precision reconstruction.
- The discrete moments are defined on a square Cartesian grid $D = \{(x,y) : x \in [-1,1], y \in [-1,1]\}$. Most digital images can be easily mapped on a square domain; therefore, almost no information will be lost when projecting the original image on the basis function.

The most relevant research on discrete moments is briefly discussed below and evaluated for each moment. The mathematical description and implementation are presented in Chapter 3.

RACAH MOMENTS (RM)

Racah polynomials are hypergeometric polynomials defined on a discrete Cartesian grid. As described before, this has the advantage that there will be no calculation errors in discretizing the basis function. Also, the image function does not need to be mapped; i.e., the image is defended on a square Cartesian domain. [54] The Racah polynomial can be seen as a more general form of the other orthogonal functions in this family, e.g., Hahn, Dual Hahn, Meixner, Krawtchouk, Chalier, Hermite. These orthogonal polynomials can be derived from the Racah polynomial by changing their limits [57]. The askey scheme describes the relations of Hypergeometric Orthogonal polynomials [39]. This also described the relations between the discrete orthogonal functions and their continuous twins, e.g., Wilson, continuous Dual Hahn, continuous Hahn, Meixner - Pollaczek, Jacobi, Laguerre.
From previous research, the Racahn moments used for image recognition better reconstructed the original image in a noise-free case than DCT, Legendre moments, and discrete. Tchebichef moments and discrete Krawtchouk moments. [54]. Racahn moments can be transformed into invariant Rachan moments that are rotational and scale invariant[58].
Another notable observation is that the ROI can be changed by changing the limits in equation 3.66. This has the property that a higher resolution can be obtained for specific areas, using the same order of moments.

HAHN MOMENTS (HM)

From the Racah polynomial as given in section 3.4 the Hahn polynomial, can be formed. See Chapter 3 for the derivation. As for the RM, the Hahn moments also have the ability to change its region of emphasis, i.e., they can change its ROI[53]. This allows them to extract local or global features. In [53], a method for adaptive feature selection is proposed using the property of highlighting the moment in the region of interest. The proposed method uses the centroid as the spread to determine the values of certain parameters. This has shown significantly better results in the reconstruction of the ROI compared to Chebyshev, Legendre, Zernike and Hahn.[53] provides empirical proof that an image can be represented in lesser moments and still contain the same level of information.

DUAL HAHN MOMENTS (DHM)

As with the Hahn polynomial, the dual Hahn can be derived from the Racah polynomial. In [59] dual Hahn moments are proposed, together with KM and CHM, to extract local and global features. This proposal aims to prove the effectiveness of these moments in near-infrared and visible light. The proposed moment shows encouraging results in terms of accuracy, error rate, and classification.

KRAWTCHOUK MOMENTS (KM)

From previous research, Krawtchouk is one of the most promising orthogonal moments, when looking at reconstruction capabilities [54], [52],[60]. The weighted Krawtchouk moment has some remarkable properties. Krawtchouk moments can change its ROI, i.e., adjusting its resolution on certain parts of the image. In [60], the performance of

the reconstruction ability is shown to depend on the location and size of the higher-resolution areas in the image. The ability to move the ROI to a high information-dense part of the data could improve the classification results [52]. These information dense parts are referred to as local features, whereas capturing the whole image is referred to as global features. In [61], an analysis is made on the classification of Chinese characters. Each frame is divided into several zones, referred to as local features. It showed that Krawtchouk and Tchebichef outperformed continuous orthogonal movements such as Zernike and Legendre in classification. In [62], the moments are used for iris recognition with a global and local feature extraction algorithm. In this case, the ability to change the ROI of the CHM is exploited. An overview was given in [63] for different OM classifications of micro Doppler signatures. Here, a selection of OM is tested, classifying signatures using a simulated rotor blade as data. Furthermore, in [15], a selection of OM is evaluated in relation to the classification performance of hand gesture signatures. The research done on micro-Doppler classification has shown promising results in classification for KM compared to ZM, PZM, CHM, OFMM and LM. In this research, there is no reference to the ROI.

TCHEBICHEF (CHDM)

The Tchebichef polynomial is related to Racah, Hahn, and Krawtchouk and can be formed from the Hahn polynomial. [53] CHDM are commonly used for image and pattern recognition[61]. In [62] the moments are used for iris recognition with a global and local feature extraction algorithm. There it showed that at lower orders, Tchebichef performs worse than DHM and KM for local features, since the other moments have the ability to change the ROI. In [15] the Tchebichef moments are introduced as classification method for micro Doppler signatures. The data on which the classification is performed are signatures of hand gestures using CVD as a representation. The results are compared to pseudo-Zernike, Krawtchouk and Doppler Time and Range (DTR) based features. The study showed that the Tchebichef moments got the best performance in the classification rate. In [64], the CHM is introduced as a tool for the Classification, Recognition and Fingerprinting of Drones. As validation, there are several experiments performed with drone/bird, loaded/unloaded drone, and a fixed-wing/multirotor classification. In general, the CHM showed good classification results compared to the KMs that are implemented as references. The classification step is performed with three different classifiers: KNN, SVM, and RF.

CHARLIER (CM)

Charlier Moments are part of the family of discrete moments. The basis function has free parameters influencing the location of the ROI. In [65] the recursive implementation of the moment is given. In addition, the moment is used for image reconstruction where it is compared to the Krawtchouk moment.

MEIXNER (MM)

Meixner Moments are part of the family of discrete moments [57]. The moments' basis function contains free parameters which enables the moments to change their ROI. In [66] the moment is used for accurate 2D and 3D image classification using translation

and scale invariants. Here the moment is compared in classification performance to Legendre, Tchebichef, Krawtchouk, and Charlier moments.

2.2.2. CONTINUOUS CIRCULAR MOMENTS

The continuous circular moments are the family most widely used in recent research due to their natural rotational invariance [38]. The general advantages/disadvantages of Continuous Circular Moments are:

- Discretization errors are introduced when discretizing the basis function. The introduction of these errors makes it less suitable for high-precision reconstruction [42].
- The radial basis function of the continuous circular moments is rotation invariant. Because of this property, great scientific interest has been given to this category of moments [38].
- The fact that the basic function is defined on a unit disk implies that when dealing with a square image, some information will be lost, e.g. the corners of the image.

JACOBI-FOURIER MOMENTS(JFM)

The JFM can be seen as the generalized case of the other exceptional cases of Jacobi-based moments. Where the Jacobi polynomial is used In [63], the JFM are simulated and compared with other simulated continuous orthogonal moments defined on the unit disk. In this work, the ability to reconstruct and classify using the moments as features and the robustness against noise were tested. The JFM did show average results in these simulations, comparable to the other moments based on the Jacobi polynomial.
In [47], The JFM are calculated for a black and white image of a letter. This is used to calculate the reconstruction error. Other moments based on the Jacobi polynomial are also calculated, and the results are compared. Also, the influence of noise and its relation to other orthogonal moments are studied.

PSEUDO JACOBI-FOURIER MOMENTS(PJFM)

In [63], the JFM are simulated and compared with other continuous orthogonal moments defined on the unit disk. This work tested the ability to reconstruct and classify using the moments as features and the robustness against noise. The PJFM showed comparable results with the JFM where only at high order the PJFM shows minimal better results when Gaussian noise is added.
In [46] PJFM are studied with respect to image reconstruction. The PFM are compared with the OFMM, and the PFM outperforms the OFMM with respect to signal-to-noise ratio and reconstruction.

ZERNIKE MOMENTS(ZM)

[63] Looks at the performance of reconstruction and classification of simulated micro-Doppler data for different Orthonogal moments. Zernike moments are compared to PZM, OFMM, LM, and KM. This research shows that KMs and PZMs perform better with

respect to reconstruction errors. The correct classification rate shows similar results between the moments on higher orders. In [67], the reconstitution of noisy images is examined using orthogonal moments. The signal-to-noise ratio is calculated for different OGs. An original image under the influence of additive zero-mean Gaussian noise is compared to the reconstructed one. This shows that Zernike moments perform better in signal-to-noise ratio after reconstructing an original image compared to Chebyshev moments i.e. Zernike is less sensitive to added noise than Chebyshev moments. The study also presents a comparative empirical study of orthogonal moments in watermarking applications.

PSEUDO-ZERNIKE MOMENTS(PZM)

In [11] the application of pseudo-Zernike moments for micro-Doppler classification is introduced. It compares the classification rates of Zernike and pseudo-Zernike moments when applied to Cadence Velocity diagrams(CVD). In addition, the results showed robustness against translation and scale variation in the microdoppler data. In [15] pseudo-Zernike moments are implemented for the micro Doppler classification of hand-gestation signatures. The main focus of this research is classification using Chebyshev moments, where the pseudo-Zernike moments are used as a reference. CVD is used as an image function that will be mapped using moments. The results show that on lower order, pseudo-Zernike outperforms Chebyshev and Krawtchouk moments; on higher order, the named moments will perform better.

CHEBYSHEV-FOURIER (CHFM)

The Chebyshev-Fourier radial basis function can be derived from the Jacobi basis function; see Chapter 3 for the derivation. In [38] a review is given on continuous orthogonal moments. Here several continuous moments are compared. Moments are evaluated and tested on image representation en classification. These results show that the CHFM performs average in classification on low noise levels and the best classification results on higher noise levels. In [45], CHFM is proposed for image reconstruction. The results of the CHFM are compared with those of OFMM, which defers a lot from its performance in representation error.

FOURIER-MELLIN (OFMM)

The Fourier-Mellin moment basis function can be derived from the Jacobi basis function. In [38], OFMMs are tested, evaluated, and compared to several other orthogonal moments. The results showed that the OFMM have an average classification result in the noiseless and noisy case compared to the other researched moments. In [68] The moments are used as a reference when introducing the BFM for classification. In this research, the OFMMs did not perform better than the ZM and BFM moments. In [63], a study is presented comparing different types of orthogonal moments for classifying microdoppler data. The classification is performed on the CVD of a simulated rotor blade. The results show that the ZM performs better than the OFMM in classification for this type of micro-Doppler data.

EXPONENT-FOURIER MOMENTS (EFM)

Exponent Fourier moments are introduced in [48], where they are compared with BFM ZM and RHFM. This research showed that the exponential Fourier polynomial has much more zeros than ZM of the same degree. This property makes FM more suitable for image description. Also, it illustrates that the unbounded imaginary part causes numerical instability. Exponent Fourier moments have shown promising results in [38], where several continuous moments are evaluated and compared. The EMF showed the best classification rate and the classification rate under the influence of noise. Although Moment is one of the most promising ones in classification, the research on using EFM for classification and, notably, using EFM for micro-Doppler classification is lacking.

BESSEL-FOURIER MOMENTS (BFM)

Bessel-Fourier moments are a type of eigenvalue-based moments.[38] The moments are proposed in [49], where the comparison is made between BFM, OFM and ZM in terms of image analysis aspects. The basis function of the proposed BFM is based on the Bessel function of the first kind and can be seen as a generalization of the orthogonalized complex moments. The results show that the BFM performs better than the OFM and ZM in image reconstruction under normal and noisy conditions. In [68], BFM are again tested and compared against OFFM and ZM; The same conclusion is drawn from the results that BFM are consistently better suited for image retrieval, rotation invariant, and image classification. Research on BFM applied to radar data is lacking.

RADIAL HARMONIC FOURIER MOMENT(RHFM)

Radial harmonic Fourier moments are proposed for the recognition of Chinese characters in [68]. RFMs are tested on simulated, experimental, and transformed experimental data. The experiments showed that RFM outperformed ZMs and JFMs in classifying and transforming experimental data, e.g., under the influence of noise. For the analysis done in [38], the RFM is compared with several continuous moments in classification. The results show that they outperform ZM as seen in other research, and the classification rate is average compared to all analyzed moments.

POLAR COMPLEX EXPONENTIAL TRANSFORMATION (PCEM)

The Polar Complex Exponential Transformation PCET is closely related to Discrete Cosine Transformation (DCT) and discrete sine transformation (DST) [56]. In [69], the PCEM is used for Zero-Watermarking for Multiple Medical Images. Further, in [38], the moments are used for classification and reconstruction, where the moment performed particularly well in the classification of images under the influence of white Gaussian noise.

2.2.3. CONTINUOUS CARTESIAN MOMENTS

The continuous cartesian moments are the moments defined on the square $D = \{(x, y) : x \in [-1, 1], y \in [-1, 1]\}$. The general advantages/disadvantages of discrete moments are listed below, followed by the literature review of the moments of this family. The mathematical derivation and implementation is given in Chapter 3.

- Discretization errors are introduced when discretizing the basis function. The introduction of these errors makes it less suitable for high-precision reconstruction.

- The continuous cartesian moments are defined on a square Cartesian grid $D = \{(x, y) : x \in [-1, 1], y \in [-1, 1]\}$. Most digital images can be easily mapped on a square domain; therefore, almost no information will be lost when projecting on the basis function.

LEGENDRE MOMENTS (LM)

Where in contrast to the variations of the Jacobi-based polynomials discussed earlier, the orthogonality region is a Cartesian square [-1,1]x[-1,1]. In [62], Legendre invariant moments are used as features to classify different types of car logos from pictures. This research is compared to Tchebychef, which showed a better classification result than Legendre. In [40], Legendre moments are compared to pseudo-Zernike, rotational, Zernike and complex moments in terms of reconstruction error. It showed that the SNR under the influence of noise is significantly worse than the other simulated moments. In [63], Legendre moments are used as a feature for micro-Doppler data. In this study, a comparison is made between the other orthogonal moments, e.g. Zernike, Pseudo-Zernike, Krawtchouk, and orthogonal Fourier-Mellin moments. The simulations have shown that the Legendre moments have the highest reconstruction error and the worst classification rate results.

GEGENBAUER MOMENTS (GM)

Gegenbauer polynomials also known as ultraspherical polynomials, is a special case of the Jacobi polynomial and can be derived from the Legendre polynomial. In [43], the GM is proposed for the classification of Chinese characters. The moments are tested on a data set with different parameters. The experiments showed promising results in classification rate but are hard to interpret because they are not compared with another type of moment or type of classification algorithm.

CONTINUOUS CHEBYSHEV MOMENTS (CHM)

The continuous Chebyshev is a special case of Jacobi polynomials. It can be derived from the Gegenbauer polynomial by setting $\lambda = 0$ for the first kind and $\lambda = 1$ for the second kind of Chebyshev polynomials.

GAUSSIAN-HERMITE MOMENTS (GHM)

In [70] the GHM are introduced for the recognition of license plates. The features that are extracted using the moments are used in a back-propagation neural net. In [41] the image reconstruction abilities and properties are evaluated. This evaluation shows that a better result separating image features based on different modes could be obtained because of the different number of zero crossings for the different orders. Also, an analysis was done on several applications, including; Fingerprint extraction and Moving object detection. Research on the GHM for the extraction of micro-Doppler features is lacking. The GHM is applied not commonly, and research regarding the usage of micro-Doppler data is lacking.

2.2.4. SUMMARY ON ORTHOGONAL MOMENTS FOR CLASSIFICATION

The research performed on OMs, as discussed previously in this section, is summarized in table 2.5 and table 2.7. The tables clearly show the research gap between the num-

ber of different moments applied in image and pattern recognition, and applied to radar micro-Doppler classification. Only a few discussed moments have been used for micro-Doppler classification. In table 2.4 an overview is given on what moments have been used for this purpose. Several of the moments that have never been used on radar data outperformed the top-performing moments in image classification, for example, EFM outperformed PZM [9]. Additionally, several moments discussed can extract local features by changing their region of emphasis/region of interest (ROI). This property has never been addressed in research related to micro-Doppler, to the best of the author's knowledge.

2.2.5. STATE OF THE ART TECHNIQUES FOR MICRO-DOPPLER SIGNATURE CLASSIFICATION

In the pipeline, as demonstrated in section 2.5, it is visible that a classification algorithm is needed after the extraction of features. The OMs used as features are, in most relevant research, using a classifier dependent on the Euclidean distance between the moments, e.g. KNN, SVM etc. In the summarizing table 2.5 the most relevant research related to the classification of micro Doppler signatures using orthogonal moments is listed, including the classifier. From this table, it is easily observed that the KNN classifier is used primarily.

NEURAL NETWORKS

Neural networks have been proven to be highly efficient in the field of classification of micro-Doppler signatures [3][71][4]. Because of this performance, neural networks are one of the most popular classification techniques in micro-Doppler classification. In [16] a study is presented in the classification of human behaviour using radar Doppler data. A comparison is made between a neural network, GoogleNet, and a transfer learning method leveraging the AlexNet network using other classifiers, e.g. KNN and SVM. It showed the best results for GoogleNet, which handles feature extraction and classification. The neural networks can be used as classifiers for the extracted features, but can also use the total data window represented as an image as input in most cases.
To conclude, because of the widely used neural networks in micro-Doppler classification, using the neural networks as a reference classification rate for new proposed techniques would be a reasonable choice.

2.3. SUMMARY AND RESEARCH GAPS

From the analysis of the open literature, several knowledge gaps are exposed. These gaps are listed below and are considered as the inspiration for the novel research of this thesis.

- Although a wide variety of moments have been evaluated for image classification purposes, only a small part of them have been studied in research regarding radar data, for example, for classification tasks based on radar micro-Doppler data. Therefore, a comprehensive investigation of the usage of the many existing OMs for radar-based classification is missing, to the best of the author's knowledge.

Moment type	Used for image-based classification	Used for micro-Doppler classification
PJFM	*Yes*	*No*
ZM	*Yes*	*Yes*
PZM	*Yes*	*Yes*
CHFM	*Yes*	*No*
OFMM	*Yes*	*Yes*
EFM	*Yes*	*No*
PCEM	*Yes*	*No*
BFM	*Yes*	*No*
RHFM	*Yes*	*No*
GM	*Yes*	*No*
LM	*Yes*	*Yes*
CHM-1st	*Yes*	*No*
CHM-2nd	*Yes*	*No*
GHM	*Yes*	*No*
RM	*Yes*	*No*
DHM	*Yes*	*No*
HM	*Yes*	*No*
CHDM	*Yes*	*Yes*
KM	*Yes*	*Yes*
MM	*Yes*	*No*
CM	*Yes*	*No*

Table 2.4: Table showing the research gap between the field of image-based classification and radar micro-Doppler signatures classification: only a few moments proposed in image processing are studied for micro Doppler classification

- Research on the influence of data topology when using orthogonal moments is missing. The topology is closely related to the type of target's movement/activity and pre-processing, but is rarely addressed as an aspect affecting the overall performances, for instance in classification problems.
- Research on the influence of pre-processing approaches when using orthogonal moments is lacking. In all the research considered in OM used for radar data, CVD was used as input. No evaluations exist on the effect of different types of pre-processing, to the best of the author's knowledge.
- The influence of noise when using orthogonal moments on radar data is not investigated in depth. Only a small number of the moments discussed have been tested under the influence of noise. Also, an evaluation of robustness against noise compared to other state-of-the-art classification algorithms is lacking.
- The optimization properties of some moments, e.g. the ROI of the KM and, in general, the local and global features that have proven to give good results in image recognition, are typically not examined for radar data.

Table 2.5: Summary of literature on radar micro-Doppler classification using orthogonal moments

Authors	Orthogonal moments discussed	Radar data representation	Classification method	Number of classes	Radar type	Target type	Method
[63]	ZM, PZM, OFMM, LM, KM	CVD	KNN	4	5 Ghz simulation	Simulated rotor blade	Comparison of the different OGMs in terms of reconstruction error and classification rate under the influence of additive noise.
[15]	PZM, KM, CHM	CVD	KNN	4	24 GHz FMCW	Empirical data, Hand gestures	Introduces CHM as a tool for feature extraction from micro doppler CVD and compares the classification with other promising OGM
[11]	PZM	CVD	KNN SVM with RBF	6	Selex ES PicoSAR system operating in DMTI 9.2 GHZ	Empirical data, Rotor blade	Introduces pseudo-Zernike moments for micro doppler classification
[32]	KM	CVD	KNN	4	24 GHz	Empirical data, Individual walking, group walking	A method of micro doppler classification with the use of feature extraction using Krawtchouk is discussed. Using a kNN classifier the performance and results are presented.
[8]	PZM	CVD	SVM with RBF	5	12-18 Ghz	Empirical data, Person/group walking and crawling	Shows the performance of the introduced Pseudo-Zernike moment-based features from micro-doppler radar data.
[64]	CHM, KM	CVD	KNN, SVM, RF	2–6	12-18 Ghz	Empirical data, Birds, drones at different payloads	Introduces Chebyshev moments-based features as a tool to classify, reconstruct and fingerprint drones using micro doppler data.
[72]	KM	CVD	KNN	2	-	Empirical data, warrhead and experimental shapes	Proposes a method of detecting warheads with the use of micro dopplers signatures. Krawchouk moments are used for the feature extraction.

Table 2.6: Summary of literature on orthogonal moment based classification

Authors	Moments discussed	Application	Data	Classification method	Method	Classes
[73]	KM	Facial recognition	100×80 pixel grayscale image	KNN	Proposes recognition method using KMs extracting both globals and local features	
[38]	ZM, PZM, OFMM, CHFM, PJFM, JFM, RHFM, EFM, PCET, PCT, PST, BFM	Pattern Recognition	128 × 128 pixel RGB	KNN, CNN(reference)	A comprehensive survey of the Othonogal moments, defined on a unit circle, for image processing. Presents reconstruction and classification results of the discussed moments	
[58]	RM, KM, HM,CHM	Image classification	128x128 Coil-20 databse	KNN	Introducing racah moments for Image classification. The results are presented for all the discussed moments, using two different data sets.	
[62]	LM, CHM	Pattern Recognition, Vehicle logo recognition	70x70 pixel grayscale image	MMD	Proposes Vehicle Logo Recognition technique, where legende and tchebichef moments are used for feature extraction	
[61]	CHM, KM, ZM, LM	Recognition of chinese characters	64x64 pixel grayscale image	DHMM, K-means	Presents a performance study of the recognition of Chinese characters using different discrete orthonogal moments	
[74]	ZM	Posture Recognition of Pigs	binary contour images	SVM	Uses a binary contour image of pigs to classify their posture. Where Zernike moments are used for feature extraction	
[59]	CHM, KM, DHM	high-security applications, Iris Recognition	320x280, 320x240 grayscale 576x768 RGB	K-NN	Different discrete orthogonal moments are proposed to extract global and local features for iris recognition. This has proven to be effective for visible and near-infrared images	

Table 2.7: Summary of literature on orthogonal moment based classification

Authors	Moments discussed	Application	Data	Classification method	Method	Classes
[75]	RHFM, JFM,ZM	Chinese Chess Character Recognition	64x64 grayscale	Minimum city block distance classifier	RHFMs are used to extract features from images of Chinese chess images and toy images. This is used to demonstrate the performance of RHFM	
[68]	BFM, ZM, OFFM	Image Retrieval and Classification	25x16 2d image, 144x144 2d colour (ALOI database)	KNN	Proposes an efficient way of extracting BFM and texture features for image retrieval and classification. Different data sets are used.	
[76]	KM, CHM, ZM	static hand gesture recognition	104x104 Normalized	Minimum distance classifier	Explores different types of OM for detection of static hand gestures, where a Minimum distance classifier is used. The results suggest that the Krawtchouk moment features are comparatively robust to view point changes.	10
[77]	HM	Classification System for Cervical Cell Images	150x150 Normalized	SVM, QSVM, fine gaussian SVM	A classification system is proposed based on Hu moments features. Where the classification is done for three different classifiers; cubic SVM, Quadradic SVM, fine Gaussian SVM,	3
[78]	LM, ZM, PZM	Classification of Ground vehicle SAR Images	128x128	LibSVM(SVM)	Several continuous moments are examined for the classification of SAR images. Where The classification is done using a SVM	3
[70]	GHM	License Plate Character Recognition	32x16 binary image	BPNN(back propagation neural net)	Proposes a method of licence plate recognition using 2d gaussian Hermite moments. The classification is done using a BPNN.	34
	RM	Image analysis, image compression	40x40, 96x96, 256x256			

3

ORTHOGONAL MOMENTS

This chapter introduces orthogonal moments as features in a classification processing pipeline. The chapter includes the mathematical derivation and implementation of the orthogonal moments investigated in this research. Next to the implementation, the individual properties are analyzed and elucidated.

3.1. MOMENTS AS FEATURE

From the literature review presented in the previous chapter, the conclusion can be drawn that orthogonal moments have the potential to be a good feature extraction technique for micro-Doppler signatures. The variety of moments, as found in the literature, will be evaluated on the problem of micro-Doppler signature classification. This is intended to fill in the research gaps found. The moments are introduced mathematically in this chapter, after which an implementation is proposed.

MATHEMATICAL MOMENTS

Let us first define moments in general; moments are scalar quantities used to describe a certain *image function*. The principle of using moments to capture significant features from a function has been around for a long time. Examples of commonly used statistical moments are Expectation, Variance, and Skewness, among many others. The mathematical definition of a general moment is given by:

$$M_{pq} = \iint_D p_{pq}(x,y)\, f(x,y) \mathrm{d}x\, \mathrm{d}y \tag{3.1}$$

Where the function $f(x,y)$ is the *image function* and is defined as a piece-wise continuous real function of two variables defined on $D \subset \mathbb{R} \times \mathbb{R}$ with a finite nonzero integral. p and q are non-zero integers defining the moment order: $r = p+q$. The function $p_{qp}(x,y)$ is the polynomial basis defined on D. Mathematically, moments can be seen as a projection of the image function onto a polynomial basis, i.e. a change in basis.

ORTHOGONAL MOMENTS

The orthogonal moments are derived by taking a polynomial basis $p_{pq}(x,y)$ in equation 3.1 that carries the condition of orthogonality. This means that two different basis functions from the same set, p_{pq} and $p_{p'q'}$, are independent of each other. Mathematically, this independence can be proved by satisfying the condition:

$$\langle p_{pq}, p_{p'q'} \rangle = \iint_D p_{pq}(x,y) p^*_{p'q'}(x,y)\,dx\,dy = \delta_{pp'}\delta_{qq'} \tag{3.2}$$

Where δ_{ij} is the Kronecker delta function defined as

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}. \tag{3.3}$$

The orthogonal polynomial basis can be constructed over different domains and categorized into different families. The mathematical formulation of the different families is discussed in this chapter, with the derivation of all the moments used in this thesis. In extension to the derivation, an implementation will be proposed. For this research, the influence of the moment order is also analysed concerning the classification abilities. An assumption is made that this could be done with an analysis stopping at 50 as the highest order. Therefore, the proposed implementations of the variety of moments will be at least suitable for calculation till the 50th order.

3.2. CONTINUOUS CIRCULAR MOMENTS

The circular moments are the family of moments that are orthogonally defined on the unit disk. This section mathematically introduces the moments and proposes an implementation. This includes the implementation of the basis function and the mapping used to extract moments from a square image function.

JACOBI-FOURIER

The Jacobi-Fourier moments are the first discussed of the Jacobi polynomial-based group of functions. The JFM can be seen as the generalized case of the other special cases of Jacobi-based moments. Where for the radial basis function from the special cases can be derived from the JFM radial basis function by changing the parameters q and p. The Jacobi Fourier moments combine the Jacobi function radial function and the Fourier exponential factor. The JMF kernel function consists of two separable parts: The Fourier exponent function $\exp(jm\vartheta)$ and the deformed Jacobi polynomial $R_n^{(\mathrm{JFM})}(p,q,r)$ [47]. The JMF kernel can be defined as

$$P_{nm}(r,\vartheta) = R_n^{(\mathrm{JFM})}(p,q,r)\exp(jm\vartheta) \tag{3.4}$$

Where the orthogonality condition holds for the region $0 \leqslant r \leqslant 1, 0 \leqslant \vartheta \leqslant 2\pi$ The Jacobi polynomial and the Fourier exponentials are separable. Where the Fourier exponent $\exp(\mathrm{j}m\theta)$ is orthogonal [47]. This means that the radial function $R_n^{(\mathrm{JFM})}(p,q,r)$ should satisfy the condition:

$$\int_0^{2\pi}\int_0^1 P_{nm}(r,\vartheta)P_{kl}(r,\vartheta)\,r\,dr\,d\vartheta = \delta_{nk}\delta_{ml} \tag{3.5}$$

The Jacobi polynomial itself is defined as :

$$G_n(p,q,r) = \frac{n!(q-1)!}{(p+n-1)!}\sum_{s=0}^{n}(-1)^s \times \frac{(p+n+s-1)!}{(n-s)!s!(q+s-1)!}r^s \tag{3.6}$$

The polynomial should satisfy equation 3.5 on the interval $0 \leqslant r \leqslant 1$ resulting in

$$\int_0^1 G_n(p,q,r)G_m(p,q,r)w(p,q,r)\mathrm{d}r \quad = b_n(p,q)\delta_{nm} \tag{3.7}$$

Where b_n is a normalisation constant and $w(p,q,r)$ the weight function.

$$b_n = \frac{n![(q-1)!]^2(p-q+n)!}{(q+n-1)!(p+n-1)!(p+2n)}, \tag{3.8}$$

$$w(p,q,r) = (1-r)^{p-q}r^{q-1} \quad p-q>-1, \quad q>0. \tag{3.9}$$

And so can the radial polynomial can be defined as

$$R_n^{(\mathrm{JFM})}(p,q,r) = \sqrt{\frac{w(p,q,r)}{b(p,q)r}}G_n(p,q,r) \tag{3.10}$$

Or when substituted, resulting in

$$R_n^{(\mathrm{JFM})}(p,q,r) = \sqrt{\frac{r^{q-2}(1-r)^{p-q}(p+2n)\cdot\Gamma(q+n)\cdot n!}{2\pi\Gamma(p+n)\cdot\Gamma(p-q+n+1)}}\sum_{k=0}^{n}\frac{(-1)^k\Gamma(p+n+k)r^k}{k!(n-k)!\Gamma(q+k)} \tag{3.11}$$

This results in the Jacobi Fourier moment, defined as

$$\phi_{nm} = \int_0^{2\pi}\int_0^1 f(r,\vartheta)R_n^{(\mathrm{JFM})}(p,q,r)e^{-jm\vartheta}rdrd\vartheta \tag{3.12}$$

By changing the parameters q and p the properties of the radial basis function and, therefore, the moments will change. In [47], several combinations of q and p are evaluated, showing that for image representation, the optimal values are $p = 4$ $q = 4$. The error near the origin is the smallest in this combination compared to the other evaluated variants. The results on the reconstruction capabilities do not guarantee that this combination will still be optimal in classification. As the data of interest could be off the origin.
For some specific values of q and p, the JFM radial basis function will reduce the function to another type of basis function. The accompanying moments will still consist of the same Fourier angular kernel $exp(j*M*\theta)$, but the radial basis function will differ. These special cases are listed below:

- **Pseudo Jacobi-Fourier**: By setting the parameters of the Jacobi radial function equation 3.2 to

  $$p = 4 \text{ and } q = 3$$

  by filling in the values for p and q, the radial basis function is obtained

  $$R_n^{(\mathrm{PJFM})}(r) = \sqrt{\frac{(n+2)\left(r-r^2\right)}{\pi(n+3)(n+1)}}\sum_{k=0}^{n}\frac{(-1)^{n+k}(n+k+3)!r^k}{k!(n-k)!(k+2)!} \tag{3.13}$$

- **Zernike**: The Zernike radial basis function is obtained when setting the parameters q and p in equation 3.2 to

$$p = |m| + 1 \text{ and } q = |m| + 1$$

resulting in the radial basis function

$$R_{nm}^{(\mathrm{ZM})}(r) = \sqrt{\frac{n+1}{\pi}} \sum_{k=0}^{\frac{n-|m|}{2}} \frac{(-1)^k (n-k)! r^{n-2k}}{k! \left(\frac{n+|m|}{2} - k\right)! \left(\frac{n-|m|}{2} - k\right)!} \tag{3.14}$$

- **Pseudo-Zernike**: The Pseudo-Zernike radial basis function is obtained when setting the parameters q and p in the Jacobi-Fourier radial basis function (equation 3.2) to

$$p = 2|m| + 2 \text{ and } q = 2|m| + 2$$

Reduces to:

$$R_{nm}^{(\mathrm{PZM})}(r) = \sqrt{\frac{n+1}{\pi}} \sum_{k=0}^{n-|m|} \frac{(-1)^k (2n+1-k)! r^{n-k}}{k!(n+|m|+1-k)!(n-|m|-k)!} \tag{3.15}$$

- **Chebyshev-Fourier**: The Chebyshev-Fourier radial basis function is obtained when setting the parameters q and p in equation 3.2 to

$$p = 2 \text{ and } q = 1.5$$

Reduces to:

$$R_n^{(\mathrm{CHFM})}(r) = \frac{2}{\pi} \left(\frac{1-r}{r}\right)^{\frac{1}{4}} \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k (n-k)! (4r-2)^{n-2k}}{k!(n-2k)!} \tag{3.16}$$

- **Fourier-Mellin**: The Fourier-Mellin radial basis function is obtained when setting the parameters p and q in equation 3.2 to

$$p = 2 \text{ and } q = 2$$

Reduces to:

$$R_n^{(\mathrm{OFMM})}(r) = \sqrt{\frac{n+1}{\pi}} \sum_{k=0}^{n} \frac{(-1)^{n+k} (n+k+1)! r^k}{k!(n-k)!(k+1)!} \tag{3.17}$$

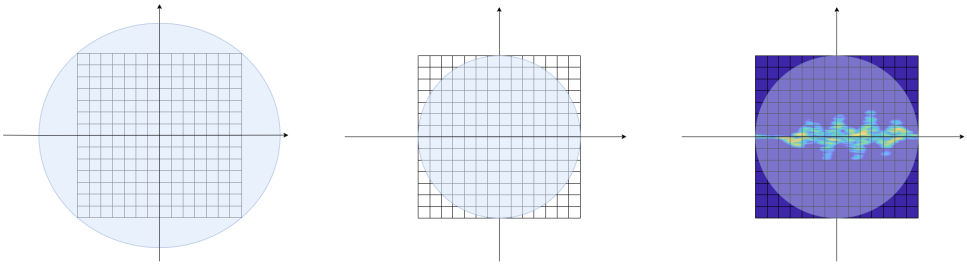
IMPLEMENTATION

The discussed moments are all defined on a unit circle. In general, radar data is represented in a square grid. This means there will be some loss in data or mapping of some kind to reduce this loss. In [38], different methods are proposed for accurately calculating square image functions using moments defined in a unit circle. The most straightforward solution would be to translate the Cartesian-defined image function to polar coordinates and draw a circular area with a radius r^{max} on the image defining the domain. Two cases are shown in figure 3.2; one is defined as a domain containing the whole

image function and a part of non-existing data, and one where some data of the image function is not contained in the domain of the moments. Due to the nature of radar data in Doppler-time representation, a case where the corners are not included would not introduce a large loss of information; this type of information loss is commonly referred to as 'Geometric Error'. An illustration of a cut out on human motion Doppler-time data is shown in figure 3.2c. Where the main part of interest lies in the domain of the calculated moment suggesting a low level of information loss, in the other case, non-existent data should be introduced to calculate the moments. This will, inevitably, introduce errors in representation capability for the calculated moments [38]. Therefore, the moments defined on the unit disk will be calculated as an incircle mapping.
After which the Cartesian-defined image function $f(x, y)$ will be mapped to the polar domain using;

$$r = \sqrt{(x^2 + y^2)} \quad \theta = \tan^{-1}(y/x)$$

A numerical approximation has to be made for the discretisation of the integral. Here a variant of the midpoint method is used because of its simplicity. As an implementation, a grid is created in polar space where $r \in [0, 1]$ and $\theta \in [0, 2\pi]$ all points outside this domain will be set to zero since they are not needed when calculating the moment. The number of pixels/data will determine the stepsize of the grid points N. The matrices containing r and θ can be directly filled in the radial polynomials as given previously in this section.



(a) r = 1, Circumcircle

(b) r = 1, Incircle

(c) r = 1, Incircle, placed on human motion spectrogram

Figure 3.1: Illustration of the square image function mapping

FAST AND ACCURATE COMPUTATION

The function, as described earlier, depends on the factorial. Higher-order factorials must be computed when entering higher-order moments, resulting in computationally expensive procedures. In addition, the software used to calculate the moments (MATLAB R2021a) cannot deal with such large factorials, causing the calculations to become numerically unstable. Therefore, the direct implementation of the moments is limited in order. As a solution, several papers have proposed [79][80] to use the recursion property of radial functions to make higher-order moments less computationally expensive. The recursive version of the Jacobi-Fourier radial function as given in the equation can be

calculated recursively using the method proposed in [79]. The radial basis function first has firstly to be rewritten in the form:

$$R_n(p,q,r) = \sqrt{(p+2n)(1-r)^{p-q}r^{q-2}}A_n(p,q)P_n(p,q,r) \tag{3.18}$$

with

$$P_n(p,q,r) = \sum_{k=0}^{n} \frac{(-1)^k(n+k+p-1)!r^k}{k!(q+k-1)!(n-k)!} \tag{3.19}$$

and

$$A_n(p,q) = \sqrt{\frac{n!(q+n-1)!}{(p+n-1)!(p-q+n)!}} \tag{3.20}$$

Where the initialization of the zeroth and first order for P_n:

$$P_0(p,q,r) = \frac{(p-1)!}{(q-1)!}, \quad P_1(p,q,r) = \frac{p!}{(q-1)!}\left(1 - \frac{(p+1)}{q}r\right), \tag{3.21}$$

The function itself is given by the following recursion

$$P_n(p,q,r) = (L_1 r + L_2)\,P_{(n-1)}(p,q,r) + L_3 P_{(n-2)}(p,q,r), \quad n = 2,3,\ldots,n_{\max} \tag{3.22}$$

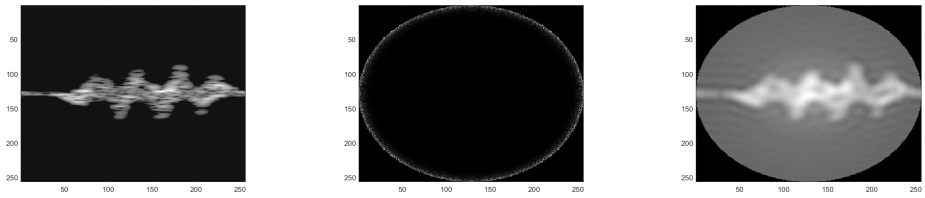
Where n_{max} defines the maximum order of the moments. The values of L_1, L_2 and L_3 are defined as:

$$\begin{aligned} L_1 &= -\frac{(2n+p-1)(2n+p-2)}{n(q+n-1)} \\ L_2 &= (p+2n-2) + \frac{(n-1)(q+n-2)}{(p+2n-3)}L_1 \\ L_3 &= \frac{(p+2n-4)(p+2n-3)}{2} + \frac{(q+n-3)(n-2)}{2}L_1 - (p+2n-4)L_2 \end{aligned} \tag{3.23}$$

Also for $A_n(p,q)$ it is possible to do the calculations recursively;

$$\begin{aligned} A_0(p,q) &= \frac{(q-1)!}{(p-1)!(p-q)!} \\ A_n(p,q) &= \frac{n(q+n-1)}{(p+n-1)(p-q+n)}A_{n-1}(p,q), n = 1,2,\ldots,n_{\max} \end{aligned} \tag{3.24}$$

The other Jacobi-based moments defined on a unit disk are mapped and computed the same way as done for the JFM. The recursive implementation is able to get accurate reconstruction results at higher-order moments. Also, the recursive method at lower orders seems more accurate when visually validating the moments. In Figure 3.2, the results can be seen for the direct and recursive approaches at order 50, illustrating the numerical instability of the direct approach.



(a) Original spectogram of human motion

(b) Order 50 reconstruction with direct implementation

(c) Order 50 reconstruction with recursive implementation

Figure 3.2: Example of radar spectrogram reconstruction of order 50 with Zernike moments, illustrating the instability of the direct implementation

EXPONENT-FOURIER

The Exponent-Fourier moments are part of the so-called harmonic function-based moments. The function is obtained using the exponent function $R_n(r) = A_n e^{jnr}$ as a radial basis function combined with the Fourier exponential function[48]. The orthogonality condition stated in equation 3.25 must hold for the region $0 \leq r \leq 1$, $0 \leq \theta \leq 2 * \pi$. From this condition, the value of A_n can be derived.

$$\int_0^1 A_n \exp(jnr) A_k \exp(-jkr) r dr = \delta_{nk} \tag{3.25}$$

When substituting $A_n = \frac{1}{\sqrt{2\pi r}}$ the final radial basis function is given as

$$R_n^{(\mathrm{EFM})}(r) = \frac{1}{\sqrt{2\pi r}} \exp(j2n\pi r) \tag{3.26}$$

The resulting moments are then calculated using 3.27, with $f(r,\theta)$ being the image function.

$$E_{nm} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 f(r,\theta) T_n^*(r) \exp(-jm\theta) r dr d\theta \tag{3.27}$$

The implementation, given by [38], is used to calculate this type of moment. This uses the algorithm proposed in [81] that uses an FFT-based calculation method in a polar coordinate system. Further, it provides the relation from this proposed; fast generic polar complex exponential transform (FGPCET) - to other harmonic function-based moments. Using this relation, a variety of harmonic-based moments can be calculated, e.g. Exponent-Fourier moments.

POLAR COMPLEX EXPONENTIAL TRANSFORM MOMENTS

As for the exponent Fourier, the Polar complex exponential transform moment (PCETM) is a harmonic function-based moment. Unlike the exponent Fourier, the radial basis function is numerically stable. The PCET is closely related to the discrete cosine transformation (DCT) and the discrete sine transformation (DST) [56]. The radial basis function is defined by the complex exponential:

$$R_n(r) = e^{j2n\pi r^2} \tag{3.28}$$

And the complete basis function as a combination of the complex exponential and the Fourier exponential:

$$H_{nl}(r,\theta) = R_n(r) * e^{il\theta} \tag{3.29}$$

Note that this radial basis function is numerically stable and satisfies the orthogonality condition:

$$\int_0^1 R_n(r)[R'_m(r)]^* r dr = \frac{1}{2}\delta_{nm'} \tag{3.30}$$

and so

$$\int_0^{2\pi}\int_0^1 H_{nl}(r,\theta)[H_{n'l'}(r,\theta)] * r dr d\theta = \pi\delta_{nn'}\delta_{ll'} \tag{3.31}$$

As can be seen from this equation, the basis function can be normalized, making the function orthonormal by:

$$\tilde{H}_{nl}(r,\theta) = \frac{1}{\sqrt{\pi}} H_{nl}(r,\theta) \tag{3.32}$$

The PCET-Moment is implemented similarly to the exponent Fourier, using an FFT-based computation technique. [81].

BESSEL-FOURIER

The Bessel-Fourier moments are based on the Bessel function; in this case, the Bessel function of the first kind is used. This function is defined as [68]

$$J_v(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(v+k+1)} \left(\frac{x}{2}\right)^{v+2k} \tag{3.33}$$

With v a real constant. The Bessel-Fourier moment is formed by taking the Bessel function in polar form and the exponent function [49]:

$$B_{nm} = \frac{1}{2\pi a_n}\int_0^{2\pi}\int_0^1 f(r,\theta) J_v(\lambda_n r)\exp(-jm\theta) r dr d\theta \tag{3.34}$$

Where $f(r,\theta)$ is a $N * M$ image function, a_n a normalization constant defined by $a_n = (J_{v+1}(\lambda_n))^2/2$, n the order and λ_n the n-th zero of the function $j_v(x)$. The orthogonality of this moment holds for $0 \leq r \leq 1$ can be shown by the orthogonality condition:

$$\int_0^1 r J_v(\lambda_n r) J_v(\lambda_k r) dr = a_n \delta_{nk} \tag{3.35}$$

With δ_{nk} the Kronecker symbol. The function can be rewritten in the same form as done with the Jacobi Fourier moments. Consisting of a radial basis function and the Fourier exponential function. Where the radial basis function can be derived from the equations above:

$$R_n^{(\mathrm{BFM})}(r) = \frac{1}{\sqrt{\pi} J_{v+1}(\lambda_n)} J_v(\lambda_n r) \tag{3.36}$$

The function is implemented using the same procedure as the Jacobi Fourier. MATLAB's first-order Bessel function is used to calculate the Bessel function. In addition, the implementation is used as provided in [38] for the moment calculation.

RADIAL HARMONIC FOURIER

The radial harmonic Fourier moments show similarities with the harmonic function-based moments discussed above. The basis function can be written again as a radial function $R_n^{(\mathrm{RHFM})}(r)$ in combination with an exponent function $exp(jm\theta)$. The basis function is defined as:

$$P_{nm}(r,\theta) = R_n^{(\mathrm{RHFM})}(r)\exp(jm\theta) \tag{3.37}$$

Satisfying the orthogonality condition in the range $0 < r < 1$:

$$\int_0^{2\pi}\int_0^1 P_{nm}(r,\theta)P_{kl}(r,\theta)r\,\mathrm{d}r\,\mathrm{d}\theta = \delta_{nmkl} \tag{3.38}$$

The radial function is defined as:

$$R_n^{(\mathrm{RHFM})}(r) = \begin{cases} \frac{1}{\sqrt{2\pi r}} & n = 0 \\ \sqrt{\frac{1}{\pi r}}\sin(\pi(n+1)r) & n > 0 \& n \text{ odd} \\ \sqrt{\frac{1}{\pi r}}\cos(\pi n r) & n > 0 \& n \text{ even} \end{cases} \tag{3.39}$$

The Radial harmonic Fourier-Moments are implemented similarly to the exponent Fourier, using an FFT- based computation technique. The implementation, given by [38], is used to calculate this type of moment.

3.3. CONTINUOUS MOMENTS DEFINED ON A SQUARE GRID

The continuous orthogonal moments defined on a square grid can be easily computed since no mapping of the image function is needed. The discretization can be done using a discrete linear approximation related to the midpoint method. The moments of this family will be first derived, and then the proposed implementation will be discussed. The implementation is just as for previously discussed moments based on recursive algorithms, in order to reduce the computational complexity.

GEGENBAUER

Gegenbauer polynomials, such as continuous Chebyshev and Legendre, are all special cases of Jacobi polynomials. The domain in which they are orthogonal is a square $D = \{(x,y) : x \in [-1,1], y \in [-1,1]\}$. The other continuous moments, Chebyshev and Legendre, can be formed from the Gegenbauer moments. The Gegenbauer basis function is given as:

$$G_n^{(\lambda)}(x) = \frac{(2\lambda)_n}{\left(\lambda+\frac{1}{2}\right)_n} P_n\left(\lambda-\frac{1}{2},\lambda-\frac{1}{2},x\right) \tag{3.40}$$

Where $P_n(\alpha,\beta)$ is the jacobi polynominal and λ a free parameter

$$P_n(\alpha,\beta)(x) = \frac{(\alpha+1)_n}{n!} \times {}_2F_1\left(\begin{matrix} -n, n+\alpha+\beta+1 \\ \alpha+1 \end{matrix} \middle| \frac{1-x}{2}\right) \tag{3.41}$$

Where the orthogonallity condition on domain $D \in [-1,1] \times [-1,1]$ holds:

$$\int_{-1}^{1} w_G(x,\lambda) G_n^{(\lambda)}(x) G_m^{(\lambda)}(x)\,dx = \rho_G(n,\lambda)\delta(n,m) \tag{3.42}$$

The weighting function and normalization function are derived from this condition:

$$w(x) = (1-x^2)^{a-\frac{1}{2}} \tag{3.43}$$

$$\rho_G(n,\lambda) = \frac{2\pi\Gamma(n+2\lambda)}{2^{2\lambda} n!(n+\lambda)[\Gamma(\lambda)]^2} \tag{3.44}$$

resulting in the final moment being defined as

$$M_{n,m}^{G} = \frac{1}{\rho_G(n,\lambda)\rho_G(m,\lambda)} \int_{-1}^{1}\int_{-1}^{1} f(x,y) G_n^{(\lambda)}(x) G_m^{(\lambda)}(y) w_G(x,\lambda) w_G(y,\lambda)\,dx\,dy \tag{3.45}$$

The function is discretized by mapping the image function $f(i,j)$ that exists in the domain defined by $1 < i,j < N$ to the region $-1 < x,y < 1$
The computational complexity becomes problematic when directly calculating the moments since the Gegenbauer polynomial depends on gamma, factorial, Pochhammer, and hypergeometric functions. Therefore, a recursive method is preferred. In [80], a fast computation method for Legendre moments based on recursion is proposed. The method is given as a pseudocode. This method can be rewritten for all moments in this family. First, the proposed algorithm is rewritten for Gegenbauer moments. From Gegenbauer implementations, the other cases can be formed by changing the parameter λ. The implementation is based on the recursion of the Gegenbauer polynomial[82] [83]:

$$G(n,a,x) = \frac{2x(n+a-1)}{n} * G(n-1,a,x) - \frac{n+2a-2}{n} G(n-2,a,x) \tag{3.46}$$

With the initialization $G(0,a,x) = 1$ and $G(1,a,x) = 2ax$. The pseudo-code for computing the GMs is given in Appendix A.

LEGENDRE MOMENTS

The Legendre two-dimensional moments of order (q + p) are given as in [84] as 3.48, where $f(x,y)$ is the image intensity function. The function is Orthogonal on the square [-1,1]x[-1,1] on a Cartesian grid. The Legendre is of the same family as the Gegenbauer polynomial and can be obtained by setting $\lambda = 0.5$. This results in a polynomial of order p described as:

$$P_p(x) \quad = \sum_{k=0}^{p} \left\{ (-1)^{\frac{p-k}{2}} \frac{1}{2^p} \frac{(p+k)!x^k}{\left(\frac{p-k}{2}\right)!\left(\frac{p+k}{2}\right)!k!} \right\}_{p-k=\text{cven}} \tag{3.47}$$

And the Legendre moment as

$$\begin{aligned} L_{pq} = & \frac{(2p+1)(2q+1)}{4} \int_{-1}^{1}\int_{-1}^{1} P_p(x) \\ & \times P_q(y) f(x,y)\mathrm{d}x\,\mathrm{d}y; \quad x,y \in [-1,1], \end{aligned} \tag{3.48}$$

The function is discretized by mapping the image function $f(i,j)$ that exists in the domain defined by $1 < i, j < N$ to the region $-1 < x, y < 1$. Resulting in

$$Ł_{pq} = \frac{(2p+1)(2q+1)}{(N-1)^2} \sum_{i=1}^{N} \sum_{j=1}^{N} P_n(x_i) P_m(y_j) f(i,j) \tag{3.49}$$

The moment is implemented in the same way as the Gegenbaur moment. Following the same procedure as proposed earlier. Where the implementation is based on the recurrence relation of the Legendre polynomials, which follows:

$$P_p(x) = \frac{(2p-1)xP_{p-1}(x) - (p-1)P_{p-2}(x)}{p} \tag{3.50}$$

and is initialized using $P_0(x) = 1$ $P_1(x) = x$

CONTINUOUS CHEBYSHEV

The continuous Chebyshev is a special case of Jacobi polynomials. It can be derived from the Gegenbauer polynomial by setting $\lambda = 0$ for the first kind and $\lambda = 1$ for the second kind of Chebyshev polynomials.

First kind: As for $\lambda = 0$ will introduce singularities in the Gegenbauer polynomial, the limit is taken, and the polynomial function is defined as:

$$T(n,x) = \begin{cases} \frac{1}{2} \lim_{\lambda \to 0} \frac{n+\lambda}{\lambda} G^{\lambda}(n,x) & \text{if } n \neq 0 \\ \lim_{\lambda \to 0} G^{\lambda}(0,x) = 1 & \text{if } n = 0 \end{cases} \tag{3.51}$$

This influences the recurrence relation:

$$T(n,x) = 2xT(n-1,x) - T(n-2,x) \tag{3.52}$$

with initialization $T(0,x) = 1$ and $T(1,x) = x$ and the weighting function and normalization function:

$$w(x) = \frac{1}{\sqrt{1-x^2}} \tag{3.53}$$

$$\rho = \begin{cases} \pi \text{ if } n = m = 0 \\ \frac{\pi}{2} \text{ if } n = m \neq 0 \end{cases} \tag{3.54}$$

SECOND KIND:

For this research, the Gegenbauer polynomial, as implemented before, is taken to calculate the second Chebyshev moments where:

$$U(n,x) = G(n,1,x) \tag{3.55}$$

GAUSSIAN-HERMITE

The pth-order Hermite polynomial is defined as

$$H_p(x) = (-1)^p \exp\left(x^2\right)\left(d^p/dx^p\right)\exp\left(-x^2\right) \tag{3.56}$$

When written as a series, the function can be described as:

$$H_p(x) = \sum_{k=0}^{p/2} \frac{(-1)^k p!}{k!(p-2k)!}(2x)^{p-2k} \tag{3.57}$$

Witch is a discrete function and can easily be implemented using the recursive properties of the factorial term:

$$H_{p+1}(x) = 2xH_p(x) - 2pH_{p-1}(x), \quad \text{for } p \geq 1 \tag{3.58}$$

with $H_0(x) = 1$ and $H_1(x) = 2x$
The orthogonality condition is given as follows:

$$\int_{-\infty}^{\infty} \exp\left(-x^2\right) H_p(x) H_q(x) dx = 2^p p!\sqrt{\pi}\delta_{pq} \tag{3.59}$$

From this condition, the weighted Hermite polynomial can be derived. This is done by modulating the Hermite function with a Gaussian function, making the function orthonormal and resulting in the Gaussian-Hermite function.

$$\widehat{H}_p(x;\sigma) = \left(2^p p!\sqrt{\pi}\sigma\right)^{-1/2} \exp\left(-x^2/2\sigma^2\right) H_p(x/\sigma) \tag{3.60}$$

$$\int_{-\infty}^{\infty} \widehat{H}_p(x;\sigma)\widehat{H}_q(x;\sigma)dx = \delta_{pq}. \tag{3.61}$$

Note that the function depends on the term σ that defines the variance of the Gaussian function. In [85], an optimization of σ_n, concerning the image reconstruction, is performed. This resulted in the relation given:

$$\sigma_n = \begin{cases} 0.9n^{-0.52} & n \geq 1 \\ 1.0 & n = 0 \end{cases} \tag{3.62}$$

The assumption is made that the optimal σ_n for image representation would also be optimal for a classification problem. The resulting moments can then be calculated using the following:

$$\eta_{pq} = \iint f(x,y)\widehat{H}_p(x,\sigma)\widehat{H}_q(y;\sigma)dxdy \tag{3.63}$$

Note that the Gaussian function is still continuous. Therefore, the image function will be mapped to the square domain of [-1,1]x[-1,1], at which the function is orthogonal, and at these locations, the function is evaluated. Resulting in a sample size equal to the size of the image.

$$\eta_{pq} = \sum_{y=-1}^{1}\sum_{x=-1}^{1} f(x,y)\widehat{H}_p(x,\sigma)\widehat{H}_q(y;\sigma)dxdy \tag{3.64}$$

3.4. DISCRETE MOMENTS

The discrete orthogonal moments defined on a square grid can be computed without the need to map the image function. The moments of this family will be first mathematically derived, and then the proposed implementation will be discussed. The implementation is as for previously discussed moments based on recursive algorithms to reduce computational complexity.

RACAH MOMENTS

Racah polynomials are hypergeometric polynomials defined on a discrete Cartesian grid. As described above, this has the advantage that there will be no calculation errors in discretizing the basis function. Also, the image function does not need to be mapped; i.e., the image is defended on a square Cartesian domain. [54] The Racah polynomial can be seen as a more general form of the other orthogonal functions in this family, e.g., Hahn, Dual Hahn, Meixner, krawtchouk, Chalier, Hermite. All orthogonal polynomials given can be derived from the Racah polynomial by changing their limits [57]. The relations of Hypergeometric Orthogonal polynomials are described by the askey scheme as represented in the figure. This also described the relations between the discrete orthogonal functions and their continuous twins, e.g., Wilson, continuous Dual Hahn, continuous Hahn, Meixner - Pollaczek, Jacobi, and Laguerre.

The Racah polynomial is defined on a non-uniform lattice, contrary to Krawtcouk, Tchebichef and Hahn. For this research, the lattice is quadratically defined as $x(s) = s(s+1))$. The nth order Racah polynomial $u_n^{(\alpha,\beta)}(s,a,b)$, defined on an NxN image, is represented by the equation:

$$u_n^{(\alpha,\beta)}(s,a,b) = \frac{1}{n!}(a-b+1)_n(\beta+1)_n(a+b+\alpha+1)_n \times {}_4F_3\left(\begin{matrix} -n, \alpha+\beta+n+1, a-s, a+s+1 \\ \beta+1, a+1-b, a+b+\alpha+1 \end{matrix}; \right), \tag{3.65}$$

$n = 0,1,\ldots,L-1, s = a, a+1, \ldots, b-1$

Where $(u)_k$ is the Pochhammer symbol and ${}_4F_3(\cdot)$ a hypergeometric function. Furthermore, the constraints 3.66 apply; note that changing the limits of these constraints changes the region of the image function that is mapped on the basis function, i.e. the region of interest (ROI) will change.

$$-1/2 < a < b, \quad \alpha > -1, \quad -1 < \beta < 2a+1, \quad b = a+N \tag{3.66}$$

Equation 3.65 satisfies its orthogonal property as described in equation 3.67

$$\sum_{s=a}^{b-1} u_n^{(\alpha,\beta)}(s,a,b)\, u_m^{(\alpha,\beta)}(s,a,b)\, \rho(s) \left[\Delta x\left(s-\frac{1}{2}\right)\right] = \delta_{nm} d_n^2, \quad n,m = 0,1,\ldots,L-1 \tag{3.67}$$

with

$$d_n^2 = \frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)\Gamma(b-a+\alpha+\beta+n+1)\Gamma(a+b+\alpha+n+1)}{(\alpha+\beta+2n+1)n!(b-a-n-1)!\Gamma(\alpha+\beta+n+1)\Gamma(a+b-\beta-n)}, \tag{3.68}$$
$$n = 0,1,\ldots,L-1.$$

To reach numerical stability, the weighting function described in Equation 3.69 is introduced.

$$\rho(s) = \frac{\Gamma(a+s+1)\Gamma(s-a+\beta+1)\Gamma(b+\alpha-s)\Gamma(b+\alpha+s+1)}{\Gamma(a-\beta+s+1)\Gamma(s-a+1)\Gamma(b-s)\Gamma(b+s+1)} \tag{3.69}$$

Resulting in the weighted Racah polynomial:

$$\hat{u}_n^{(\alpha,\beta)}(s,a,b) = u_n^{(\alpha,\beta)}(s,a,b)\sqrt{\frac{\rho(s)}{d_n^2}\left[\Delta x\left(s-\frac{1}{2}\right)\right]}, \quad n=0,1,\ldots,L-1 \tag{3.70}$$

Then the orthogonality condition reduces to;

$$\sum_{s=a}^{b-1} \hat{u}_n^{(\alpha,\beta)}(s,a,b)\,\hat{u}_m^{(\alpha,\beta)}(s,a,b) = \delta_{nm}, \quad n,m=0,1,\ldots,L-1., \tag{3.71}$$

In [54] these normalized orthogonal Racahn functions are used to create moments and are defined, for a NxN domain and a (n+m)th order, as:

$$U_{nm} = \sum_{s=a}^{b-1}\sum_{t=a}^{b-1} \hat{u}_n^{(\alpha,\beta)}(s,a,b)\,\hat{u}_m^{(\alpha,\beta)}(t,a,b)\,f(s,t) \quad n,m=0,1,\ldots,L-1 \tag{3.72}$$

IMPLEMENTATION

The Racah polynomials are dependent on a higher-order hypergeometric function which is a computationally expensive procedure to implement. Furthermore, the weighting functions $\rho(s)$ and the square norm d_n^2 are dependent on the gamma functions, which also require a lot of computational power to implement directly. In [54], a recursive method is proposed with respect to n to calculate Racah polynomials. Where the following recursive relation is proposed:

$$A_n\hat{u}_n^{(\alpha,\beta)}(s,a,b) = B_n\frac{d_{n-1}}{d_n}\hat{u}_{n-1}^{(\alpha,\beta)}(s,a,b) + C_n\frac{d_{n-2}}{d_n}\hat{u}_{n-2}^{(\alpha,\beta)}(s,a,b) \tag{3.73}$$

Where :

$$A_n = \frac{n(\alpha+\beta+n)}{(\alpha+\beta+2n-1)(\alpha+\beta+2n)} \tag{3.74}$$

$$\begin{aligned} B_n =& x - \frac{a^2+b^2+(a-\beta)^2+(b+\alpha)^2-2}{4} + \frac{(\alpha+\beta+2n-2)(\alpha+\beta+2n)}{8} \\ &- \frac{\left(\beta^2-\alpha^2\right)\left[(b+\alpha/2)^2-(a-\beta/2)^2\right]}{2(\alpha+\beta+2n-2)(\alpha+\beta+2n)} \end{aligned} \tag{3.75}$$

$$\begin{aligned} C_n =& -\frac{(\alpha+n-1)(\beta+n-1)}{(\alpha+\beta+2n-2)(\alpha+\beta+2n-1)}\left[\left(a+b+\frac{\alpha-\beta}{2}\right)^2-\left(n-1+\frac{\alpha+\beta}{2}\right)^2\right] \\ &\times\left[\left(b-a+\frac{\alpha+\beta}{2}\right)^2-\left(n-1+\frac{\alpha+\beta}{2}\right)^2\right] \end{aligned} \tag{3.76}$$

and the initialization is done using:

$$\hat{u}_0^{(\alpha,\beta)}(s,a,b) = \sqrt{\frac{\rho(s)}{d_0^2}\left[\Delta x\left(s-\frac{1}{2}\right)\right]} \tag{3.77}$$

$$\hat{u}_1^{(\alpha,\beta)}(s,a,b) = -\frac{1}{\rho(s)}\frac{\rho_1(s)-\rho_1(s-1)}{x(s+1/2)-x(s-1/2)}\sqrt{\frac{\rho(s)}{d_1^2}\left[\Delta x\left(s-\frac{1}{2}\right)\right]} \tag{3.78}$$

With $\Delta x\left(s-\frac{1}{2}\right) = 2s+1 > 0$ for $s > -1/2$. This implementation greatly reduces the computational complexity of the Racah polynomial although the initialization terms are still dependent on equations 3.69 and 3.68, which on themselves are dependent on gamma functions. Therefore these functions will be rewritten in a recursive manner. The same procedure is used as proposed in [86] for calculating Dual Hahn polynomials but rewritten for the Racah polynomial. $\rho(s)$ can be rewritten using property: $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$ and $\Gamma(1) = 1$, as the recursive function:

$$\rho(s) = \frac{\Gamma(a+s+1)\Gamma(s-a+\beta+1)\Gamma(b+\alpha-s)\Gamma(b+\alpha+s+1)}{\Gamma(a-\beta+s+1)\Gamma(s-a+1)\Gamma(b-s)\Gamma(b+s+1)}$$

$$\rho(s) = \frac{(a+s)\Gamma(a+s)(s-a+\beta)\Gamma(s-a+\beta)(b+\alpha-s)^{-1}\Gamma(b+\alpha-s-1)(b+\alpha+s)\Gamma(b+\alpha+s+1)}{(a-\beta+s)\Gamma(a-\beta+s+1)(s-a)\Gamma(s-a+1)(b-s)^{-}1\Gamma(b-s-1)(b+s)\Gamma(b+s+1)}$$

$$\rho(s) = \rho(s-1) * \frac{(a+s)(s-a+\beta)(b+\alpha+s)(b-s-1)}{(a-\beta+s)(s-a)(b+s)(b+\alpha-s)}$$

For initialization $s = a$:

$$\rho(a) = \frac{\Gamma(2a+1)\Gamma(\beta+1)\Gamma(b+\alpha-a)\Gamma(b+\alpha+a+1)}{\Gamma(2a-\beta+1)\Gamma(b-a)\Gamma(b+a+1)} \tag{3.79}$$

Note that this initialization is not stable when calculating the higher values of the gamma function. Therefore, a second recursive calculation is proposed with respect to b; since $b = a + image\_size$.

$$\rho(b) = \frac{\Gamma(b+\alpha-a)\Gamma(b+\alpha+a+1)}{\Gamma(b-a)\Gamma(b+a+1)} * \frac{\Gamma(2a+1)\Gamma(\beta+1)}{\Gamma(2a-\beta+1)} \tag{3.80}$$

Initialization $b_i = a+1$

$$\rho(b_i = a+1) = \frac{\Gamma(\alpha+1)\Gamma(\alpha+2a+2)}{(2a+1)} * \frac{\Gamma(\beta+1)}{\Gamma(2a-\beta+1)} \tag{3.81}$$

$$\rho(b_i) = \frac{(b_i+\alpha-a-1)(b_i+\alpha+a)}{(b_i-a-1)(b_i+a)} * \rho(b_i-1) \tag{3.82}$$

for $b_i = a+2, a+3, \ldots, b$ The same procedure can be done for d_n^2 resulting in :

$$d_n^2 = \frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)\Gamma(b-a+\alpha+\beta+n+1)\Gamma(a+b+\alpha+n+1)}{(\alpha+\beta+2n+1)n!(b-a-n-1)!\Gamma(\alpha+\beta+n+1)\Gamma(a+b-\beta-n)} \tag{3.83}$$

$$(\alpha+\beta+2n+1)*d_n^2 = d_{n-1}^2 * \frac{(\alpha+n)(\beta+n)(b-a+\alpha+\beta+n)(a+b+\alpha+n)(b-a-n)(a+b-\beta-n)}{n(\alpha+\beta+n)} \tag{3.84}$$

With initialization $n = 0$:

$$d_0^2 = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)\Gamma(b-a+\alpha+\beta+1)\Gamma(a+b+\alpha+1)}{(\alpha+\beta+1)(b-a-1)!\Gamma(\alpha+\beta+1)\Gamma(a+b-\beta)} \tag{3.85}$$

Again, due to the complexity of the initialization, a second iterative solution is proposed, where the iteration is performed with respect to b:

$$d_{0,b}^2 = d_{0,b-1}^2 \frac{(b-a+\alpha+\beta)(a+b+\alpha)}{(b-a-1)(a+b-\beta-1)} \tag{3.86}$$

initialization for $b = a+1$

$$d_0^2 = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)\Gamma(\alpha+\beta+2)\Gamma(2a+\alpha+2)}{(\alpha+\beta+1)\Gamma(\alpha+\beta+1)\Gamma(2a+1-\beta)} \tag{3.87}$$

Using this method will give an accurate result for d_n^2 till the 12th order and then numbers will become too large for Matlab. To be able to go into higher orders the terms $\frac{d_{n-1}^2}{d_n^2}$ and $\frac{d_{n-2}^2}{d_n^2}$ from equation 3.73 recursively calculated using previous derivations. Rewriting equation 3.84.

$$\begin{aligned}&\frac{(\alpha+\beta+2n+1)}{(\alpha+\beta+2(n-1)+1)} * \frac{d_n^2}{d_{n-1}^2} = \\ &\frac{(\alpha+n)(\beta+n)(b-a+\alpha+\beta+n)(a+b+\alpha+n)(b-a-n)(a+b-\beta-n)}{n(\alpha+\beta+n)}\end{aligned} \tag{3.88}$$

and for $\frac{d_{n-2}^2}{d_n^2}$:

$$\begin{aligned}&\frac{(\alpha+\beta+2n+1)}{(\alpha+\beta+2(n-2)+1)} * \frac{d_n^2}{d_{n-2}^2} = \\ &\frac{(\alpha+n)(\beta+n)(b-a+\alpha+\beta+n)(a+b+\alpha+n)(b-a-n)(a+b-\beta-n)}{n(\alpha+\beta+n)} * \\ &\frac{(\alpha+n-1)(\beta+n-1)(b-a+\alpha+\beta+n-1)(a+b+\alpha+n-1)(b-a-n-1)(a+b-\beta-n-1)}{(n-1)(\alpha+\beta+n-1)}\end{aligned} \tag{3.89}$$

This method gives accurate results and can be seen in the values in Table 3.1.

EVALUATION PARAMETERS

Equation 3.72 has the ability to adjust different parameters. Each parameter will determine how the basis functions which define the resulting moment will behave. To be able to determine the best possible realization of the Racah moment for the purpose of classifying micro-Doppler signatures, an analysis is performed. This is done by illustrating

n	d recursive	d direct	d/(d-1) recursive	d/(d-1) direct	d/(d-1) fraction recursive
10	1.19602903500493e+79	1.19602903500499e+79	41143120.7478261	41143120.7478247	41098431.1688312
15	1.38132936819189e+117	-	40460155.4823529	-	40696056.6532258
20	1.38456519804585e+155	-	38716260.4389006	-	39132489.8587933
50	-	-	-	-	18179296.1144854
80	-	-	-	-	156492.302853998
	d/(d-2) recursive	**d/(d-2) direct**	**d/(d-2) fraction recursive**		
10	1.69168446014574e+15	-	1.68383447049679e+15		
15	1.62581559364224e+15	-	1.66395097380975e+15		
20	1.48166935679476e+15	-	1.54639831758786e+15		
50	-	-	345726449017378		
80	-	-	48371242043.0769		

Table 3.1: Table containing the calculated parameter d values for the different calculation methods.

the influence of each parameter on the basis function and then optimizing it to actual micro-Doppler signatures.
The free variables in the basis function are a, b, α, β. Variables are constrained by the equation 3.66. The variable a can be seen as the starting point of the region of interest. Variations in the values of a and b can be seen in Figure 3.3. Note that the region size will not change due to the earlier-mentioned constraints. In the case of an image, the horizontal and vertical basis functions are defined by the same values of a and b. Examples of the influence on image representations can be seen in Figures 3.4 and 3.5.

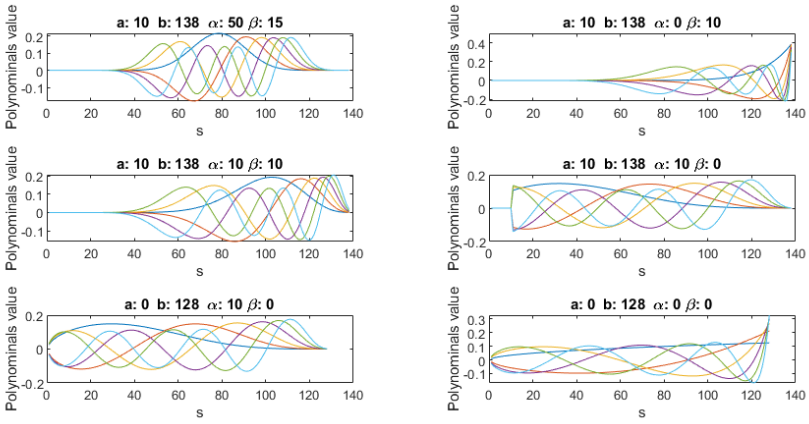


Figure 3.3: Racah basis functions of order 1 to 5 under different values of the parameters a, b, α, and β, the s axis represents the non-uniform lattice.

Where also variations of alpha and beta are shown. It is clearly visible that the value of a has a shifting property that places emphasis on the part at pixel locations larger than a. The influence of alpha and beta is also visible. Whereas β increases, the basis function will contract and shift [54]. This property also holds for α. This means that the distribution of the zeros depends on the values of α and β [58]. This property can

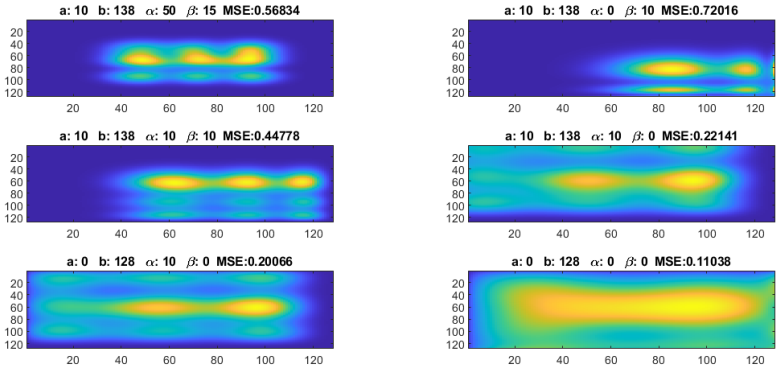


Figure 3.4: Example of reconstructed images with different parameters in Racah basis function, order 5

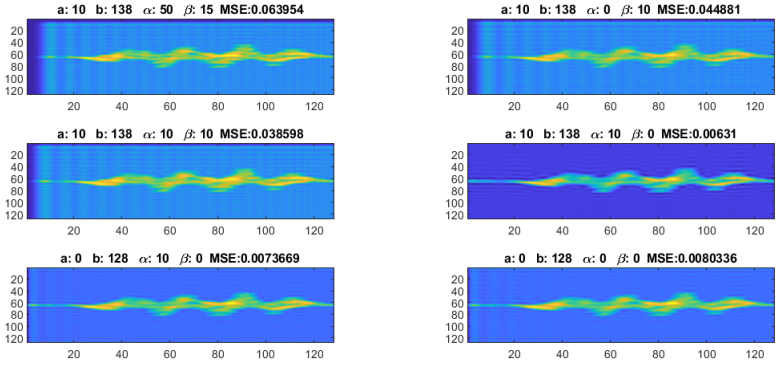


Figure 3.5: Example of reconstructed images with different parameters in Racah basis function, order 50

potentially increase classification rates, as the distribution of zeros illustrates the region of emphasis of a basis function. When taking a simple Micro-Dopler data set, Dopnet, the information is spread across the middle, vertically, and horizontally. Therefore, a distribution is chosen where the basis functions have a high zero density in the middle region of our image function. This results in a change in emphasis to the region with the most relevant information. Various parameter variations are shown in Figures 3.4 and 3.5. In these figures, zero distributions can be easily observed. These parameter settings are used in Chapter 6, where they are evaluated as individual moments and applied to different datasets.

DUAL HAHN

Compared to the Hahn moment, the dual Hahn moment is defined on a non-uniform grid as $x(s) = s(s+1)$. The Dual Hahn moment of order n and size N is defined by the

hypergeometric function[54]:

$$H_n{}^{(a,b,c)}(s) = \frac{(a-b+1)_n(a+c+1)_n}{n!} \\ {}_3F_2(-n, a-s, a+s+1; a-b+1, a+c+1; 1) \tag{3.90}$$

Where $n = 0,1,...N-1$, $s = a, a+1, .., b-1$, $(u)_k$ the Pochhammer symbol[1] and ${}_3F_2$ a hypergeometric function[2]. The orthogonality condition is given as follows

$$\sum_{s=a}^{b-1} H_n^{(a,b,c)}(s) H_m^{(a,b,c)}(s) \rho(s) \left[\Delta x\left(s-\frac{1}{2}\right)\right] = \delta_{nm} d_n^2 \tag{3.91}$$

With $n, m = 0,1,\ldots,N-1$ and $\Delta x\left(s-\frac{1}{2}\right) = 2s+1 > 0$ for $s > -1/2$ and

$$\rho(s) = \frac{\Gamma(a+s+1)\Gamma(c+s+1)}{\Gamma(s-a+1)\Gamma(b-s)\Gamma(b+s+1)\Gamma(s-c+1)} \tag{3.92}$$

Where the constraint holds :

$$-\frac{1}{2} < a < b,\ |c| < 1+a,\ b = a+N \tag{3.93}$$

and

$$d_n^2 = \frac{\Gamma(a+c+n+1)}{n!(b-a-n-1)!\Gamma(b-c-n)} \tag{3.94}$$

Note that the Dual Hahn polynomial, as described before, is not numerically stable. Therefore, will the function be normalized using the relation from equation 3.91 resulting in:

$$\widetilde{H}_n^{(a,b,c)}(s) = H_n^{(a,b,c)}(s) \sqrt{\frac{\rho(s)}{d_n^2}\left[\Delta x\left(s-\frac{1}{2}\right)\right]} \tag{3.95}$$

The resulting weighted dual Hahn moments for an image function $f(s,t)$ can then be calculated using the following:

$$W_{nm} = \sum_{s=a}^{b-1}\sum_{t=a}^{b-1} \widetilde{H}_n{}^{(a,b,c)}(s) \widetilde{H}_m{}^{(a,b,c)}(t) f(s,t),\ \ n,m = 0,1,....N-1 \tag{3.96}$$

IMPLEMENTATION

The equation for the Dual Hahn polynomial is computationally complex due to its dependence on Gamma functions and factorial operators. Due to this complexity, MATLAB cannot compute the high-order moments. Therefore, a recursive implementation is preferred. In [86], a recursive method is proposed. This follows the recursion:

$$\widetilde{H}_n^{(a,b,c)}(s) = A\sqrt{B} \times \widetilde{H}_{n-1}{}^{(a,b,c)}(s) + C\sqrt{D} \times \widetilde{H}_{n-2}{}^{(a,b,c)}(s) \tag{3.97}$$

[1] $(a)_k = a(a+1)(a+2)\ldots(a+k-1) = \frac{\Gamma(a+k)}{\Gamma(a)}$

[2] ${}_3F_2(a_1,a_2,a_3;b_1,b_2;z) = \sum_{k=0}^{\infty} \frac{(a_1)_k(a_2)_k(a_3)_k}{(b_1)_k(b_2)_k} \cdot \frac{z^k}{k!}$

With

$$\begin{array}{ll} A & \frac{1}{n}\left[s(s+1)-ab+ac-bc-(b-a-c-1)(2n-1)+2(n-1)^2\right] \\ B & \frac{n}{(a+c+n)(b-a-n)(b-c-n)} \\ C & -\frac{1}{n}(a+c+n-1)(b-a-n+1)(b-c-n+1) \\ D & \frac{n(n-1)}{(a+c+n)(a+c+n-1)(b-a-n+1)(b-a-n)(b-c-n+1)(b-c-n)} \end{array} \tag{3.98}$$

And initialization:

$$\begin{aligned} \widetilde{H}_0{}^{(a,b,c)}(s) &= \sqrt{\frac{\rho(s)}{d_0^2}\left[\Delta x\left(s-\frac{1}{2}\right)\right]} \\ \widetilde{H}_1{}^{(a,b,c)}(s) &= \frac{(s-a)(s+b)(s-c)+(s+a+1)(s+c+1)(s-b+1)}{(2s+1)} \\ &\quad \times \sqrt{\frac{\rho(s)}{d_1^2}\left[\Delta x\left(s-\frac{1}{2}\right)\right]} \end{aligned} \tag{3.99}$$

Note that the terms $\rho(s)$ and n^2 still depend on the Gamma function and factorial term. Therefore, for the calculation of the initialization of the polynomial, another recursion over n is introduced:

$$\widetilde{H}_0{}^{(a,b,c)}(s) = \sqrt{\frac{(a+s)(c+s)(b-s)(2s+1)}{(s-a)(b+s)(s-c)(2s-1)}}\,\widetilde{H}_0{}^{(a,b,c)}(s-1) \tag{3.100}$$

Where $H_0{}^{(a,b,c)}(a)$ is rewritten using the $ln\Gamma(*)$ function. This function is chosen to reduce the size of the output of the gamma function and so reduce the computational complexity. This results in :

$$H_0{}^{(a,b,c)}(a) = \sqrt{2a+1} * e^{\frac{E-F}{2}} \tag{3.101}$$

with

$$E = ln\Gamma(2a+1) + ln\Gamma(b+c) \tag{3.102}$$

$$F = ln\Gamma(b+a+1) + ln\Gamma(a-c+1) \tag{3.103}$$

And the second initialization :

$$H_1{}^{(a,b,c)}(s) = A * \frac{(s-a)(s+b)(s-c)+(s+a+1)(s+c+1)(s-b+1)}{(2s+1)} * H_0{}^{(a,b,c)}(s) \tag{3.104}$$

with

$$A = \sqrt{\frac{1}{(a+c+1)(b-c-1)(b-a-1)}} \tag{3.105}$$

EVALUATION PARAMETERS

The Dual Hahn polynomial, as described in equation 3.90 has free parameters a, b and c. As can be deduced from the region where s is defined, the parameter a states the beginning location of the polynomial evaluation on the image function, i.e. the starting point. Because the constraint 3.93, $b = a + N$ will determine the largest point of s that will be evaluated. Therefore, a and b together determine the location of the window of size N of the polynomial function that is evaluated. This can be observed in figure 3.6. Note that the size of the reconstructed image does not vary as illustrated in figure 3.7.
The influence of c is shown in figure 3.6 where the values of b and a are kept constant, and three cases of c are plotted for the first ten orders of the polynomial. This shows that when increasing c, the polynomial will shift from left to right and the other way around. The three cases of parameter selection, as illustrated in the figures presented, will be tested in the classification pipeline to get an idea of the influence of the varying values for c on the classification results.

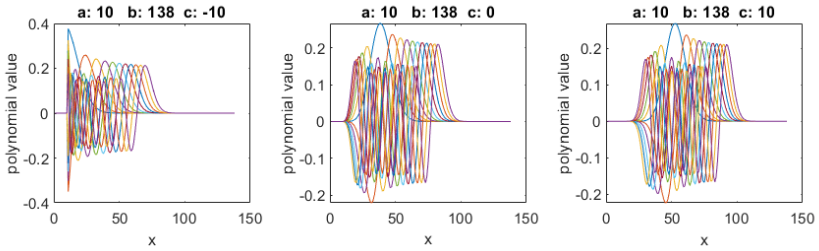


Figure 3.6: 1D representation of 10th order Dual-Hahn moment with varying values of the parameter c

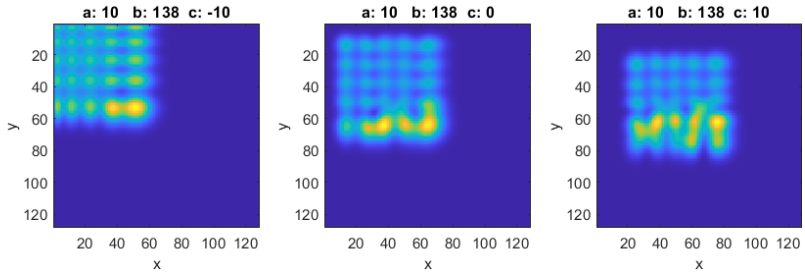


Figure 3.7: Example of image reconstruction of human motion spectrogram with 10th order Dual-Hahn moment and varying values of the parameter c

HAHN

From the Racah polynomial given in Section 3.4 the Hahn polynomial, $Q_n(x;\alpha,\beta,N)$, can be formed by setting $\gamma + 1 = -N$ and let $\delta \to \infty$ in $R_n(x(x+\gamma+\delta+1);\alpha,\beta,\gamma,\delta)$[53]. Note that in contrast to the Racah and Dual- Hahn moments, the Hahn moment is defined on a uniform lattice. This results in the Hahn polynomial as described in

$$\mathscr{H}_n^{\alpha,\beta}(x;N) = \frac{(-1)^n(\beta+1)_n(N-n)_n}{n!}\,{}_3F_2\left(\begin{matrix} -n,-x,n+1+\alpha+\beta \\ \beta+1,1-N \end{matrix}\middle|\, 1\right) \tag{3.106}$$

Here $(u)_k$ the Pochhammer symbol and ${}_3F_2$ a hypergeometric function. To establish numerical stability, the same step as for Racah is applied. Here the polynomial is normalized as in 3.107, resulting in the weighted Hahn polynomial.

$$\tilde{\mathscr{H}}_n^{\alpha,\beta}(x;N) = \mathscr{H}_n^{\alpha,\beta}(x;N)\sqrt{\frac{\omega_{\mathscr{H}}}{\rho_{\mathscr{H}}}}. \tag{3.107}$$

With the weight function defined as

$$w(x;\alpha,\beta,N) = \left(\begin{matrix} \alpha+x \\ x \end{matrix}\right)\left(\begin{matrix} \beta+N-x \\ N-x \end{matrix}\right) \tag{3.108}$$

and Norm

$$\rho(n;\alpha,\beta,N) = \frac{(-1)^n(n+\alpha+\beta+1)_{N+1}(\beta+1)_n n!}{(2n+\alpha+\beta+1)(\alpha+1)_n(-N)_n N!} \tag{3.109}$$

The Hahn polynomial has the ability to change its region of emphasis, i.e., it can change its ROI. This is done by changing the parameters $\alpha_1 = p_1 t_1, \beta_1 = (1-p_1)t_1$, and $\alpha_2 = p_2 t_2, \beta_2 = (1-p_2)t_2$.[53] In the case of $\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 0$ the Moment is set into global mode i.e. capturing the whole image. In case of $\{\alpha_1, \beta_1, \alpha_2, \beta_2\} > 0$ it will extract local features.

IMPLEMENTATION

As with the Racah moments and Dual Hahn moments, the calculation of the Hahn moments will depend on the hypergeometric function and factorial functions. Therefore, the order of the moments that can be computed in e.g. MATLAB will be limited. There are several techniques for rewriting equation 3.106 in recursive ways. In [87], a method is proposed to calculate the Hahn moments in a recursive and fast manner. For this research, the implementation of recursive computation over x is chosen and is implemented as follows:
-Initialization:

$$\tilde{h}_0^{(\alpha,\beta)}(x) = \begin{cases} \sqrt{\frac{(N-x)(\beta+x)}{x(N+\alpha-x)}}\,\tilde{h}_0^{(\alpha,\beta)}(x-1); & x>0 \\ \sqrt{\frac{(\alpha+\beta+1)\Gamma(\alpha+1+\beta)\Gamma(N+\alpha)}{\Gamma(\alpha+1)\Gamma(N+\alpha+\beta+1)}} & ;x=0 \end{cases} \tag{3.110}$$

-Recursive method to lose gamma functions :

$$\tilde{h}_n^{(\alpha,\beta)}(0) = -\sqrt{\frac{(\alpha+\beta+2n+1)(\alpha+\beta+n)(N-n)(n+\beta)}{n(\alpha+n)(\alpha+\beta+N+n)(\alpha+\beta+2n-1)}}\,\tilde{h}_{n-1}^{(\alpha,\beta)}(0) \tag{3.111}$$

$$\widetilde{h}_n^{(\alpha,\beta)}(1) = ((n+\beta+1)(N-n-1)-n(N+\alpha-1))\sqrt{\frac{1}{(\beta+1)(N-1)(N+\alpha-1)}}\widetilde{h}_n^{(\alpha,\beta)}(0) \quad (3.112)$$

-The iteration loop itself is then defined as:

$$\widetilde{h}_n^{(\alpha,\beta)}(x) = A \times \widetilde{h}_n^{(\alpha,\beta)}(x-1) + B \times \widetilde{h}_n^{(\alpha,\beta)}(x-2) \quad (3.113)$$

with

$$A = \frac{2(x-1)(N+\alpha-x+1)+(\beta+1)(N-1)-(\alpha+\beta+2)(x-1)-n(\alpha+\beta+n+1)}{(x-1)(N+\alpha-x+1)+(\beta+1)(N-1)-(\alpha+\beta+2)(x-1)}\sqrt{\frac{(\beta+x)(N-x)}{(N+\alpha-x)x}} \quad (3.114)$$

and

$$B = -\frac{(x-1)(N+\alpha-x+1)}{(x-1)(N+\alpha-x+1)+(\beta+1)(N-1)-(\alpha+\beta+2)(x-1)}\sqrt{\frac{(\beta+x)(N-x)}{(N+\alpha-x)x}\frac{(\beta+x-1)(N-x+1)}{(N+\alpha-x+1)(x-1)}} \quad (3.115)$$

EVALUATION PARAMETERS

The variables α and β in the Hahn polynomial, as described in Equation 3.106, can change the topology of the basis function. As previously described for the Racah moments, the change of shape of the basis function will affect the moments' classification ability for micro-Doppler signatures, where the values of a and b define the ROI of the basis functions. Therefore, a selection of parameter values is evaluated to optimize these parameters with respect to the classification rate. A selection of parameter settings is chosen for evaluation on micro-Doppler data. This selection consists of three cases where the values of a and b are appointed to move the ROI of the Hahn polynomial in the right, left and whole image. The 1D representation of the different polynomials is shown in figure 3.8 illustrating the shape change of the basis functions when varying a and b.

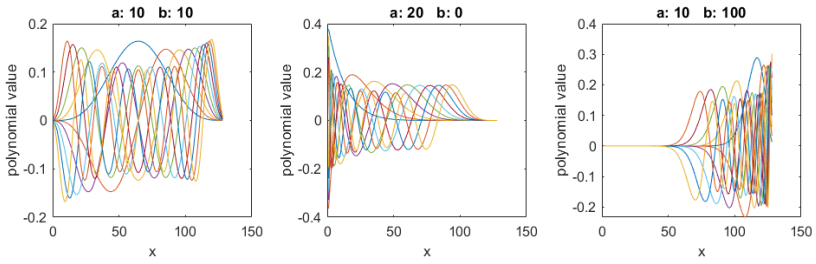


Figure 3.8: 1D representation of 10th order Hahn moment with varying values of the parameters a and b

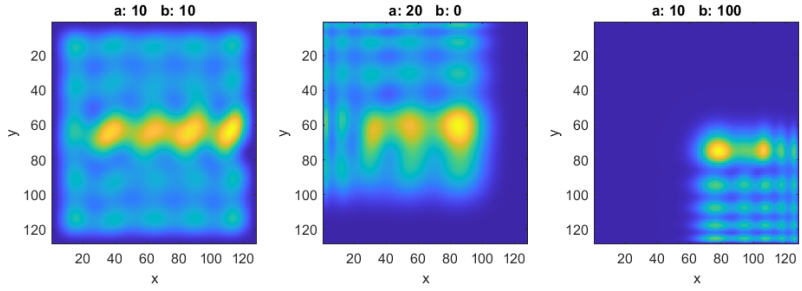


Figure 3.9: Example of image representation of 10th order Hahn moment with varying values of the parameters a and b

The 2D image representation of the different Hahn polynomials is shown in figure 3.9, where a spectrogram is reconstructed from the 10th-order moments.
The different cases are evaluated for classification of their classification performance by evaluating them in the same manner as the other discussed moments. The results are presented in Chapter 6.

KRAWTCHOUK MOMENTS

As with the Racah moment, a hypergeometric function defines the basis function. The basis function can again be obtained by the Racah polynomial and the Hahn polynomial as $t \to \infty$. For a 1D case, the Krawtchouk polynomial is defined as equation 3.117 and 3.118 with $(a)_k$ the Pochammer symbol.

$$\lim_{t\to\infty} \bar{h}_n(x;\alpha,\beta,N) = \bar{k}_n(x;p,N) \tag{3.116}$$

$$K_n(x;p,N) = \sum_{k=0}^{N} a_{k,n,p} x^k = {}_2F_1\left(-n,-x;-N;\frac{1}{p}\right) \tag{3.117}$$

with

$${}_2F_1(a,b;c;z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!} \tag{3.118}$$

$$(a)_k = a(a+1)\dots(a+k-1) = \frac{\Gamma(a+k)}{\Gamma(a)} \tag{3.119}$$

The weighted version, by means of a form of normalization, is derived for numerical stability. The weighting function is constructed by means of a binomial distribution as described below, with the orthogonality condition given, and this will be reduced to just δ_{nm} after normalization.

$$w(x;p,N) = \begin{pmatrix} N \\ x \end{pmatrix} p^x (1-p)^{N-x} \tag{3.120}$$

$$\sum_{x=0}^{N} w(x;p,N) K_n(x;p,N) K_m(x;p,N) = \rho(n;p,N)\delta_{nm} \tag{3.121}$$

where $n, m = 1, 2, \ldots, N$ and

$$\rho(n; p, N) = (-1)^n \left(\frac{1-p}{p}\right)^n \frac{n!}{(-N)_n}$$

$$\bar{K}_n(x; p, N) = K_n(x; p, N) \sqrt{\frac{w(x; p, N)}{\rho(n; p, N)}} \tag{3.122}$$

This results in the 2D weighted Krawtchouk moments described as

$$Q_{nm} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \bar{K}_n\left(x; p_1, N-1\right) \bar{K}_m\left(y; p_2, M-1\right) f(x, y) \tag{3.123}$$

IMPLEMENTATION

As with the previous discrete moments discussed, a recursive implementation is desired to lose the binomial operator and factorial terms. In [88] a recursive implementation for calculating the Krawtchouk polynomial is proposed. The recursion is given by:

$$p(N-n)\bar{K}_{n+1}(x) = A(Np - 2np + n - x)\bar{K}_n(x) - Bn(1-p)\bar{K}_{n-1}(x) \tag{3.124}$$

for

$$n = 1, 2, \ldots, N-2; \quad x = 0, 1, \ldots, N-1; \quad p \in (0, 1)$$

Where A and B are defined by

$$\begin{aligned} A &= \sqrt{\frac{p(N-n)}{(1-p)((n+1)}} \\ B &= \sqrt{\frac{p^2(N-n)(N-n+1)}{(1-p)^2(n+1)n}} \end{aligned} \tag{3.125}$$

And the initialization of $\bar{K}$

$$\begin{aligned} \bar{K}_0(x) &= \sqrt{\frac{w(x; p, N)}{\rho(0; p, N)}} \\ \bar{K}_1(x) &= \left(1 - \frac{x}{pN}\right) \sqrt{\frac{w(x; p, N)}{\rho(1; p, N)}} \end{aligned} \tag{3.126}$$

for $x = 0, 1, \ldots, N-1$.

EVALUATION PARAMETERS

The weighted Krawtchouk moment has some remarkable properties. Krawtchouk moments can change its ROI, by changing its p_1 and p_2 corresponding to the binomial distribution in equation 3.123 and 3.120 i.e., adjusting its resolution on certain parts of the image. The p values change the basis functions so that, in lower-order Krawtchouk moments, the selection of the p values could be seen as specifying the window location in

which the moments capture information. On higher orders, the value of p can be interpreted as the point at which the basis function begins to expand, e.g., defining the starting point of the ROI. To be able to exploit this property for micro-Doppler classification purposes, an analysis is performed using a selection of p values. As an illustration of the influence of this variable, three different cases are shown in figures 3.10 and 3.11, where $p_1 = p_2$ and values 0.3, 0.5 and 0.7 are chosen, evaluated on two orders as data a spectrogram of a human micro Doppler signature is used as a sample image. From the figures, it can be easily observed that the window captured by the moments depends on the order and value of p, where p will influence the starting point of expansion of the basis function and the order of the size and resolution of the frame. Therefore, the topology of the classification data will influence the classification potential of the Krawtchouk moment, given a particular value of p. To illustrate their influence on classification performances, the three selected cases are used to classify three different micro-Doppler data sets. The results and analysis of this classification are presented in Chapter 6.

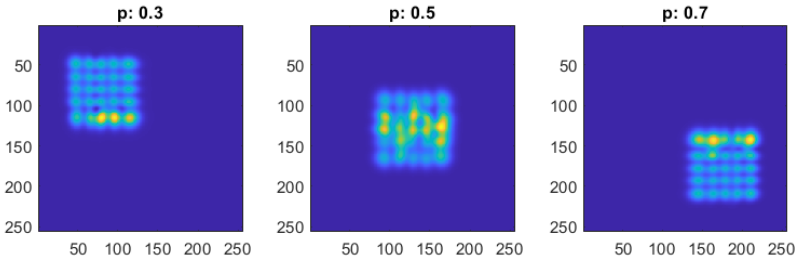


Figure 3.10: Example of image representation of 10th order Krawtchouk moment with varying values of the parameter p

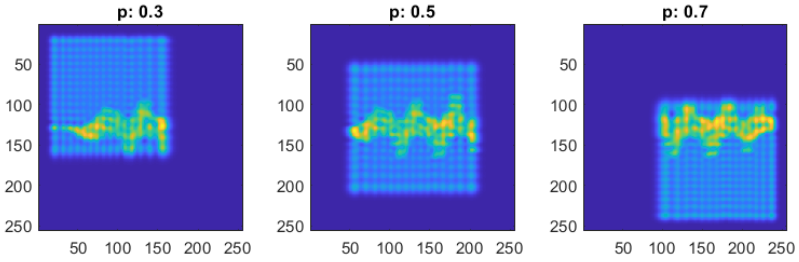


Figure 3.11: Example of image representation of 30th order Krawtchouk moment with varying values of the parameter p

TCHEBICHEF

The Tchebichef polynomial is related to Racah, Hahn, and Krawtchouk.[53] The polynomial can be formed from the Hahn polynomial by letting $t \to 0$.

$$\lim_{t \to 0} h_n(x; \alpha, \beta, N) = t_n(x) \tag{3.127}$$

This results in

$$t_n(x) = (1-N)_n {}_3F_2(-n,-x,1+n;1,1-N;1) \\ n,x,y = 0,1,2,\ldots N-1 \tag{3.128}$$

which can be written in the form

$$t_n(x) = n! \sum_{k=0}^{n} (-1)^{n-k} \left(\begin{array}{c} N-1-k \\ n-k \end{array} \right) \left(\begin{array}{c} n+k \\ n \end{array} \right) \left(\begin{array}{c} x \\ k \end{array} \right) \tag{3.129}$$

The scaling factor $\beta(n,N) = N^n$ is introduced for numerical stability.

$$\beta(n,N) = N^n. \tag{3.130}$$

The scaled Chebyshev polynomial is then given by

$$\tilde{t}_l(x) = \frac{t_l(x)}{\beta(l,L)}. \tag{3.131}$$

The norm or amplitude factor is given as

$$\rho(n,N) = (2n)! \left(\begin{array}{c} N+n \\ 2n+1 \end{array} \right). \tag{3.132}$$

The dependence on the binomial operator $\left(\begin{array}{c} N+n \\ 2n+1 \end{array} \right)$ and the factorial term $(2n)!$ will limit the maximum order due to the computational complexity. Therefore a recursive the function is rewritten in a recursive manner:

$$\rho(n,N) = \frac{(2n)!(N+n)!}{(2n+1)!(N-n-1)!} \tag{3.133}$$

using $n! = n(n-1)!$, this can be rewritten as

$$(2n+1)\rho(n,N) = \rho(n-1,N)(N^2-n^2) \tag{3.134}$$

and the scaled version $\tilde{\rho}(n,N) = \frac{\rho(n,N)}{\beta(n,N)^2}$

$$(2n+1)\tilde{\rho}(n,N) = \frac{\tilde{\rho}(n-1,N)(N^2-n^2)}{N^2} \tag{3.135}$$

$$\tilde{t}_n(x) = \frac{(2n-1)\tilde{t}_1(x)\tilde{t}_{n-1}(x) - (n-1)\left(1-\frac{(n-1)^2}{N^2}\right)\tilde{t}_{n-2}(x)}{n}, \quad n = 2,3,\ldots N-1 \tag{3.136}$$

The resulting moment is defined as

$$T_{l,h} = \frac{1}{\tilde{\rho}(l,L)\tilde{\rho}(h,H)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \tilde{t}_l(x)\tilde{t}_h(y) f(x,y), \tag{3.137}$$

With $f(x,y)$, the image function. Note that when $t \to 0$ results in the Hahn polynomial $\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 0$. As described in section 3.4, it would be the same as the Hanh moment that captures the whole frame. Note that the Tchebychef moment cannot emphasize a certain region.

MEIXNER MOMENTS

The Meixner moments are part of the family of discrete moments. The polynomial is defined as

$$\omega_n^{(\beta,\mu)}(x) = (\beta)_{n^2}F_1(-n,-x;\beta;1-1/\mu) \tag{3.138}$$

Which is based on a hypergeometric function and has free parameters β and μ

$$\text{where } \beta \text{ and } \mu \text{ are restricted to } 0<\mu<1 \text{ and } \beta>0. \tag{3.139}$$

And the normalized Meixner polynomial is then defined as:

$$\tilde{\omega}_n^{(\beta,\mu)}(x) = \sqrt{\frac{w(x)}{d_n^2}}n = 0,1,\ldots,N-1 \tag{3.140}$$

With

$$w(x) = \frac{\mu^{x+n}\Gamma(n+\beta+x)}{\Gamma(\beta)x!} \tag{3.141}$$

and

$$d_n^2 = \frac{n!(\beta)_n}{\mu^n(1-\mu)^\beta} \tag{3.142}$$

The resulting Meixner moments of order n are then calculated for an image function $f(x,y)$ over the domain NxM using:

$$MM_{nm} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\tilde{\omega}_n^{(\beta,\mu)}(x)\tilde{\omega}_n^{(\beta,\mu)}(y)f(x,y) \tag{3.143}$$

IMPLEMENTATION

The basis function is implemented similarly to the other discrete moments. Here, a recursive implementation is chosen to be able to calculate the basis functions up to high orders. As presented in [55], the approach was used to achieve a recursion on n. The recursion is implemented as follows:

$$\tilde{\omega}_n^{(\beta,\mu)}(x) = \frac{BD}{A}\tilde{\omega}_{n-1}^{(\beta,\mu)}(x) - \frac{CE}{A}\tilde{\omega}_{n-2}^{(\beta,\mu)}(x) \tag{3.144}$$

With

$$\begin{aligned} A &= \tfrac{\mu}{\mu-1} \\ B &= \tfrac{x-x\mu-n+1-\mu n+\mu-\beta\mu}{1-\mu} \\ C &= \tfrac{(n-1)(n-2+\beta)}{1-\mu} \\ D &= \sqrt{\tfrac{u}{n(\beta+n-1)}} \\ E &= \sqrt{\tfrac{u^2}{n(n-1)(\beta+n-2)(\beta+n-1)}} \end{aligned} \tag{3.145}$$

The initialization is of $\tilde{\omega}_0^{(\beta,\mu)}(x)$ and $\tilde{\omega}_1^{(\beta,\mu)}(x)$ is calulated using:

$$\tilde{\omega}_0^{(\beta,\mu)}(x) = \sqrt{\frac{\omega(x)}{d_n^2(0)}}$$
$$\tilde{\omega}_1^{(\beta,\mu)}(x) = \left(\gamma + x - \frac{x}{\mu}\right)\sqrt{\frac{\omega(x)}{d_n^2(1)}} \tag{3.146}$$

Note that because of the Gamma functions and the exponential terms in equation 3.141, the function cannot be calculated to high orders and high values of x and β. Therefore the function will be recursively implemented, resulting in the procedure:

$$w(x) = \frac{x+\beta-1}{x} * w(x-1) * \mu \tag{3.147}$$

with the initialization $w(0) = 1$

EVALUATION PARAMETERS

The parameters β and μ are free under the condition 3.139. Where the two parameters will change the shape of the basis functions. The combination of the two variables will define the ROI of the functions. β can be seen as an indication of the center on which the function starts expanding. The parameter μ defines the distribution of the zeros. In the case of $\mu < 0.5$ this will emphasize the left side of β, $\mu = 0.5$ centred around β and lastly, $\mu > 0.5$ right of β. Four variations in parameter selections are chosen to analyze the effect on the moments, selected in a way that gives a good overview of the variety in the characteristics. The different cases are shown in figure 3.12. In figure 3.13, the image reconstruction of the Meixner moments is shown when calculated for a 128x128 spectrogram of a micro-doppler signature. The variations of the moments are evaluated the same way as the other moments analyzed. This results in an overview of the influence on the classification result when varying parameters. The results can be found in Chapter 6

CHARLIER MOMENTS

Charlier moments are part of the family of discrete moments. The basis function is defined using a hypergeometric function:

$$c_n^{a_1}(x) = {}_2F_0\left(-n, -x; \frac{-1}{a_1}\right) \tag{3.148}$$

for $0 < a_1$. When rewriting the Hypergeopmatic term, the equation function becomes:

$$C_n^{a_1}(x) = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} \begin{pmatrix} x \\ k \end{pmatrix} k!(-a_1)^{-k} \tag{3.149}$$

The normalized Charlier polynomial follows from the orthogonality condition and is defined as

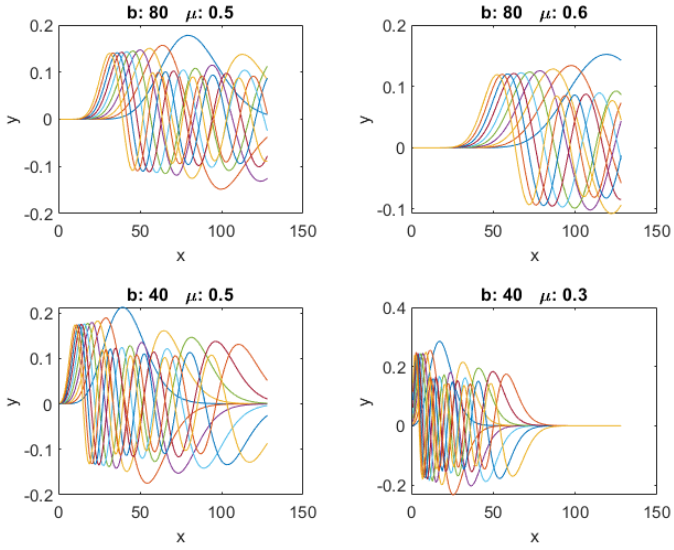


Figure 3.12: Representation of 10th order Meixner basis function with varying values of the parameters β and μ

$$\tilde{C}_n^{a_1}(x) = C_n^{a_1}(x)\sqrt{\frac{\omega(x)}{d_n^2}} \tag{3.150}$$

with

$$\omega(x) = \frac{e^{-a_1}a_1^x}{x!} \tag{3.151}$$

$$d_n^2 = \frac{n!}{{a_1}^n} \tag{3.152}$$

As with the previous discrete moments discussed, a recursive implementation is desired to lose the binomial operator and factorial terms. In

$$\tilde{C}_{n+1}^{a_1}(x) = \frac{a_1 - x + n}{a_1}\sqrt{\frac{a_1}{n+1}}\tilde{C}_n^{a_1}(x) - \sqrt{\frac{n}{n+1}}\tilde{C}_{n-1}^{a_1}(x) \tag{3.153}$$

With initialization:

$$\begin{aligned}\tilde{c}_0^{a_1}(x) &= \sqrt{\frac{w(x)}{d_0^2}} = \sqrt{\frac{\mathrm{e}^{-\mu}\mu^x}{x!}} \\ \tilde{c}_1^{a_1}(x) &= \frac{\mu - x}{\mu}\sqrt{\frac{w(x)}{d_1^2}} = \frac{\mu - x}{\mu}\sqrt{\frac{\mathrm{e}^{-\mu}\mu^{x+1}}{x!}}\end{aligned} \tag{3.154}$$

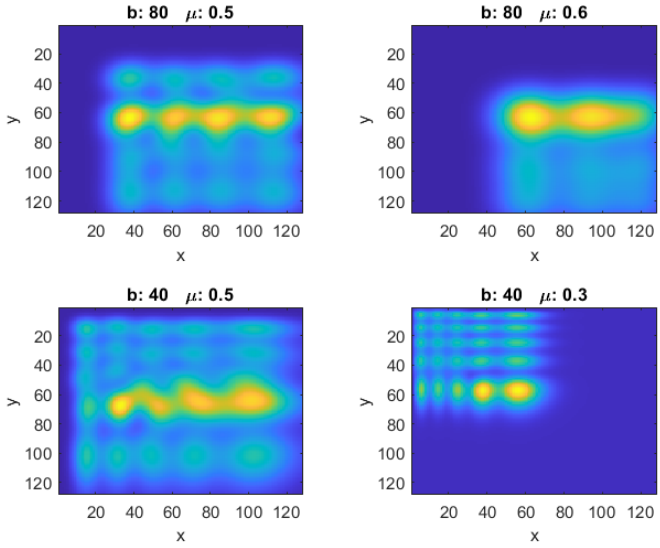


Figure 3.13: Example of image representation of 10th order Meixner moment with varying values of the parameters β and μ

The Resulting Charlier moments are then calculated for an NxM case, using the following:

$$C_{nm} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \tilde{C}_n^{a_1}(x) \tilde{C}_m^{a_1}(y) f(x,y) \tag{3.155}$$

EVALUATION PARAMETERS

The Charlier basis function has a free parameter a_1. This parameter will influence the ROI of the basis function where a_1 can be mainly seen as the point where the basis function is centered. This behavior can be observed in figure 3.14. The effect on the 2d ROI of the moment can be observed in figure 3.15. Here, also the compression in the distribution of zeros is visible. As with the earlier discussed moments, the selection of the free parameter values, as given in the presented figures, will be evaluated on the classification pipeline to see their influence on the classification performance.

3.5. TO SUMMARIZE

All the moments investigated in this research have been mathematically derived, and an implementation is proposed for each moment. With the proposed implementations, the moments can be calculated to a maximum order equal to 50 without issues of stability. This opens up the possibility of using the moments as features. Furthermore, for the moments with free parameters in their formulation, various variables are chosen to test further regarding their classification abilities. The implementation of the moments as

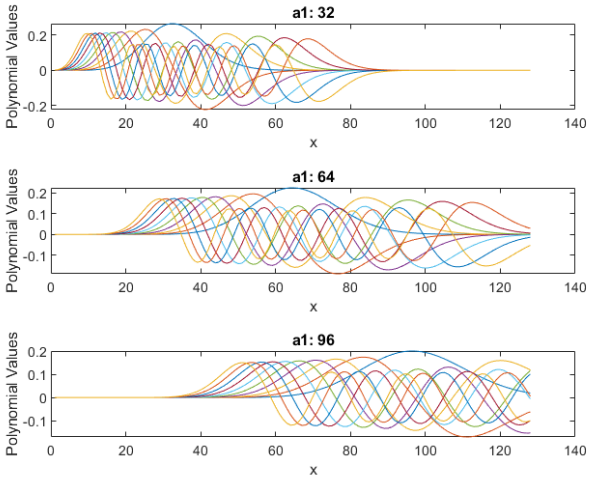


Figure 3.14: Basis functions of 10th order Charlier moment with varying values of the parameter a_1

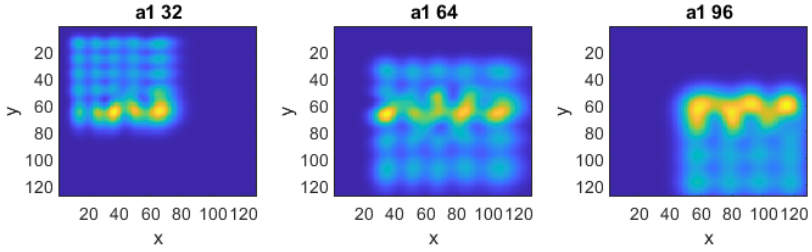


Figure 3.15: Image representation of 10th order Charlier moment with varying a_1 values

presented in this chapter is used to construct the feature vector discussed in Chapter 5. The final classification results are then presented in Chapter 6.

3.6. CONTRIBUTION OF CODE

As a further contribution, the MATLAB scripts that implement the moments as described in this chapter are shared after the publication of this thesis. The scripts and documentation can be found at https://doi.org/10.4121/efd059aa-f2c2-4c54-be85-4e600876f444.v1[10].

4

MICRO-DOPPLER SIGNATURE DATA

This chapter describes the experimental micro-Doppler datasets used for the verification of the moment-based classification approaches formulated in this thesis. The analysis contains an evaluation of the topology and the influence of the different moments, an evaluation of the data concerning the physical aspects they capture, and an analysis of a selection of pre-processing techniques and their influence on the classification potential of the orthogonal moments.

4.1. DATA GENERATION

This research focuses on the classification of micro-Doppler signatures and tries to generalize the micro-Doppler classification problem as a general research question. However, the micro-Doppler signatures will not come from a constant source, as different factors will influence the data and therefore their appearance and the way relevant information is encoded in them. As described in Chapter 2, first various objects are of interest, which causes a variety of signatures. In addition to the diversity of objects, the type of radar and settings will also influence the data gathered, resulting in a time-Doppler representation that differs from case to case. This means that micro-Doppler signature data cannot be treated as homogeneous. To generalize the classification results as much as possible, a variety of data sets will be evaluated; these sets will have different settings and targets. This will show the classification abilities of the moments for these various cases.

4.1.1. MICRO-DOPPLER SIGNATURES

The Micro-Doppler signatures are generated by taking the short-time Fourier transform (STFT) over the data generated from the radar. The Doppler-time representation of the data is generally represented by the spectrogram. The spectrogram is obtained by representing the signal power in time and Doppler frequency. This is illustrated in the equation 4.1. Here $\tilde{s}(n)$ indicates the radar signal and $h(n-k)$ a window filter.

$$\chi(v,k) = |STFT(x)|^2 = \left| \sum_{n=0}^{N-1} \left(\tilde{s}(n) \cdot h(n-k) \cdot e^{\frac{-j2\pi vn}{N}} \right) \right|^2, k = 0, \ldots, K-1, \tag{4.1}$$

Most datasets and classification pipelines treat digitally the signatures as images, whereas several datasets provide the data in spectrograms saved as images. The literature study provided in Chapter 2 shows that the data in such proposed pipelines are commonly presented as a spectrogram for micro-Doppler classification problems. Note that raw data are not available for the cases where the datasets are provided as a spectrogram. Different techniques have been proposed to transform the raw radar data, e.g., the wavelet transform or other time-frequency distributions. Evaluation of all transformation possibilities for raw radar data is considered outside of the scope of this research. The input data for the proposed classification pipeline is assumed to be in the format of a spectrogram described as an image.

4.2. DATASETS USED FOR THE ANALYSIS

In Chapter 2, a selection of different micro-Doppler data sets is presented. From this selection, three data sets are used to evaluate the moments. The data sets used are selected on the basis of the analysis presented regarding the general properties of the data, e.g. presence of harmonics and topologies. Where a variation in data set properties is desired to subject the moments to distinct cases, the different cases will open up the opportunity to test the classification abilities of the moments more broadly.
Knowing the characteristics of the data can make a possible link on why/which orthogonal moments will perform the way they do. Section 3 shows that the basis function of the moments has different characteristics. The functions expand in different ways, resulting in a difference in how the ROI of the functions behaves. Therefore, the way a function places emphasis on a certain region will be different between the functions. The difference in the data topology of the different classes could influence the performance of the moments. All classes of the sets will be evaluated to identify how the data distributions look. Connecting the data with the physical event they represent, e.g. waving hands with a particular repetition frequency and duration in Doppler time. From this evaluation, the class-defining properties of classes can be estimated. These properties can be compared to the properties of the moments to explain their classification behaviour.
To get an idea of the general distribution of the data, we used the mean and variance of the data set. Variance can be seen as a measure of the spread of the data. In the case of a dataset defined by classes, the assumption can be made that when taking the variance over the whole set, the variance between the data from different classes will be dominant, i.e. measuring the variability between the classes. In the case of a classification problem, a region of significant statistical spread could identify a class-defining region, as the data differs from class to class in this location. The variance will be defined as:

$$v_{ij} = \frac{1}{N-1} \sum_{k=1}^{N} \left| x_{ij} - \mu_{ij} \right|^2 \tag{4.2}$$

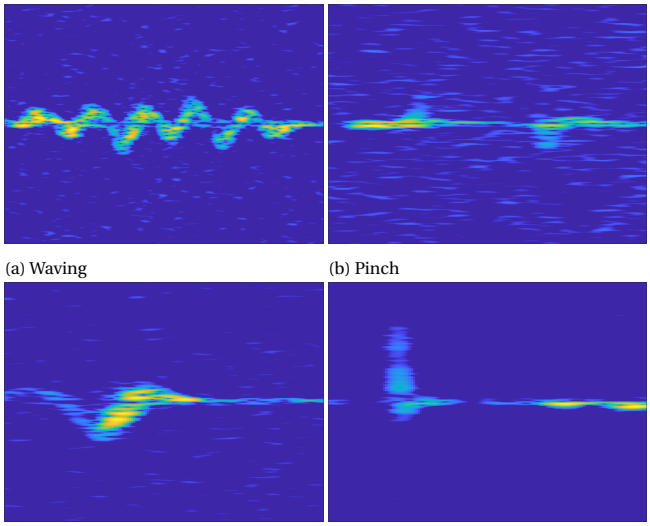
Where i and j define the pixel location of the image, N is the number of the data sample, and μ_{ij} indicates the average value of the pixel at location (i, j) for the whole data set.

4.2.1. DOP-NET DATASET

The Dop-NET data set [22] is a set of four classes of human hand motion. The data is recorded using a 24GHz FMCW radar with a bandwidth of 750 MHz and a chirp period

of 1 ms. Four human hand signatures define the classes. The data are published as a pre-processed spectrogram. One sample of each class is shown in figure 4.1. The different classes and their physical interpretation are listed below:

- **Wave**: A waving motion of the human hand. Note that this creates a harmonic response in time Doppler, as shown in figure 6.5a. This continuous harmonic means that frequency-based preprocessing, e.g. CVD, could improve classification rates. Also notable is that the same harmonic spreads over the whole horizontal axes.
- **Pinch**: The thumb and index finger move to each other and back. This creates a temporal Doppler-time response of a positive and negative frequency shift, figure 6.5b. In this case, the information is only present in a certain part of the horizontal-time axes.
- **Click**: The thumb moves to the middle finger, similar to the pinch, but faster. Therefore, the Doppler-time response is similar, but the positive and negative parts are closer in time, figure 6.5c.
- **Swipe**: The hand is moved in the direction of the pointing direction of the thumb. This creates a single positive response in Doppler time, figure 6.5c. Again, we create a temporal response in Doppler time where the class information is situated in only a small part of the time axes.



(a) Waving (b) Pinch

(c) Click (d) Swipe

Figure 4.1: Spectrogram examples for the four classes of the Dop-NET dataset

The variance of the different pixel locations across the whole data set is shown in Figure 4.2. This figure shows the variability of the pixels; this can be used as an indication of a region of interest. It can be observed that the whole horizontal axis, representing time,

has high variability. Therefore, defining an ROI of the horizontal pixel location will lead to a potential loss in classification. This would require using a moment that captures the whole horizontal axes. In addition, regions of high variability are present above and below the central x values, indicating positive and negative Doppler values. This region has the highest values around the vertical center of the image. Vertically, the classification rate could benefit from defining an ROI and the appropriate moment close to the vertical center of the image. A moment type that performs well in the illustrated region is expected to also perform well in terms of classifying this dataset.

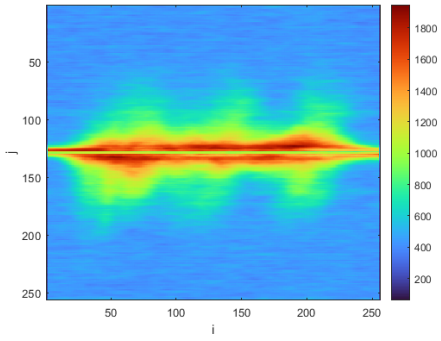


Figure 4.2: The representation of the variance over the pixel location of the whole Dop-NET dataset

4.2.2. GLASGOW DATASET

The Glasgow set [2] is a dataset containing micro-Doppler signatures of different human activities. The set is measured using a 5.8 GHz radar with a 1ms chirp duration and 400 MHz bandwidth. The set used contains six different events defined by their human activity. The class activities and their micro-Doppler signatures are discussed below.

- **Walking**: The person walks away, turns and walks towards the radar. This causes a micro-Doppler response that is negative when walking away and positive when walking towards. The multibody dynamics produced by the limbs produce harmonics around this burst Doppler velocity, figure 4.3a. Note that the harmonics are spread over the total length of the data frame. Also, the continuous presence of harmonics could indicate that, for this class, a frequency-separating preprocessing step, e.g., CVD, would improve classification results.

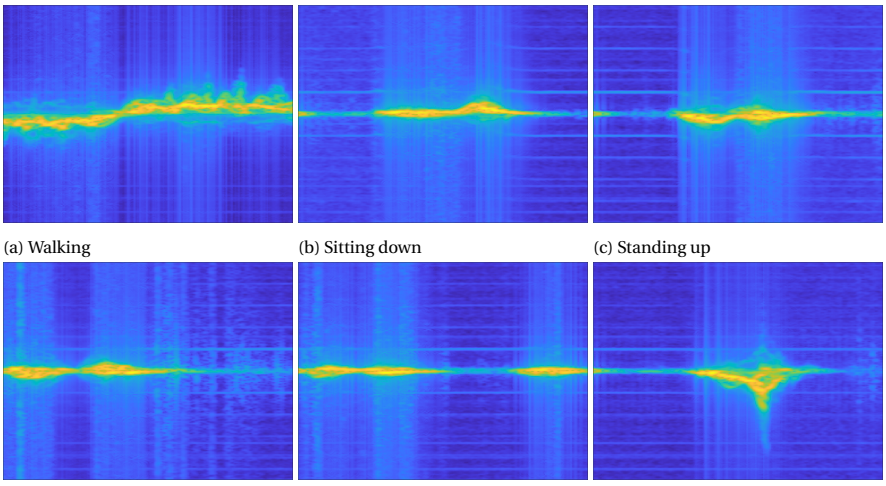
- **Sitting down**: A person sits in a chair. This introduces a bulk of complex moving multi-dynamic bodies. It is hard to say which body part and motion will produce the dominant part of the signature. Figure 6.5b shows a temporal Doppler response. Furthermore, are there no clear harmonics present

- **Standing up**: A person stands up from a sitting position. This movement is similar to sitting down class, where this produces the inverse motions. The produced Doppler response is temporal and without any clean, continuous harmonics.

- **Pick up an object**: A person is kneeling or bending down to pick up an object from the floor. This motion is similar to sitting down and standing up, as it creates a bulk of complex motions. The Doppler response is temporal and without any clear harmonics.

- **Drink water**: Pouring and drinking a glass of water. The arm movements of the person will dominate this activity. The Doppler velocity will be low as the arms will not move at high speeds. This can be seen in figure 4.3e. Also, the data are spread over more time. Here, the Doppler response stretches almost the whole data frame.

- **Fall**: A person will fall on the ground. The falling motion will produce a high negative Doppler velocity, as can be seen in Figure 4.3f. The limbs will cause less dominant responses in different motions. The Doppler response will be temporal an without any continuous harmonics.



(a) Walking (b) Sitting down (c) Standing up

(d) Pick up an object (e) Drink water (f) Fall

Figure 4.3: Spectrogram examples for the six classes of the Glasgow dataset

The variance over the dataset of each pixel location is again used as an indication of interclass variability. Figure 4.4 shows the variance figure. This figure shows that the most variable pixes are located at the vertical center of the image, that is, the time axes. Horizontally, the variance is spread uniformly, with the exception of the left side. There an increase in the variability of the pixels is shown. Emphasizing this could potentially increase the classification results. For the horizontal axes, the emphasis should be on the center.

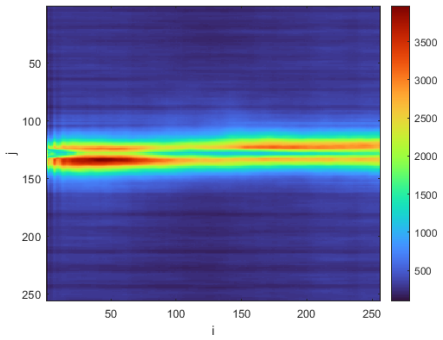
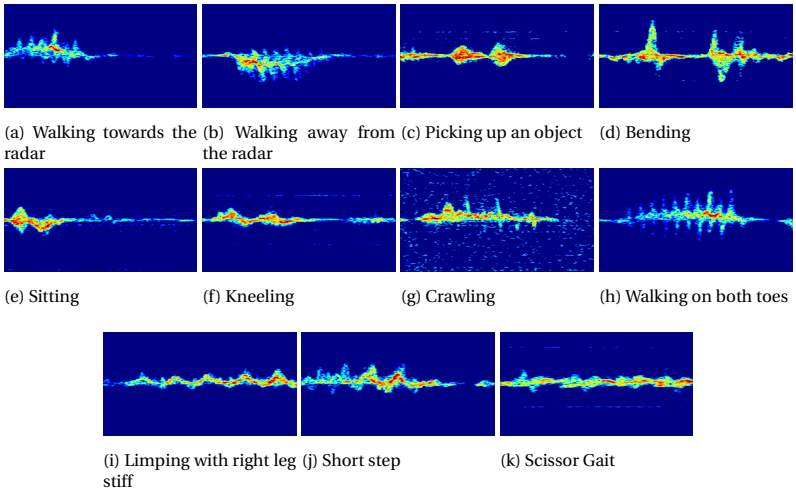


Figure 4.4: The representation of the variance over the pixel location of the whole Glasgow dataset

4.2.3. CROSS-FREQUENCY SET - XETHRU

The Cross-frequency set used [25] is initially made with three types of radar, each providing its own data set. The set is produced to analyze the influence of different types of radar on a classification problem. For this research, one of the data sets is used. This set was recorded with a XeTru UWB impulse radar with a frequency band of roughly 7GHz - 8GHz. The set consists of 11 classes of human activities. The activities are more complex in motion than the earlier ones and are described in figure 4.5. Note that because of the larger amount of classed and less separable activities, this data set will be harder to classify.



(a) Walking towards the radar (b) Walking away from the radar (c) Picking up an object (d) Bending

(e) Sitting (f) Kneeling (g) Crawling (h) Walking on both toes

(i) Limping with right leg stiff (j) Short step (k) Scissor Gait

Figure 4.5: Spectrogram examples for the eleven classes of the Cross-frequency XeThru dataset

The pixel variance over the dataset is shown in figure 4.6. Where the highlighter regions indicate a location of high variability, this region is mainly spread around the vertical

center and the whole horizontal part of the image. Note that the vertical spread of the variability is wider than in the previous presented dataset. Moments that emphasize this region could potentially increase the classification results.

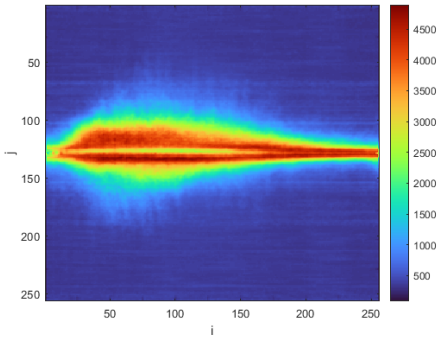


Figure 4.6: The representation of the variance over the pixel location of the whole Cross-frequency XeThru dataset

4.2.4. TOPOLOGY OVERVIEW

The variance figures of all the selected datasets show variability that is concentrated in the vertical center of the image function and spreads throughout the horizontal direction. As the variance over the classes is assumed to be an indicator for class-defying regions, the moments that will capture this region will potentially have better performances. The mathematical evaluation of the moments presented in 3 shows that in general most moments capture this region. However, those moments that can move the ROI, i.e. the moments with free parameters, do not always capture the whole vertical region at lower orders. Therefore, it is assumed that these moments will perform better at higher orders. This hypothesis is tested and the results are presented in Chapter 6.

4.3. CVD AS PRE-PROCESSING METHOD

The literature study in chapter 1 shows that CVD is always used in the general processing pipeline of previous research using orthogonal moments. The Cadence Velocity Diagram is created when taking the FFT over time at each Doppler frequency bin in the time-frequency spectrogram [19], i.e. it returns the motion harmonics of the targets measured. As the analysis done on the different data sets shows that not all classes contain continuous, well-defined harmonics, the benefits of this mapping can be questioned. The orthogonal moments will be analyzed on both spectrograms and CVD to analyze the influence of using one instead of the other. The CVD will be calculated using:

$$\Delta(v,\epsilon) = \left| \sum_{k=0}^{K-1} \left(\chi(v,k) \cdot e^{\frac{-j2\pi k\epsilon}{K}} \right) \right|, \tag{4.3}$$

Where $\chi(v,k)$ is the spectrogram defined in the equation 4.1. The data characteristics of the data sets, as analyzed above, will change when taking the CVD, i.e., the whole data

topology will be different due to this mapping.

PRESENCE OF HARMONICS IN DATASETS

The presence of harmonics in a data set would favor the argument for using CVD as an additional preprocessing step. For the selected datasets, some classes show the presence of a harmonic. Although only a small portion of the classes have a recognizable continuous harmonic presence over the whole length of the captured frame. Therefore, it would be hard to say if techniques such as CVD would be beneficial for these datasets. This is evaluated for all moments, and the results are presented in Chapter 6

4.4. NOISE INTRODUCTION

The robustness against noise introduction will be tested for the moments discussed. The data used in this research is measured in different settings and environments. This introduces a problem on where the data is not noise-free, essentially it will be unknown what the contribution of this noise is compared to the captured signal.
To evaluate the robustness against noise, the default data from the data sets will be seen as a clean signal. Noise will then be introduced into the data in an additive manner. For this research, the choice is made to introduce additive Gaussian noise to the spectrogram representation of the radar data. Note that this noise introduction does not have a direct physical meaning. The noise introduction in this way only evaluates the robustness with random Gaussian variations. The evaluation of the robustness against more physical varieties of noise and distortions is out of the scope of this research for time limitations. This method of testing the influence of noise on classification ability is however similar to the method used in image recognition [38]. For this evaluation, the radar data are not tested on the physical implementation of noise, but the spectrogram of the data is treated as done in image processing, so the data are treated as an image. This method shows the robustness of the moments regarding zero-mean Gaussian noise and provides a first indication of the general noise robustness of the moments.

5

CLASSIFICATION PROCESSING STEP

This chapter discusses and analyses the different aspects of the proposed classification pipeline exploiting a set of orthogonal moment-based features. These aspects include specifically the feature vector generation from the discussed moments, the classifier selection, and the feature selection out of an initially established set. The analysis of these aspects illustrates the choices available in these components of the classification pipeline. Different methods are introduced to optimize the overall classification performance of the discussed moments.

5.1. FEATURE VECTOR GENERATION

A feature vector must be made from the generated moments to use a distance-based classifier, e.g. KNN or SVM. Specific properties of the moments can be exploited to generate a vector that optimizes the overall classification performances. Not all moments have the same properties. Therefore, different approaches are discussed for different types of moments.

5.1.1. COMPLEX MOMENTS

The moments that are defined on a circle have an imaginary part. This imaginary part is a problem for most simple Euclidean classifiers, since they are designed to deal with only real numbers. There are several ways to deal with the imaginary parts. In [38], the characteristic vector is made up of taking the magnitude. The rotational invariance can still be proven as it satisfies the equation:

$$V_{nm}(r\cos\theta, r\sin\theta) \equiv V_{nm}(r,\theta) = R_n(r)A_m(\theta) \tag{5.1}$$

Where $A_m(\theta) = \exp(jm\theta)$ and $R_n(r)$ is the radial basis function. This invariance holds since, when taking the length of the exponential function, the length will be constant,

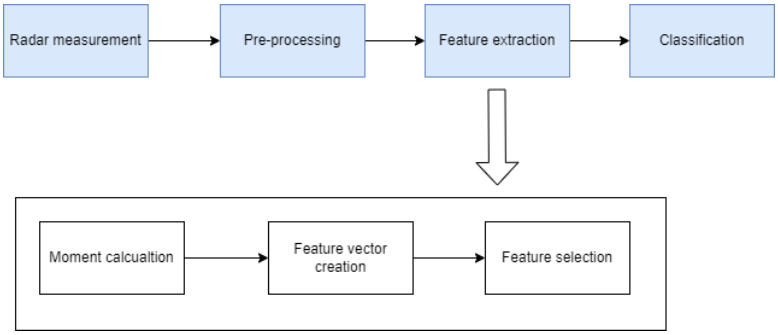


Figure 5.1: Zoom in on the feature selection block of the overall classification pipeline

independent of the value of θ. Suppose there will be a rotation of angle θ_0; then:

$$|V_{nm}(r,\theta_0)| = |R_n(r)exp(jm\theta_0)| = |R_n(r)| \tag{5.2}$$

and for the angle:

$$\phi_{nm}^{rot} = \phi_{nm} - m * \theta_0 \tag{5.3}$$

In most research, the importance of the rotational invariance property implies that the magnitude of the complex moments is taken to construct the feature vector. This procedure excludes the phase part of the moments as a source of information. In [89] it is shown that the phase of Zernike moments does capture information about the image function. It was proved that the expected distortion caused by the variation in magnitude is less than or equal to the distortion caused by the variation in phase, indicating that the phase contains information of importance similar to the magnitude.

PROPOSED FEATURE VECTOR CONSTRUCTION METHOD

Considering the aforementioned discussion, for this research the importance of the rotational invariance property can be questioned, because the micro Doppler signatures are defined on a non-rotational square grid. Therefore, the imaginary part can be added to the feature vector, contributing the additional phase information. The vector will be constructed using the normalized real part of the orders 0 to nth. The complex normalized part of the moment is added to the end vector.

$$\tilde{\mathbf{F}}_{real} = \frac{\mathbf{F}_r - \mu_{\mathbf{F}_r}}{\sigma_{\mathbf{F}_r}} \quad , \quad \tilde{\mathbf{F}}_{Imag} = \frac{\mathbf{F}_i - \mu_{\mathbf{F}_i}}{\sigma_{\mathbf{F}_i}} \tag{5.4}$$

With $\mathbf{F}_i = \Im\mathfrak{m}(\mathbf{F})$ and $\mathbf{F}_r = \Re\mathfrak{e}(\mathbf{F})$ for a feature vector $\mathbf{F}$, where μ indicates the mean and σ the standard deviation. The final vector is constructed as:

$$\tilde{\mathbf{F}} = [\tilde{\mathbf{F}}_{real}, \tilde{\mathbf{F}}_{Imag}] \tag{5.5}$$

5.1.2. FEATURE VECTOR NORMALIZATION

The moments defined on a square grid are commonly normalized in their feature vector. This procedure ensures that the impact of all features is normalized, removing a potential bias in the classification process [90]. The normalization of the feature vector is defined as:

$$\tilde{\mathbf{F}} = \frac{\mathbf{F} - \mu_{\mathbf{F}}}{\sigma_{\mathbf{F}}} \tag{5.6}$$

Here, $\sigma_{\mathbf{F}}$ is defined as the standard deviation of the feature vector $\mathbf{F}$ and $\mu_{\mathbf{F}}$ the mean of $\mathbf{F}$. The feature vector itself is composed by taking all non-zero components from the moment matrices. The normalization step can be performed on all moment-based feature vectors. In section 6, the results are shown in terms of the classification performance with normalization on the complex feature vectors.

5.2. EUCLIDEAN DISTANCE-BASED CLASSIFIERS

To focus the investigation on the different properties and parameters of the studied moments, a simple classifier is preferred. Note that the property of interest is the extraction ability of class-defying information of the data of each moment. A classifier based on the distance of the features between the classes will preserve these individual properties of the moments. When choosing a relatively advanced classification algorithm, e.g. Neural Networks, the classification results cannot be traced back as an achievement by the moment investigated. A selection of the most commonly used classifying algorithms is discussed and implemented.

5.2.1. K-NEAREST NEIGHBOR (KNN)

From the literature review presented in Chapter 1, a KNN is the most commonly used classification algorithm in research related to orthogonal moments as a feature for micro Doppler signature classification. This algorithm is commonly chosen due to its simple nature. The parameter k has to be chosen. From the literature, the most common cases are 1, 3 and 7 [64]. It does not specify whether the values used are optimal for the evaluated moment. Also, it would be hard to reason which value would be optimal for each case. Therefore a variety of k: 1 to 10 is chosen and tested using the moments discussed and the selected datasets. The results can be found in Chapter 6

5.2.2. SUPPORT VECTOR MACHINE (SVM)

A Support Vector Machine (SVM) is commonly chosen as a classifier. This classifier is chosen for the same reason as the KNN, i.e. as a simple distance-based classifier. The SVM algorithm is dependent on a specific type of kernel. This kernel influences the classification ability, which depends on the distribution of the features. In previous research, several kernels are used in classifying problems in combination with orthogonal moments. The most common kernels are listed below, where x_j and x_k are vectors of dimension p representing observations j and k in x. q represents the polynomial kernel order.

- Linear
$$G(x_j, x_k) = x_j{}' x_k$$
- Quadratic
$$G(x_j, x_k) = (1 + x_j{}' x_k)^q$$
- Radial Basis Function
$$G(x_j, x_k) = \exp\left(-\|x_j - x_k\|^2\right)$$

The choice of kernel and if this kernel is optimal for a classification problem with a particular type of moment is not stated. There are algorithms proposed for the automated selection of kernel and regularization parameters. A complete derivation of optimal kernels and parameters is beyond the scope of this research. To give an initial idea of kernel selection, the kernels listed above are selected and used in Chapter 6. This gives an initial optimal approach to using an SVM in micro-Doppler classification problems with the applied moments.

5.3. FEATURE SELECTION

The general idea behind using orthogonal moments for micro-Doppler classification is to extract class-defining features directly from them. Data will be mapped from a large set of data points to a smaller set, where the class-defining aspects of the data are preserved. In the simplest case, the moments will be used directly for classification. This means that the moments generated are used as features in a particular type of classifier. Section 5.2 will go deeper into selecting a classifier, but for now, the assumption will be made that the classifier will be limited in the number of features, i.e., the classifier will be sensitive to overfitting. The compressing data rate of the moments differs from moment to moment.

Moment	**Feature vector size**
ZM	(K+1)(K+2)/2
PZM, GM,LM,GHM,CHM,CHDM,KM,HM,DHM,RM,CM,MM	(K+1)^2
OFMM, CHFM, PJFM, JFM, RHFM, BFM	(K+1)(2K+1)
EFM, PCEM	(2K+1)^2

Table 5.1: Table illustrating the size of the feature vector of the calculated moments. K represents the order of the calculated moments

An overview of the moment matrix size created concerning their order is given in Table 5.1. The table lists the non-zero components, as not all moments expand squarely. As can be seen from the table, the compression rate differs. When going to higher-order moments, it can be expected that the size of the feature vector would become a problem. In most relevant research, this is not a commonly addressed problem. It can be argued that reducing the number of data samples in a selected manner could improve the classification results of higher-order moments. Therefore, an analysis is included in

the selection of the moments feature. Here, the emphasis lies on preserving the properties of moments.
The rest of this section analyses different feature selection methods when applied to the discussed orthogonal moments. This analysis aims to see if further feature selection, performed on the moment-based feature vector, would benefit the classification rate. The classification results using a feature selection algorithm are presented and discussed in Chapter 6

5.3.1. UNIVARIATE FEATURE RANKING FOR REGRESSION USING F-TESTS

The F-test is a statistical test that assesses the significance of the relationship between a characteristic and the target variable in a regression model. Each moment is treated as an individual feature when using univariate feature ranking with F-tests. Moments with higher scores on the F-test indicate stronger correlations with the target variable, suggesting that they have more predictive power. On that basis, the most relevant features are selected. This feature selection approach helps identify the moments that contribute the most to explaining the variance in the target variable.
For this research, the MATLAB implementation of this feature selection algorithm is used. The performance of this method depends on the number of features included and does not guarantee a better or optimal solution. The method will be used as a first indication of the effect of feature selection when using orthogonal moments. It is chosen because of the short calculation time in comparison to other commonly used feature selection methods, e.g., sequential feature selection which require many multiple assessments of different candidate feature vectors with the chosen classifier. The classification results can be found in Chapter 6.

5.3.2. UNIVARIATE FEATURE RANKING FOR CLASSIFICATION USING CHI-SQUARE TESTS

The univariate feature ranking for classification using the chi-square test [91] feature selection method is similar to the previously proposed univariate feature ranking for regression using F-tests. In this method, a chi-square test is used to estimate the dependence of individual characteristics on the outcome. This method has as advantage that it is more suitable for selecting features in classification tasks with categorical target variables, whereas F-tests are utilized for regression tasks. Both methods are used for a first indication on the influence of feature selection on orthogonal moments. Here, the function provided in MATLAB is used. The results are presented in Chapter 6.

5.4. EVALUATION METRIC AND PIPELINE

The results from the different pipelines will be compared between the different moments, classifiers, data sets, and moment orders. A measure is required to properly illustrate the classification performance of the proposed pipelines in the most uniform way. For this purpose, the F1 score is chosen as a measure. The F1 score does take the recall and precision into account and is defined as:

$$F1 = \frac{TP}{TP + \frac{1}{2}(TP + FN)} \tag{5.7}$$

TP is the true positive classification, and the FN is the false negative. For this research, the F1 score brings an advantage over other metrics, e.g. the sample size or the difference between the classes (imbalances) will not influence the performance measure. Note that the calculation given in Equation 5.7 is calculated over each class, resulting in the so-called Micro-F1 score. The Macro-F1 score, i.e., the F1 score over the whole set, is then calculated by averaging over all the classes. In the limit case of $TP = 0$, the fraction does not exist, and the micro-F1 score is interpreted as zero.

5.4.1. PROPOSED PIPELINE

To evaluate all the aspects of the classification pipeline described earlier, a revised pipeline is proposed for evaluation. This is built up of the combination of the blocks described in figure 5.2. Each block will vary between the previously discussed options, resulting in the classification results for all possible combinations. From there, the influence of the different blocks can be evaluated.

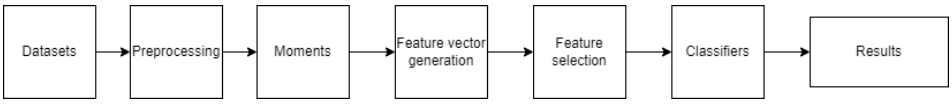


Figure 5.2: Revised proposed classification pipeline, with more specific processing blocks explored in this thesis

For clarity, the possible variations of each block are summarised below:

- **Datasets:** There are three datasets evaluated, as described in Chapter 4.
- **Preproscessing:** There are two preprocessing options evaluated, namely spectrogram and the cadence velocity diagram (CVD).
- **Moments:** There are 20 moments evaluated as described in Chapter 2, for which various parameters will be evaluated.
- **Feature vector generation:** There is one type of feature vector generation algorithm for the real-valued moments. For the complex-valued moments, there are two: including the imaginary part and without.
- **Feature selection:** Three feature selection methods are evaluated: no selection at all, CHI-square, and F-test.
- **Classifiers:** There are 13 different classifiers; a K-NN varying with k = [1,..,10] and an SVM with the kernel options of linear, quadratic and radial basis function.

The final composed pipeline will be evaluated with all possible combinations. The results are presented and discussed in Chapter 6. In addition, the pipeline is tested on the influence of added noise. Specifically, these two cases are evaluated: the case where the noise is introduced in the training and test data, and the case where the training process is done using normal data and the testing performed with the introduction of noise.

6

TESTING & RESULTS

This chapter presents the results of the proposed classification pipeline. Several variations in the pipeline and its components are investigated to fully evaluate the classification abilities of the different orthogonal moments. The results are presented and discussed component-wise, illustrating the different influences of the components on the overall results, and drawing the main take-home messages from each part of the analysis.

6.0.1. NOTE ON THE PRESENTATION OF THE RESULTS

It should be first noted that the generated results from the proposed classification pipeline are multidimensional. A visual representation of the data structure is given in figure 6.1, where the different variables to be accounted for are illustrated, namely classifier, moment order, and moment type. Note that this case is only for one pre-processing type and only for one dataset, so variability across datasets and across pre-processing will also add another dimension to evaluate the results across. Given this complex multidimensional nature of the analysis, the discussion on the results will be segmented over the different variables of interest to assess their influences.

6.1. GENERAL CLASSIFICATION RESULTS

As a step to reduce the number of possible combinations of the different aforementioned variables, the maximum value of the F1 score as performance metric will be taken across the different evaluated classifiers (described in Chapter 5) in the results reported in this section. Hence, the reported results will first look at the different moment types and order, for the different datasets. For the moments with various tunable parameters, the best-performing variation will be initially selected, and the influence of the parameter variation will be further evaluated in section 6.1.7. Moreover, the influence of the classifiers will be further discussed in Section 6.2.

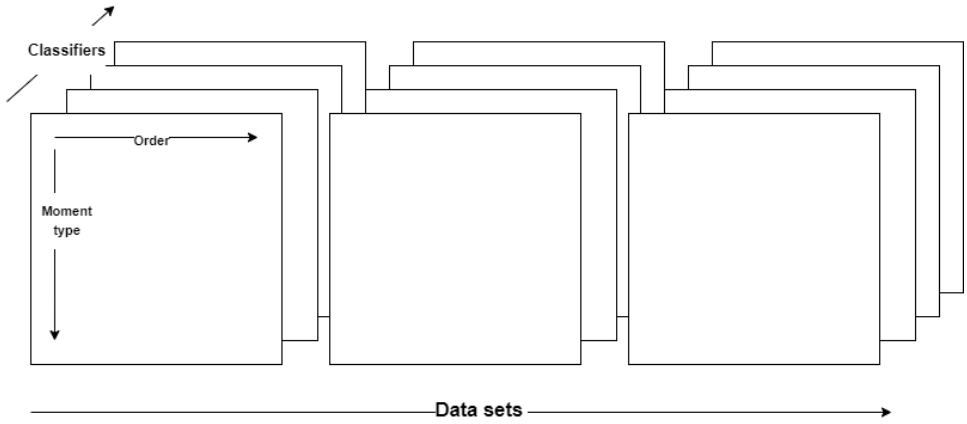


Figure 6.1: Sketch of the data structure generated by the proposed pipeline given one type of pre-processing; essentially, the variables of interest are the type of moments considered, their order, and the choice of the classifier. This can be repeated for each dataset.

6.1.1. INFLUENCE OF THE DIFFERENT DATA SETS

In table 6.1, the general classification results are presented across the three considered data sets. The maximum classification F1 score is taken for each evaluated moment. The five best performing F1 scores for each data set are highlighted in green, and the lowest five are highlighted in red. From the table, it can be observed that the relative performance of the moments fluctuates over the different datasets. The relative performance is, for this case, defined as the performance relative to the other moments applied to the same data set. Because of this fluctuation, no optimal moment in an absolute sense can be appointed to give the overall best performance in terms of classification. The moments that showed the most promising performance on average will be the CM, CHM-2nd and ZM.

Based on the findings presented in this study, it can be concluded that the performances of the moments do not exhibit significant differences. However, it is worth noting that using different classifiers tends to cause the performances to converge towards a relatively narrow range of overall performance. These results suggest that the choice of classifier may significantly impact the overall system's performance. Note that the optimal classifier changed for the different datasets for most moments; this will further be elucidated in the analysis performed in Section 6.2. Regarding overall performance under the same classifier for all datasets, RM and CHM-2nd moments will be reasonable options.

Moments	Dopnet dataset		Glasgow dataset		Xethru dataset	
	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier
PJFM	0.960	'Quadratic svm'	0.900	'RBF svm'	0.775	'RBF svm'
ZM	0.946	'Quadratic svm'	0.921	'RBF svm'	0.805	'Knn, k =1'
PZM	0.954	'Quadratic svm'	0.914	'RBF svm'	0.772	'Knn, k =1'
CHFM	0.959	'Quadratic svm'	0.883	'RBF svm'	0.766	'Knn, k =4'
OFMM	0.950	'Quadratic svm'	0.889	'RBF svm'	0.777	'Knn, k =4'
EFM	0.917	'Quadratic svm'	0.901	'RBF svm'	0.748	'RBF svm'
PCEM	0.941	'RBF svm'	0.817	'RBF svm'	0.775	'Knn, k =5'
BFM	0.952	'Knn, k =3'	0.846	'RBF svm'	0.760	'RBF svm'
RHFM	0.967	'Knn, k =4'	0.894	'RBF svm'	0.783	'RBF svm'
GM	0.949	'Knn, k =3'	0.895	'RBF svm'	0.794	'RBF svm'
LM	0.939	'Knn, k =5'	0.889	'RBF svm'	0.756	'RBF svm'
CHM-1st	0.934	'RBF svm'	0.891	'Knn, k =5'	0.772	'RBF svm'
CHM-2nd	0.968	'RBF svm'	0.878	'RBF svm'	0.808	'RBF svm'
GHM	0.971	'RBF svm'	0.871	'RBF svm'	0.787	'RBF svm'
RM max	0.967	'RBF svm'	0.880	'RBF svm'	0.810	'RBF svm'
DHM max	0.970	'RBF svm'	0.873	'Knn, k =10'	0.802	'RBF svm'
HM max	0.967	'RBF svm'	0.876	'RBF svm'	0.787	'RBF svm'
CHDM	0.945	'Knn, k =7'	0.883	'RBF svm'	0.760	'RBF svm'
KM max	0.971	'RBF svm'	0.870	'RBF svm'	0.794	'RBF svm'
MM max	0.966	'RBF svm'	0.873	'Knn, k =8'	0.834	'Knn, k =5'
CM max	0.968	'RBF svm'	0.873	'RBF svm'	0.811	'Knn, k =2'

Table 6.1: Classification results over the different datasets for different moments. The green colour indicates the top five performing and the red colour the worst five performing F1-scores

Network	Dopnet F1	Glasgow F1	XeThru F1
Googlenet	0.967	0.762	0.648
RESnet-50	0.975	0.875	0.703

Table 6.2: Classification results over the different datasets using the convolutional neural networks GoogLeNet and RESnet-50

6.1.2. COMPARISON TO NEURAL NETWORKS

A classification of the datasets using convolutional neural networks (CNN) is performed to compare with the results obtained with the proposed moments, presented in Table 6.1. The CNNs are chosen to represent conventional methods from the literature to process radar data, such as spectrograms as input images. For this analysis, GoogleNet and RESnet-50 networks are chosen, mainly because of their usage in related research [38].

The classification results are shown in Table 6.2. The best-performing approach is the RESnet-50. Comparing this result with the classification results of the moments shows that the best-performing moments outperform the CNN for the Glasgow and Xethru dataset. However, for the Glasgow dataset, the differences in F1 scores are small. For the Dopnet dataset, the RESnet-50 CNN outperforms all moments. Again, the difference in the F1 score is small.
Summarizing, the comparison of Table 6.1 and Table 6.2 shows that the investigated moments are similar to the two investigated CNNs in terms of performance for the Dopnet and Glasgow datasets, and even better performing for the more complex Xethru dataset. In the broader context, it should be noted that CNNs tend to have higher computational complexity, which consequently makes the utilization of orthogonal moments more advantageous for classification tasks of this nature. The rigorous comparative analysis of algorithmic complexities can serve as a useful future research diretion to expand from this work.

6.1.3. ASSESSMENT OF MOMENTS DEFINED ON A DISK

To further evaluate how the moments behave with respect to their order, they will be evaluated for each family of moments due to the expected similarities in behavior. This will being with the family of continuous moments defined on a disc. Table 6.3 shows the results in terms of moment order for the Dopnet dataset. The results for this data set illustrate the general behavior; the results for the other data sets can be found in Appendix B for completeness.
The table shows that the optimum values are found at low orders centered around order ten, and that the lower orders generally provide high classification scores. Furthermore, it can be observed that, at high orders, the scores of the moments that use the SVM classifier with quadratic kernel drop in the F1 score. This can be explained as a result of the larger feature vector size on higher orders.
The findings of this study suggest that the use of this family of moments produces favorable outcomes at lower orders. This is a desirable characteristic, as lower orders typically necessitate fewer computational resources.

6.1.4. PROPOSED FEATURE VECTOR FOR COMPLEX MOMENTS

In Section 5.1.1, a feature vector construction method is proposed for moments that contain complex values. This holds for all the moments defined on a disk. The proposed method is tested and compared to the method used in previous research based on the absolute value of the moments, i.e. real-valued moments. The results are presented in Table 6.4, where 'old method' refers to the conventional usage of real-valued moments, and 'new method' to the proposed complex-valued feature vector.

The results show a clear improvement using the proposed method for all the complex moments evaluated. This shows that adding the phase information is beneficial to the classification abilities when classifying micro-Doppler signatures. This research uses this complex-valued feature construction method for all the tests involving complex moments.

Moment	Max F1 score	Best Classifier	F1 score							
			n=5	n=10	n=15	n=20	n=25	n=30	n=40	n=50
PJFM	0.960	'Quadratic svm'	0.960	0.903	0.138	0.117	0.122	0.665	0.563	0.546
ZM	0.946	'Quadratic svm'	0.817	0.920	0.946	0.263	0.118	0.338	0.547	0.609
PZM	0.954	'Quadratic svm'	0.893	0.954	0.324	0.142	0.117	0.236	0.166	0.150
CHFM	0.959	'Quadratic svm'	0.959	0.925	0.113	0.117	0.113	0.207	0.663	0.522
OFMM	0.950	'Quadratic svm'	0.950	0.922	0.113	0.113	0.113	0.113	0.589	0.510
EFM	0.917	'Quadratic svm'	0.917	0.594	0.102	0.097	0.113	0.117	0.117	0.227
PCEM	0.941	'RBF svm'	0.788	0.941	0.794	0.617	0.483	0.409	0.364	0.364
BFM	0.952	'Knn, k =3'	0.922	0.952	0.846	0.675	0.626	0.637	0.637	0.665
RHFM	0.967	'Knn, k =4'	0.907	0.967	0.744	0.623	0.556	0.521	0.467	0.456

Table 6.3: Classification results of the continuous moments defined on a disk as a function of their order, when applied to the Dopnet dataset

6.1.5. ASSESSMENT OF CONTINUOUS MOMENTS DEFINED ON A SQUARE GRID

Table 6.5 shows the results of the continuous moments defined on a square grid, with respect to their order when applied to the Dopnet dataset. The results when applied to the other datasets, can be found in Appendix B for completeness.

The study shows that the spread in optimal order is wider compared to the continuous moments defined on a disk analyzed in section 6.1.3, and on average these moments show optimum values at higher orders.

6.1.6. ASSESSMENT OF DISCRETE MOMENTS DEFINED ON A SQUARE GRID

Table 6.6 shows the results of the discrete moments defined on a square grid as a function of their order, when applied to the Dopnet dataset. When applied to the other data sets discussed, the results can be found in Appendix B for completeness. One must mention that the moments evaluated as a function of their various parameters are presented in this table by taking the combination providing the best performances.

The table shows that the discrete moments find their optimum values at the highest orders compared to the other families of moments. Specifically, the moments that do not capture the whole input image at low orders, for example RM, DHM, HM, KM, MM, CM, show a required higher order at which the optimum is found, compared to the moments that capture the whole image at lower orders, for example, CHDM. This is explained by the nature of the data sets described in Chapter 4, where the variance plots over the different classes illustrate the region defining the class-related information. This showed that valuable information is present over the whole horizontal length of the image. This information is not captured at low order for the moments that expand from a center point outwards. As the result shows, higher orders are needed to capture the class-defining information, potentially resulting in additional complexity.

Moments	Dopnet dataset				Glasgow dataset			
	New method		Old method		New method		Old method	
	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier
PJFM	0.960	'Quadratic svm'	0.908	'RBF svm'	0.900	'RBF svm'	0.831	'RBF svm'
ZM	0.946	'Quadratic svm'	0.908	'RBF svm'	0.921	'RBF svm'	0.768	'RBF svm'
PZM	0.954	'Quadratic svm'	0.899	'RBF svm'	0.914	'RBF svm'	0.798	'RBF svm'
CHFM	0.959	'Quadratic svm'	0.914	'RBF svm'	0.883	'RBF svm'	0.827	'RBF svm'
OFMM	0.950	'Quadratic svm'	0.908	'RBF svm'	0.889	'RBF svm'	0.831	'RBF svm'
EFM	0.917	'Quadratic svm'	0.844	'RBF svm'	0.901	'RBF svm'	0.858	'RBF svm'
PCEM	0.941	'RBF svm'	0.901	'RBF svm'	0.817	'RBF svm'	0.859	'RBF svm'
BFM	0.952	'Knn, k =3'	0.887	'Quadratic svm'	0.846	'RBF svm'	0.794	'RBF svm'
RHFM	0.967	'Knn, k =4'	0.925	'RBF svm'	0.894	'RBF svm'	0.828	'RBF svm'

Table 6.4: F1 scores of the continuous moments defined on a disk when applied to the Dopnet and Glasgow datasets. Two feature vectors are tested, based on the old method (i.e., real-valued moments computed via the absolute value) and the newly proposed method (i.e., complex-valued moments).

Moment	Max F1 score	Best Classifier	F1 score								
			n=5	n=10	n=15	n=20	n=25	n=30	n=35	n=40	n=50
GM	0.949	'Knn, k =3'	0.887	0.928	0.949	0.918	0.916	0.891	0.797	0.761	0.648
LM	0.939	'Knn, k =5'	0.841	0.925	0.939	0.910	0.814	0.751	0.690	0.621	0.540
CHM-1st	0.934	'RBF svm'	0.544	0.748	0.798	0.905	0.929	0.934	0.897	0.734	0.620
CHM-2nd	0.968	'RBF svm'	0.667	0.727	0.768	0.816	0.908	0.955	0.967	0.968	0.959
GHM	0.971	'RBF svm'	0.607	0.726	0.826	0.933	0.963	0.967	0.971	0.968	0.946

Table 6.5: Classification results of the continuous moments defined on a square grid as a function of their order, when applied to the Dopnet dataset

Moment	Max F1 score	Best Classifier	F1 score									
			n=5	n=10	n=15	n=20	n=25	n=30	n=35	n=40	n=45	n=50
RM max	0.967	'RBF svm'	0.675	0.691	0.800	0.900	0.962	0.965	0.967	0.965	0.965	0.951
DHM max	0.970	'RBF svm'	0.548	0.703	0.861	0.947	0.956	0.967	0.970	0.944	0.908	0.617
HM max	0.967	'RBF svm'	0.610	0.753	0.770	0.837	0.922	0.953	0.967	0.962	0.959	0.952
CHDM	0.945	'Knn, k =7'	0.846	0.925	0.945	0.885	0.835	0.759	0.696	0.645	0.592	0.592
KM max	0.971	'RBF svm'	0.637	0.733	0.831	0.908	0.939	0.959	0.959	0.968	0.971	0.944
MM max	0.966	'RBF svm'	0.414	0.621	0.707	0.776	0.895	0.942	0.958	0.962	0.959	0.966
CM max	0.968	'RBF svm'	0.672	0.714	0.835	0.895	0.962	0.965	0.965	0.968	0.962	0.961

Table 6.6: Classification results for the discrete moments defined on a square grid when applied to the Dopnet dataset

6.1.7. EFFECT OF VARIATION OF MOMENT PARAMETERS

Several moments of the family of discrete moments defined on a square are evaluated for different parameters in their basis functions. The exact parameter values are described also in Chapter 3, where the moments were first introduced. Classification F1 scores for

the variety of Racah moments, applied to the Dopnet data set, are presented in Figure 6.2. The scores are plotted for each case against the order of the moment. It is shown that the cases where the expansion of the ROI covered by the moment starts away from the vertical center perform worse at low orders. When the order is increased, the ROI of the moments expands and the classification results converge to the same F1 score.
This illustrates that selecting parameters for the Racah moment is only important when high classification rates are requested on low orders. At high orders, the influence of the parameter selection regarding the classification score will be little.
This behavior can be seen for all moments shown in figure 6.3, where the results of the

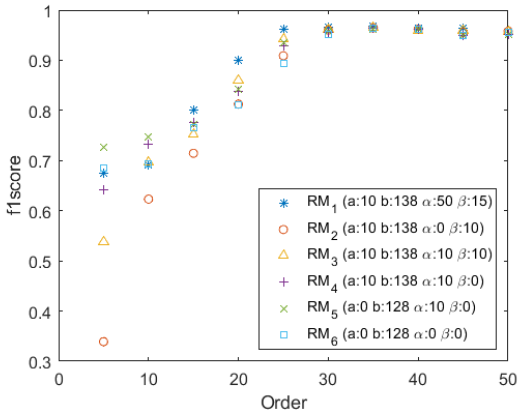
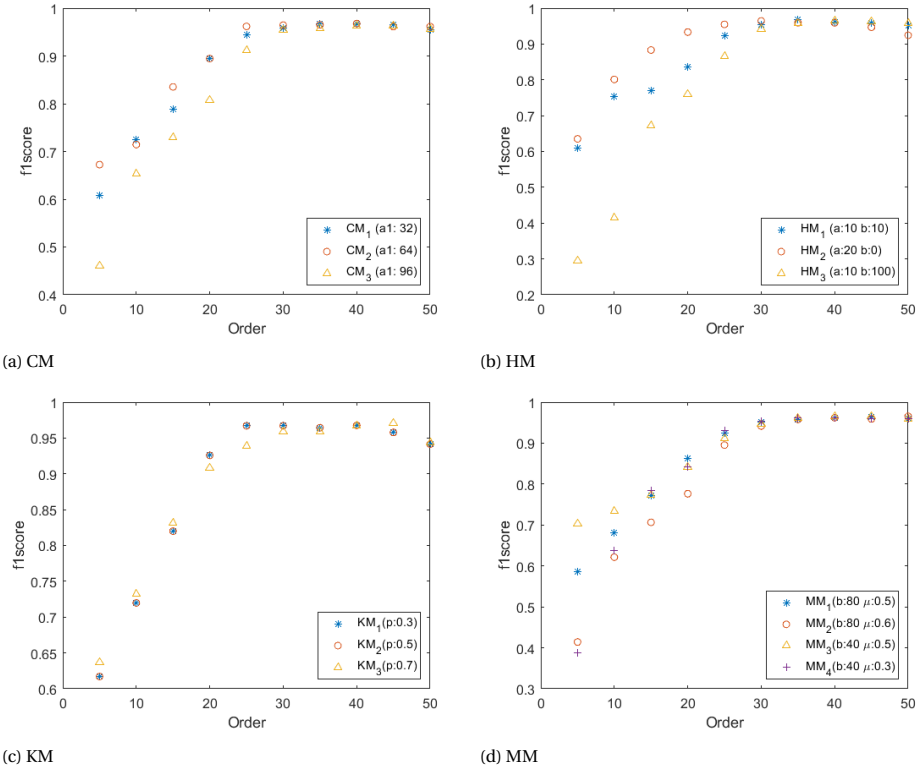


Figure 6.2: F1-score results as a function of the different variations in parameters of the Racah moment, using an SVM with an RBF kernel and the Dopnet data set.

parameter variation are shown for CM, HM, KM, and MM when applied to the Dopnet dataset. These results show that parameter selection can increase classification performances at low orders and that at high orders, the different varieties converge, making it less valuable to tune the parameters.



(a) CM (b) HM

(c) KM (d) MM

Figure 6.3: The F1 scores of the CM, HM, KM, and MM for different parameter selections when classifying the Dop-net dataset. This illustrates the convergence to the same scores at higher orders, making the parameters tuning less valuable.

Further results for the Dual Hahn moments are shown in Figure 6.4; again, it is shown that the performance at low order differs in F1-scores for the different parameter variations, and when the order increases, the F1-scores converge. After order 30, the results show a rapidly decreasing performance for DHM_1. In contrast to the previously observed trends, this appears to indicate that the parameter selection does matter to avoid decreasing performance for high moments orders, after the maximal F1-score is reached.

6.1.8. OVERVIEW EVALUATION OF REQUIRED MOMENT ORDER

The results presented so far looked at the influence of the order of the moments for the different families of moments. It was demonstrated that continuous moments defined on a disk need the lowest order of moments to obtain their maximal classification score. Further, it was shown that moments that expand in their ROI need higher orders to obtain the maximum classification rate. This matches the hypotheses of the expected ROI presented in Chapter 4, where it was illustrated that valuable information is spread over the entire horizontal axis of the data. Only at higher orders will the ROI cover this whole

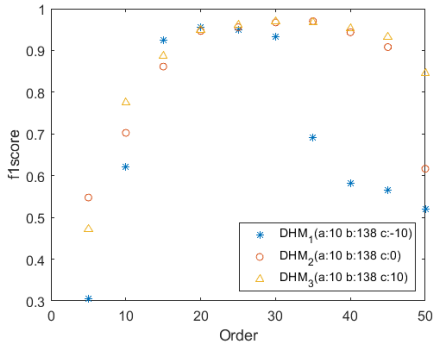


Figure 6.4: Classification results as a function of the different variations in parameters of the Dual-Hahn moment, using an SVM with an RBF kernel and the Dopnet data set

axis for all moments. This is shown in the convergence to the same classification rates at higher orders, irrespective of the type of moments considered.

6.2. CLASSIFIER INFLUENCE

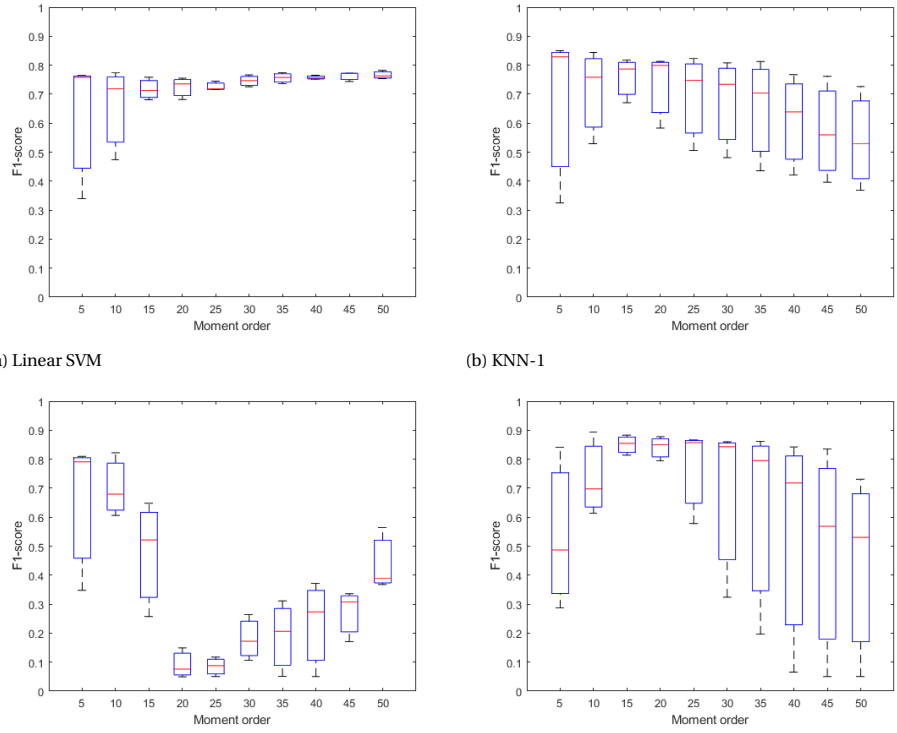
The type of chosen classifier influences the classification results. As it can be seen from the results presented above, the maximum F1 scores are obtained by different classifiers for the different moments. Table 6.7 presents the difference in the F1 scores obtained using the Glasgow dataset. The highest- and lowest-scoring classifiers are highlighted, showing that, on average, an SVM with a radial basis function performs best on this dataset. In addition, the table shows that the influence of the classifier can be considered significant, with about 0.1 difference in the F1 scores between the best and worst-performing classifiers. The results of the other data sets are presented in Appendix B for completeness. They similarly show widespread performance over the different classifiers.

The classifiers' performance shows a dependency on the order of the moments. Figure 6.5 presents the box plots for the results of all the different classifiers used, with respect to the moment order when using the Glasgow data set. It can be seen that the classifiers show different behaviors. For instance, the SVM-RBF has the highest values of F1 scores at orders 15 and 20, and becomes more unstable when dealing with lower orders, whereas the KNN shows the highest average performances. These plots show that the linear SVM becomes the most stable and the highest-performing classifier at the higher orders.

In general, the results show that the classification performance of the moments is highly dependent on the type of classifier and order. The classifier that shows on average the best classification results is the SVM with a radial basis function as a kernel function. A comment has to be made that this only would be the case with orders 15 to 20. When dealing with lower orders, KNN shows better performance; at higher orders, the SVM with linear kernel becomes the most favorable classifier.

Moments	RBF SVM	Quadratic SVM	Linear SVM	KNN-1	KNN-5	KNN-10	Difference
PJFM	0.900	0.839	0.807	0.836	0.822	0.817	0.093
ZM	0.921	0.836	0.790	0.814	0.805	0.826	0.132
PZM	0.914	0.857	0.770	0.849	0.839	0.843	0.143
CHFM	0.883	0.848	0.800	0.812	0.819	0.827	0.083
OFMM	0.889	0.834	0.814	0.821	0.820	0.804	0.084
EFM	0.901	0.781	0.808	0.829	0.814	0.821	0.120
PCEM	0.817	0.748	0.784	0.768	0.748	0.765	0.068
BFM	0.846	0.811	0.777	0.811	0.827	0.809	0.069
RHFM	0.894	0.810	0.771	0.850	0.814	0.814	0.123
GM	0.895	0.827	0.764	0.842	0.852	0.865	0.132
LM	0.889	0.839	0.779	0.847	0.865	0.859	0.110
CHM-1st	0.864	0.847	0.803	0.860	0.891	0.886	0.088
CHM-2nd	0.878	0.822	0.774	0.844	0.858	0.869	0.104
GHM	0.871	0.736	0.814	0.800	0.819	0.810	0.134
RM max	0.880	0.817	0.814	0.789	0.835	0.828	0.091
DHM max	0.857	0.745	0.789	0.830	0.859	0.873	0.128
HM max	0,876	0,817	0,782	0,847	0,864	0,863	0,094
CHDM	0.883	0.850	0.780	0.855	0.872	0.859	0.103
KM max	0.870	0.705	0.782	0.842	0.848	0.844	0.164
MM max	0.870	0.807	0.793	0.791	0.845	0.839	0.079
CM max	0.873	0.754	0.786	0.796	0.800	0.796	0.119

Table 6.7: F1-scores for different moments when different classifiers are used to classify the Glasgow dataset. The difference in F1-score between the best-performing (green) and worst-performing (red) classifier for each moment is also included.



(a) Linear SVM

(b) KNN-1

(c) Quadratic SVM

(d) RBF SVM

Figure 6.5: Boxplot of the F1 score values using all the moments considered in this thesis. This illustrates the influence of the moment order on the used classifiers.

6.3. ASSESSMENT OF PERFORMANCE OF THE CVD

The influence of using CVD as a type of preprocessing and data formatting is tested in terms of classification rates. Table 6.8 presents the maximum F1 scores obtained by the moments evaluated applied to the Dopnet data set using CVD and spectrograms. The F1 scores for the spectrogram case are shown as a reference, and the maximum value is highlighted in gray. This indicates that the CVD performs worse in most cases than the spectrogram. Only in the cases of EFM and PCEM, the F1 score shows a slight increase. These results show that the usage of the CVD in this format can be questioned. The technique does not appear to increase the classification results and generates more computational complexity in the classification pipeline.

Moments	Spectrogram		CVD	
	Max F1	Best Classifier	Max F1	Best Classifier
PJFM	0.960	'Quadratic svm'	0.885	'Knn, k =3'
ZM	0.946	'Quadratic svm'	0.838	'linear svm'
PZM	0.954	'Quadratic svm'	0.858	'linear svm'
CHFM	0.959	'Quadratic svm'	0.897	'Knn, k =1'
OFMM	0.950	'Quadratic svm'	0.873	'RBF svm'
EFM	0.917	'Quadratic svm'	0.928	'RBF svm'
PCEM	0.941	'RBF svm'	0.951	'RBF svm'
BFM	0.952	'Knn, k =3'	0.875	'Knn, k =6'
RHFM	0.967	'Knn, k =4'	0.866	'Knn, k =8'
GM	0.949	'Knn, k =3'	0.934	'RBF svm'
LM	0.939	'Knn, k =5'	0.927	'RBF svm'
CHM-1st	0.934	'RBF svm'	0.921	'RBF svm'
CHM-2nd	0.968	'RBF svm'	0.940	'RBF svm'
GHM	0.971	'RBF svm'	0.934	'RBF svm'
RM max	0.967	'RBF svm'	0.948	'RBF svm'
DHM max	0.970	'RBF svm'	0.940	'RBF svm'
HM max	0.967	'RBF svm'	0.948	'RBF svm'
CHDM	0.945	'Knn, k =7'	0.934	'RBF svm'
KM max	0.971	'RBF svm'	0.943	'RBF svm'
MM max	0.966	'RBF svm'	0.948	'RBF svm'
CM max	0.968	'RBF svm'	0.946	'RBF svm'

Table 6.8: Classification results comparison for the spectrogram vs CVD used as pre-processing techniques classifying the DOP-net dataset

6.3.1. ASSESSMENT OF THE INFLUENCE OF NOISE INTRODUCTION

The classification performance of the moments is tested with respect to the noise influence. The disturbance in the data is simulated by adding zero-mean Gaussian noise to the spectrogram. Note that this procedure does not have direct physical meaning. It is similar to the noise analysis techniques proposed in image processing [9] as explained in Chapter 4. The classification pipeline is subjected to noise in two ways; first, where the classifier is trained and tested on the noisy data, and second, a case where the classifier is trained on clean data and tested on data with introduced noise.

NOISE IN TRAINING AND VALIDATION DATA

The results of the case where the training and testing data are subjected to noise are shown in Table 6.9 for the Glasgow data set, and Table 6.10 for the Dopnet data set. Noise is added as zero-mean Gaussian noise with variances 0.1 0.2, and 0.3, respectively. It can be observed from the tables that the increases in noise variance cause the F1-scores to drop as the data get more disrupted. The orthogonal moment-based classification pipeline is compared with two neural networks. Table 6.11 presents the classification results of these two neural networks using the Glasgow data set. Comparing these results shows that the best-performing moments and best-performing neural networks have similar, comparable performances.

Moments	Spectogram		0.1		0.2		0.3	
	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier
PJFM	0.900	'RBF svm'	0.794	'Quadratic svm'	0.763	'Quadratic svm'	0.748	'Quadratic svm'
ZM	0.921	'RBF svm'	0.742	'linear svm'	0.728	'linear svm'	0.714	'linear svm'
PZM	0.914	'RBF svm'	0.738	'linear svm'	0.716	'linear svm'	0.703	'linear svm'
CHFM	0.883	'RBF svm'	0.781	'Quadratic svm'	0.771	'Quadratic svm'	0.784	'Quadratic svm'
OFMM	0.889	'RBF svm'	0.794	'Quadratic svm'	0.777	'Quadratic svm'	0.763	'Quadratic svm'
EFM	0.901	'RBF svm'	0.749	'Knn, k =8'	0.710	'Knn, k =10'	0.710	'Knn, k =9'
PCEM	0.817	'RBF svm'	0.731	'Quadratic svm'	0.713	'Quadratic svm'	0.674	'Quadratic svm'
BFM	0.846	'RBF svm'	0.751	'Quadratic svm'	0.747	'Quadratic svm'	0.736	'Quadratic svm'
RHFM	0.894	'RBF svm'	0.784	'Quadratic svm'	0.773	'Quadratic svm'	0.774	'Quadratic svm'
GM	0.895	'RBF svm'	0.777	'RBF svm'	0.759	'Knn, k =7'	0.715	'Knn, k =9'
LM	0.889	'RBF svm'	0.754	'RBF svm'	0.738	'linear svm'	0.684	'Knn, k =10'
CHM-1st	0.891	'Knn, k =5'	0.783	'RBF svm'	0.734	'RBF svm'	0.712	'RBF svm'
CHM-2nd	0.878	'RBF svm'	0.810	'RBF svm'	0.787	'RBF svm'	0.776	'RBF svm'
GHM	0.871	'RBF svm'	0.818	'RBF svm'	0.775	'RBF svm'	0.757	'RBF svm'
RM max	0.880	'RBF svm'	0.819	'RBF svm'	0.782	'RBF svm'	0.752	'RBF svm'
DHM max	0.873	'Knn, k =10'	0.823	'RBF svm'	0.804	'RBF svm'	0.786	'RBF svm'
HM max	0.876	'RBF svm'	0.815	'RBF svm'	0.779	'RBF svm'	0.766	'Knn, k =8'
CHDM	0.883	'RBF svm'	0.770	'RBF svm'	0.743	'linear svm'	0.704	'Knn, k =10'
KM max	0.870	'RBF svm'	0.813	'RBF svm'	0.779	'RBF svm'	0.749	'RBF svm'
MM max	0.873	'Knn, k =8'	0.813	'RBF svm'	0.792	'RBF svm'	0.779	'RBF svm'
CM max	0.873	'RBF svm'	0.821	'RBF svm'	0.785	'RBF svm'	0.778	'RBF svm'

Table 6.9: Classification results using the Glasgow dataset when noise is introduced in the training and test data. Noise is introduced as a zero-mean Gaussian with σ variance of 0.1, 0.2 and 0.3

Moments	Spectogram		0.1		0.2		0.3	
	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier
PJFM	0.960	'Quadratic svm'	0.926	'Quadratic svm'	0.914	'Quadratic svm'	0.916	'Quadratic svm'
ZM	0.946	'Quadratic svm'	0.852	'RBF svm'	0.844	'RBF svm'	0.840	'RBF svm'
PZM	0.954	'Quadratic svm'	0.781	'RBF svm'	0.785	'RBF svm'	0.768	'RBF svm'
CHFM	0.959	'Quadratic svm'	0.916	'Quadratic svm'	0.907	'Quadratic svm'	0.911	'Quadratic svm'
OFMM	0.950	'Quadratic svm'	0.906	'Quadratic svm'	0.907	'Quadratic svm'	0.913	'Quadratic svm'
EFM	0.917	'Quadratic svm'	0.873	'RBF svm'	0.861	'RBF svm'	0.837	'RBF svm'
PCEM	0.941	'RBF svm'	0.850	'RBF svm'	0.830	'Knn, k =4'	0.839	'Knn, k =4'
BFM	0.952	'Knn, k =3'	0.900	'Quadratic svm'	0.896	'Knn, k =3'	0.910	'Quadratic svm'
RHFM	0.967	'Knn, k =4'	0.923	'Quadratic svm'	0.916	'Quadratic svm'	0.913	'Knn, k =2'
GM	0.949	'Knn, k =3'	0.914	'Knn, k =1'	0.865	'Knn, k =3'	0.865	'Knn, k =5'
LM	0.939	'Knn, k =5'	0.907	'RBF svm'	0.870	'RBF svm'	0.859	'RBF svm'
CHM-1st	0.934	'RBF svm'	0.922	'RBF svm'	0.878	'RBF svm'	0.841	'RBF svm'
CHM-2nd	0.968	'RBF svm'	0.887	'RBF svm'	0.872	'Knn, k =5'	0.869	'Knn, k =5'
GHM	0.971	'RBF svm'	0.884	'RBF svm'	0.840	'RBF svm'	0.774	'RBF svm'
RM max	0.967	'RBF svm'	0.896	'Knn, k =4'	0.903	'Knn, k =5'	0.898	'Knn, k =5'
DHM max	0.970	'RBF svm'	0.930	'RBF svm'	0.916	'RBF svm'	0.900	'RBF svm'
HM max	0.967	'RBF svm'	0.914	'Knn, k =7'	0.914	'Knn, k =3'	0.905	'Knn, k =7'
CHDM	0.945	'Knn, k =7'	0.910	'RBF svm'	0.878	'RBF svm'	0.878	'RBF svm'
KM max	0.971	'RBF svm'	0.882	'RBF svm'	0.843	'RBF svm'	0.777	'RBF svm'
MM max	0.966	'RBF svm'	0.921	'Knn, k =3'	0.917	'Knn, k =3'	0.912	'Knn, k =2'
CM max	0.968	'RBF svm'	0.911	'RBF svm'	0.882	'RBF svm'	0.814	'RBF svm'

Table 6.10: Classification results using the Dopnet dataset when noise is introduced in the training and test data. Noise is introduced as a zero-mean Gaussian with σ variance of 0.1, 0.2 and 0.3

NOISE IN VALIDATION DATA

The results for the Glasgow dataset, where noise is only added to the test set but not to the training set, are shown in Table 6.12. In the presented table, the top five performing moments are highlighted for each case, i.e. for the original case and for the three cases where noise has been added. The classification results show that the moments that had

Network	Gaussian white noise (variance)			
	0	**0.1**	**0.2**	**0.3**
Googlenet	0.762	0.753	0.692	0.716
RESnet-50	0.876	0.841	0.798	0.699

Table 6.11: Classification results using the Glasgow dataset and a neural network as a classifier, when noise is introduced in the training and test data. Noise is introduced as a zero mean Gaussian with σ variance of 0.1, 0.2 and 0.3

their optimum values at lower orders (see Sections 6.1.3-6.1.6) perform better under the influence of noise in their test set compared to the other moments. This behavior can be explained by the low-pass filtering behavior of the lower-order moments. The lower-order basis functions will be less affected by the high-frequency components of the white noise and, therefore, be more robust. The moments that show best performances for this type of noise introduction are the CHDM, PCEM,OFMM, PJFM and RHFM.

Moments	Spectrogram		Gaussian white noise (variance)					
			0.1		0.2		0.3	
	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier
PJFM	0.900	'RBF svm'	0.831	'RBF svm'	0.716	'RBF svm'	0.637	'RBF svm'
ZM	0.921	'RBF svm'	0.618	'Knn, k =1'	0.531	'Knn, k =1'	0.438	'Knn, k =1'
PZM	0.914	'RBF svm'	0.599	'RBF svm'	0.511	'Knn, k =1'	0.454	'RBF svm'
CHFM	0.883	'RBF svm'	0.796	'RBF svm'	0.683	'RBF svm'	0.595	'RBF svm'
OFMM	0.889	'RBF svm'	0.838	'RBF svm'	0.698	'RBF svm'	0.634	'RBF svm'
EFM	0.901	'RBF svm'	0.321	'linear svm'	0.299	'linear svm'	0.227	'linear svm'
PCEM	0.817	'RBF svm'	0.823	'RBF svm'	0.737	'RBF svm'	0.671	'RBF svm'
BFM	0.846	'RBF svm'	0.731	'RBF svm'	0.648	'RBF svm'	0.578	'RBF svm'
RHFM	0.894	'RBF svm'	0.829	'RBF svm'	0.684	'RBF svm'	0.625	'RBF svm'
GM	0.895	'RBF svm'	0.687	'Knn, k =1'	0.441	'Knn, k =3'	0.377	'Knn, k =3'
LM	0.889	'RBF svm'	0.768	'Knn, k =1'	0.569	'Knn, k =1'	0.429	'Knn, k =1'
CHM-1st	0.891	'Knn, k =5'	0.326	'linear svm'	0.276	'linear svm'	0.262	'linear svm'
CHM-2nd	0.878	'RBF svm'	0.364	'linear svm'	0.288	'linear svm'	0.287	'linear svm'
GHM	0.871	'RBF svm'	0.374	'Knn, k =7'	0.325	'Knn, k =7'	0.301	'Knn, k =9'
RM max	0.880	'RBF svm'	0.353	'Knn, k =8'	0.333	'Quadratic svm'	0.318	'linear svm'
DHM max	0.873	'Knn, k =10'	0.480	'Knn, k =8'	0.454	'Knn, k =10'	0.459	'Knn, k =8'
HM max	0.876	'RBF svm'	0.391	'Knn, k =6'	0.341	'Knn, k =7'	0.328	'Knn, k =9'
CHDM	0.883	'RBF svm'	0.997	'Knn, k =1'	0.979	'Knn, k =1'	0.881	'Knn, k =1'
KM max	0.870	'RBF svm'	0.382	'Knn, k =7'	0.365	'Knn, k =2'	0.344	'RBF svm'
MM max	0.873	'Knn, k =8'	0.400	'Knn, k =1'	0.392	'Knn, k =2'	0.395	'Knn, k =2'
CM max	0.873	'RBF svm'	0.492	'Knn, k =5'	0.450	'Knn, k =6'	0.397	'Knn, k =6'

Table 6.12: Classification results using the Glasgow dataset when noise is introduced only in the test data. Noise is introduced as a zero-mean Gaussian with σ variance of 0.1,0.2 and 0.3. The best-performing moments are highlighted for each scenario.

Network	Gaussian white noise (variance)			
	0	0.1	0.2	0.3
Googlenet	0.762	0.258	0.0501	0.0502
RESnet-50	0.876	0.163	0.0712	0.0502

Table 6.13: Classification results using the Glasgow dataset and a neural network as a classifier, when noise is introduced only in test data. Noise is introduced as a zero-mean Gaussian with σ variance of 0.1,0.2 and 0.3

6.4. CLASSIFICATION RESULTS WITH FEATURE SELECTION

The earlier proposed F-test- and chi-square-based methods are used to investigate the potential of feature selection in the feature set of the generated orthogonal moments. The Dopnet data set is used to show its influence. The results are shown in Table 6.14 and 6.15. The results show that when using feature selection for a feature vector length of 100 and 200, the classification results will be similar, where some moments see a small increase in F1-score. Although these results do not show a clear increase in performance, they do show that the feature selection could be beneficial, as fewer moments may need to be calculated. This could potentially speed up the classification pipeline and reduce

Moments	Unbounded		nf = 100		nf = 200	
	Max F1	Best Classifier	Max F1	Best Classifier	Max F1	Best Classifier
PJFM	0.960	'Quadratic svm'	0.970	'Quadratic svm'	0.950	'Quadratic svm'
ZM	0.946	'Quadratic svm'	0.928	'RBF svm'	0.923	'RBF svm'
PZM	0.954	'Quadratic svm'	0.920	'RBF svm'	0.922	'RBF svm'
CHFM	0.959	'Quadratic svm'	0.962	'Quadratic svm'	0.962	'Quadratic svm'
OFMM	0.950	'Quadratic svm'	0.963	'Quadratic svm'	0.959	'Quadratic svm'
EFM	0.917	'Quadratic svm'	0.914	'Quadratic svm'	0.940	'RBF svm'
PCEM	0.941	'RBF svm'	0.945	'Knn, k =3'	0.941	'Quadratic svm'
BFM	0.952	'Knn, k =3'	0.954	'Quadratic svm'	0.954	'Quadratic svm'
RHFM	0.967	'Knn, k =4'	0.956	'Quadratic svm'	0.956	'Quadratic svm'
GM	0.949	'Knn, k =3'	0.960	'Knn, k =7'	0.963	'Knn, k =3'
LM	0.939	'Knn, k =5'	0.934	'Knn, k =3'	0.942	'Knn, k =1'
CHM-1st	0.934	'RBF svm'	0.948	'Knn, k =3'	0.954	'RBF svm'
CHM-2nd	0.968	'RBF svm'	0.954	'Knn, k =3'	0.960	'RBF svm'
GHM	0.971	'RBF svm'	0.909	'Quadratic svm'	0.933	'Knn, k =5'
RM max	0.967	'RBF svm'	0.943	'Quadratic svm'	0.966	'RBF svm'
DHM max	0.970	'RBF svm'	0.958	'Knn, k =3'	0.958	'Knn, k =3'
HM max	0.967	'RBF svm'	0.945	'Knn, k =3'	0.955	'Knn, k =1'
CHDM	0.945	'Knn, k =7'	0.935	'RBF svm'	0.979	'Knn, k =1'
KM max	0.971	'RBF svm'	0.938	'Knn, k =1'	0.942	'Knn, k =1'
MM max	0.966	'RBF svm'	0.949	'Knn, k =1'	0.949	'Knn, k =1'
CM max	0.968	'RBF svm'	0.945	'Quadratic svm'	0.955	'Quadratic svm'

Table 6.14: Classification results using F-tests based feature selection, applied to the Dopnet dataset where nf denotes the feature vector size. The best classification value is highlighted for each moment

its complexity.

Moments	Unbounded		nf = 100		nf = 200	
	Max F1	**Best Classifier**	**Max F1**	**Best Classifier**	**Max F1**	**Best Classifier**
PJFM	0.900	'RBF svm'	0.872	'RBF svm'	0.898	'RBF svm'
ZM	0.921	'RBF svm'	0.814	'RBF svm'	0.881	'RBF svm'
PZM	0.914	'RBF svm'	0.883	'RBF svm'	0.883	'RBF svm'
CHFM	0.883	'RBF svm'	0.873	'RBF svm'	0.902	'Knn, k =8'
OFMM	0.889	'RBF svm'	0.874	'RBF svm'	0.894	'RBF svm'
EFM	0.901	'RBF svm'	0.905	'RBF svm'	0.907	'RBF svm'
PCEM	0.817	'RBF svm'	0.829	'RBF svm'	0.840	'RBF svm'
BFM	0.846	'RBF svm'	0.864	'Knn, k =6'	0.901	'RBF svm'
RHFM	0.894	'RBF svm'	0.872	'RBF svm'	0.880	'RBF svm'
GM	0.895	'RBF svm'	0.898	'Knn, k =6'	0.904	'Knn, k =8'
LM	0.889	'RBF svm'	0.912	'Knn, k =10'	0.889	'RBF svm'
CHM-1st	0.891	'Knn, k =5'	0.899	'Knn, k =6'	0.891	'Knn, k =5'
CHM-2nd	0.878	'RBF svm'	0.887	'Knn, k =6'	0.879	'Knn, k =8'
GHM	0.871	'RBF svm'	0.841	'RBF svm'	0.864	'RBF svm'
RM max	0.880	'RBF svm'	0.888	'RBF svm'	0.882	'Knn, k =6'
DHM max	0.873	'Knn, k =10'	0.886	'Knn, k =6'	0.879	'Knn, k =6'
HM max	0.876	'RBF svm'	0.887	'Knn, k =9'	0.874	'RBF svm'
CHDM	0.883	'RBF svm'	0.907	'Knn, k =10'	0.902	'Knn, k =8'
KM max	0.870	'RBF svm'	0.857	'Knn, k =9'	0.881	'RBF svm'
MM max	0.873	'Knn, k =8'	0.873	'RBF svm'	0.873	'RBF svm'
CM max	0.873	'RBF svm'	0.870	'Knn, k =7'	0.891	'RBF svm'

Table 6.15: Classification results using chi-square tests based feature selection, applied to the Glasgow data set where nf denotes the feature vector size. The best classification value is highlighted for each moment

6.5. SUMMARY OF RESULTS

The results presented in this chapter show the influence of the data sets on the investigated moments' classification abilities. It is shown that the classification rates of the moments vary as a function of the dataset, and that different moments are optimal for different datasets, whereas CM, CHM-2nd and ZM exhibit the strongest performance across all evaluated data sets. For the case where the classifier is kept constant, RM and CHM-2nd demonstrate promising results in terms of overall performance with the same classifier across all datasets.

Further, it is shown that the moments defined on a disk perform better at lower orders compared to the other moments. In addition, this family of moments showed the best results for noise robustness when noise is introduced in the validation set while the training is done on clean data. Also, for this case, it is shown that the proposed moments provide better performances compared to the evaluated convolutional neural networks. The moments with free parameters showed to converge to the same classification rates when computed for high orders. At lower orders, the classification ability depends on the region of the original input radar data that the moment ROI captures, where the middle part of the spectrogram around the 0 Hz region is proven to be important for micro-Doppler signature classification. The CVD showed no clear increase in classification rate compared to spectrograms. This result shows that the usage of the CVD in these classification problems can be questioned since no performance increase is achieved, and computational complexity is added to the pipeline.

For the complex-valued moments, a feature vector construction method is proposed. This method showed an increase for all the complex moments and shows that adding phase information is beneficial for micro-Doppler classification problems.

Feature selection applied to the calculated orthogonal moments showed that the feature vector can be reduced without decreasing the classification abilities. This could potentially be used to speed up the classification process since complexity can be reduced.

7

CONCLUSION AND FUTURE WORK

7.1. SUMMARY AND CONCLUSION

This thesis focuses on the investigation of orthogonal moments for the classification of radar micro-Doppler signatures. It provides an analysis of the most frequently used orthogonal moments found in the literature, including the assessment of a proposed classification pipeline to identify the most critical aspects affecting the classification rates. Notably, to the best of the author's knowledge, the vast majority of the analysed moments has not been used for the classification of radar micro-Doppler data as discussed in this thesis. The main conclusions are summarized below as contributions to answer the formulated research questions, namely:

- **What moments show the best performances in the classification of micro-Doppler signatures?**
  The analysis of three different experimental datasets and order of moments has provided valuable information on the most effective moments for classification. The results demonstrate that CM, CHM-2nd, and ZM exhibit the strongest, consistent performance across all the evaluated datasets. Furthermore, RM and CHM-2nd demonstrate promising results in terms of overall performance with the same classifier across all datasets.

- **What is the influence of the dataset on the classification performance of orthogonal moments used as a feature?**
  The influence of the datasets in an orthogonal moment-based classification pipeline is shown to be significant. The change in the dataset causes a change in classification performance in terms of different optimal orthogonal moments, classifiers, and moment orders. This highlights the fact that relying solely on the evaluation of moment performance on a single dataset, as conducted in prior research, is a precarious approach, and on the contrary one should be mindful of adapting the moments-based classification approach to the specific radar data and classification problem at hand.

- **What is the influence of the classifier on the classification performance of orthogonal moments used as a feature?**
  The type of classifier employed in the pipeline for micro-Doppler signature classification using orthogonal moments has been demonstrated to influence the classification outcomes significantly. The performance of orthogonal moments varies depending on the specific classifier used in combination with them, a phenomenon observed also for other features proposed in the literature. Furthermore, the classification behavior of moments exhibits variations at different orders, implying that the optimal classifier at lower orders may differ from that at higher orders. These findings emphasize that the choice of classifier cannot be considered homogeneous within the classification pipeline of micro-Doppler signature classification using orthogonal moments. Essentially, it highlights the necessity of carefully considering the interaction between moments and classifiers to achieve the most effective and accurate classification results.

- **How robust are the orthogonal moments with respect to the introduction of noise?**
  The highest-performing moments have similar classification results as the state-of-the-art comparison method (Convolutional Neural Networks) in case noise is introduced in the test and train data. However, for the case where noise is only introduced in the test data, the moments outperformed the neural networks by a great margin. In this case, the best performing moments are of the family of the continuous moments defined on a disk, where the CHDM, PJFM, OFMM, PCETM, and RHFM showed the best results. These results favor the use of orthogonal moments in case the general noise distributions are not known when establishing the training set.

- **What is the influence of the variation in parameters for the moments with free parameters?**
  Moments with free parameters to be selected converge to the same classification performance at higher moment orders. This is explained as the moments capture the whole class-defying region of the data at high orders. At lower orders, the parameter selection will influence the classification performances depending on how much of the class-defying region is captured within their ROI.

- **What is the influence of using the CVD as a pre-processing step in the classification pipeline?**
  The use of the CVD does not appear to show a significant increase in classification performance with respect to using the spectrogram. For all three evaluated datasets, the majority of the orthogonal moments do not see an increase in F1 score. Therefore, the usage of CVD can be questioned for the studied classification problems, as it does not increase classification rates and yet introduces additional computational complexity.

- **How is the performance of orthogonal moments compared to other classification methods?**
  Orthogonal moments exhibit a modest but significant improvement in performance

compared to the convolutional neural networks investigated in this thesis. The classification results have shown that the best performing moments generally perform better in comparison to the evaluated neural networks. In addition, as mentioned above, the moments outperformed the evaluated neural networks in the cases where noise is introduced in the test data and the networks are trained on clean data.

In conclusion, the comprehensive evaluation of orthogonal moments for radar micro-Doppler classification demonstrates their efficacy across the essential aspects of the classification pipeline. This investigation highlights the significance of considering orthogonal moments as a valuable option in the pursuit of accurate and reliable micro-Doppler based classification.

7.2. FUTURE WORK

The presented evaluation of orthogonal moments indicated the potential to use orthogonal moments for radar classification. While several aspects have been evaluated, there are still some potential subjects that can be examined in future research.

- **Evaluation of realistic noise distributions**
  Measured radar data will inevitably be influenced by noise. This research assumed that the noise was white Gaussian and added it to the spectrogram image, which has no direct physical/electromagnetic meaning. Further research is needed to evaluate the performance of moments under more realistic noise conditions.

- **Evaluation of computational complexity**
  Orthogonal moments-based classification is assumed to be computationally simpler compared to neural networks. This property has not yet been evaluated in a rigorous manner, even because the complexity of either moments or networks can be highly variable depending on moment type or network architecture. Future research is needed to assess rigorously the computational benefits of using orthogonal moments.

- **Optimization of feature selection**
  The evaluated feature selection methods showed that the size of the feature vectors could be greatly reduced while still obtaining similar classification performances. The evaluation served as an indication of the influence of feature selection. In further research, the methods used for selection of features can be optimized. Furthermore, this process could potentially speed up classification. Training and classification will be less time-consuming when dealing with smaller feature vectors, and, in addition, fewer moments have to be calculated, decreasing the computational complexity of the feature vector construction.

- **Adaptation of moments in different scenarios**
  The optimal performing moments showed to differ from dataset to dataset. Further investigation is needed to evaluate moments on a more diverse selection of activities. This evaluation could indicate average best-performing moments in

real-life scenarios where many different activities are mixed in a continuous sequence.

BIBLIOGRAPHY

[1] V. Chen, "Analysis of radar micro-doppler with time-frequency transform," in *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing (Cat. No.00TH8496)*, Pocono Manor, PA, USA: IEEE, 2000, pp. 463–466, ISBN: 978-0-7803-5988-8. DOI: 10.1109/SSAP.2000.870167.

[2] D. F. Fioranelli, D. S. A. Shah, H. Li1, A. Shrestha, D. S. Yang, and D. J. L. Kernec, "Radar sensing for healthcare: Associate editor francesco fioranelli on the applications of radar in monitoring vital signs and recognising human activity patterns," *Electronics Letters*, vol. 55, no. 19, pp. 1022–1024, Sep. 2019, ISSN: 0013-5194, 1350-911X. DOI: 10.1049/el.2019.2378.

[3] Y. Li, Z. Peng, R. Pal, and C. Li, "Potential active shooter detection based on radar micro-doppler and range-doppler analysis using artificial neural network," *IEEE Sensors Journal*, vol. 19, no. 3, pp. 1052–1063, Feb. 1, 2019, ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2018.2879223.

[4] Y. Kim and B. Toomajian, "Hand gesture recognition using micro-doppler signatures with convolutional neural network," *IEEE Access*, vol. 4, pp. 7125–7130, 2016, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2016.2617282.

[5] D. Banerjee, S. Rani, A. M. George, *et al.*, "Application of spiking neural networks for action recognition from radar data," in *2020 International Joint Conference on Neural Networks (IJCNN)*, Glasgow, United Kingdom: IEEE, Jul. 2020, pp. 1–10, ISBN: 978-1-72816-926-2. DOI: 10.1109/IJCNN48605.2020.9206853.

[6] F. Fioranelli, M. Ritchie, and H. Griffiths, "Centroid features for classification of armed/unarmed multiple personnel using multistatic human micro-doppler," *IET Radar, Sonar & Navigation*, vol. 10, no. 9, pp. 1702–1710, Dec. 2016, ISSN: 1751-8792, 1751-8792. DOI: 10.1049/iet-rsn.2015.0493.

[7] S. Björklund, H. Petersson, and G. Hendeby, "Features for micro-doppler based activity classification," *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1181–1187, Dec. 2015, ISSN: 1751-8792, 1751-8792. DOI: 10.1049/iet-rsn.2015.0084.

[8] L. Pallotta, C. Clemente, A. De Maio, J. J. Soraghan, and A. Farina, "Pseudo-zernike moments based radar micro-doppler classification," in *2014 IEEE Radar Conference*, Cincinnati, OH, USA: IEEE, May 2014, pp. 0850–0854, ISBN: 978-1-4799-2035-8 978-1-4799-2034-1. DOI: 10.1109/RADAR.2014.6875709.

[9] Q. Li, Z. Wen, Z. Wu, *et al.*, "A survey on federated learning systems: Vision, hype and reality for data privacy and protection," *arXiv:1907.09693 [cs, stat]*, Jan. 13, 2021. arXiv: 1907.09693.

[10] B. van Nifterik, *Repository containing several implementations for the calculation of orthogonal moments*, in collab. with M. C. S. (TU Delft, suffix=Faculty Of Electrical Engineering, suffixi=F. O. E. E., version 1, Language: en Medium: *.m, Jun. 22, 2023. DOI: 10.4121/EFD059AA-F2C2-4C54-BE85-4E600876F444.V1.

[11] C. Clemente, L. Pallotta, A. De Maio, J. J. Soraghan, and A. Farina, "A novel algorithm for radar classification based on doppler characteristics exploiting orthogonal pseudo-zernike polynomials," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 417–430, Jan. 2015, ISSN: 0018-9251, 1557-9603, 2371-9877. DOI: 10.1109/TAES.2014.130762.

[12] R. Boulic, N. M. Thalmann, and D. Thalmann, "A global human walking model with real-time kinematic personification," *The Visual Computer*, vol. 6, no. 6, pp. 344–358, Nov. 1990, ISSN: 0178-2789, 1432-8726. DOI: 10.1007/BF01901021.

[13] F. Fioranelli, M. Ritchie, and H. Griffiths, "Multistatic human micro-doppler classification of armed/unarmed personnel," *IET Radar, Sonar & Navigation*, vol. 9, no. 7, pp. 857–865, Aug. 2015, ISSN: 1751-8792, 1751-8792. DOI: 10.1049/iet-rsn.2014.0360.

[14] D. Wang, J. Park, H.-J. Kim, K. Lee, and S. H. Cho, "Noncontact extraction of biomechanical parameters in gait analysis using a multi-input and multi-output radar sensor," *IEEE Access*, vol. 9, pp. 138 496–138 508, 2021, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2021.3117985.

[15] L. Pallotta, M. Cauli, C. Clemente, F. Fioranelli, G. Giunta, and A. Farina, "Classification of micro-doppler radar hand-gesture signatures by means of chebyshev moments," in *2021 IEEE 8th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, Naples, Italy: IEEE, Jun. 23, 2021, pp. 182–187, ISBN: 978-1-72817-556-0. DOI: 10.1109/MetroAeroSpace51421.2021.9511751.

[16] S. A. Shah and F. Fioranelli, "Human activity recognition : Preliminary results for dataset portability using FMCW radar," in *2019 International Radar Conference (RADAR)*, TOULON, France: IEEE, Sep. 2019, pp. 1–4, ISBN: 978-1-72812-660-9. DOI: 10.1109/RADAR41533.2019.171307.

[17] J. Martin and B. Mulgrew, "Analysis of the theoretical radar return signal form aircraft propeller blades," in *IEEE International Conference on Radar*, Arlington, VA, USA: IEEE, 1990, pp. 569–572. DOI: 10.1109/RADAR.1990.201091.

[18] S. Yang, S. Yeh, S. Bor, S. Huang, and C. Hwang, "Electromagnetic backscattering from aircraft propeller blades," *IEEE Transactions on Magnetics*, vol. 33, no. 2, pp. 1432–1435, Mar. 1997, ISSN: 00189464. DOI: 10.1109/20.582526.

[19] W. Zhang and G. Li, "Detection of multiple micro-drones via cadence velocity diagram analysis," *Electronics Letters*, vol. 54, no. 7, pp. 441–443, Apr. 2018, ISSN: 0013-5194, 1350-911X. DOI: 10.1049/el.2017.4317.

[20] D. Tahmoush, "Review of micro-doppler signatures," *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1140–1146, Dec. 2015, ISSN: 1751-8792, 1751-8792. DOI: 10.1049/iet-rsn.2015.0118.

[21] L. Liu, D. McLernon, M. Ghogho, W. Hu, and J. Huang, "Ballistic missile detection via micro-doppler frequency estimation from radar return," *Digital Signal Processing*, vol. 22, no. 1, pp. 87–95, Jan. 2012, ISSN: 10512004. DOI: 10.1016/j.dsp.2011.10.009.

[22] M. Ritchie, R. Capraru, and F. Fioranelli, "Dop-NET: A micro-doppler radar data challenge," *Electronics Letters*, vol. 56, no. 11, pp. 568–570, May 2020, ISSN: 0013-5194, 1350-911X. DOI: 10.1049/el.2019.4153.

[23] E. Kurtoglu, A. C. Gurbuz, E. A. Malaia, D. Griffin, C. Crawford, and S. Z. Gurbuz, *ASL trigger recognition in mixed activity/signing sequences for RF sensor-based user interfaces*, Nov. 18, 2021. arXiv: 2111.05480[eess].

[24] M. J. Bocus, W. Li, S. Vishwakarma, *et al.*, "OPERAnet, a multimodal activity recognition dataset acquired from radio frequency and vision-based sensors," *Scientific Data*, vol. 9, no. 1, p. 474, Dec. 2022, ISSN: 2052-4463. DOI: 10.1038/s41597-022-01573-2.

[25] S. Z. Gurbuz, M. M. Rahman, E. Kurtoglu, T. Macks, and F. Fioranelli, "Cross-frequency training with adversarial learning for radar micro-doppler signature classification (rising researcher)," in *Radar Sensor Technology XXIV*, A. M. Raynal and K. I. Ranney, Eds., Online Only, United States: SPIE, May 11, 2020, p. 16, ISBN: 978-1-5106-3593-7 978-1-5106-3594-4. DOI: 10.1117/12.2559155.

[26] R. I. A. Harmanny, J. J. M. de Wit, and G. P. Cabic, "Radar micro-doppler feature extraction using the spectrogram and the cepstrogram," in *2014 11th European Radar Conference*, Italy: IEEE, Oct. 2014, pp. 165–168, ISBN: 978-2-87487-037-8. DOI: 10.1109/EuRAD.2014.6991233.

[27] X. Qiao, T. Shan, R. Tao, X. Bai, and J. Zhao, "Separation of human micro-doppler signals based on short-time fractional fourier transform," *IEEE Sensors Journal*, vol. 19, no. 24, pp. 12 205–12 216, Dec. 15, 2019, ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2019.2937989.

[28] Z. Qun, G. Hua, G. Ying, and B. Youqing, "Separation of micro-doppler signal using an extended hough transform," in *2006 International Conference on Communications, Circuits and Systems*, Guilin, Guangzi, China: IEEE, Jun. 2006, pp. 361–365, ISBN: 978-0-7803-9584-8 978-0-7803-9585-5. DOI: 10.1109/ICCCAS.2006.284654.

[29] X. Qiao, M. G. Amin, T. Shan, Z. Zeng, and R. Tao, "Human activity classification based on micro-doppler signatures separation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022, ISSN: 0196-2892, 1558-0644. DOI: 10.1109/TGRS.2021.3105124.

[30] S. Rani, A. Chowdhury, T. Chakravarty, and A. Pal, "Exploiting unique state transitions to capture micro-doppler signatures of human actions using CW radar," *IEEE Sensors Journal*, vol. 21, no. 24, pp. 27 878–27 886, Dec. 15, 2021, ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2021.3126436.

[31] S. Bjorklund, T. Johansson, and H. Petersson, "Evaluation of a micro-doppler classification method on mm-wave data," in *2012 IEEE Radar Conference*, Atlanta, GA: IEEE, May 2012, pp. 0934–0939, ISBN: 978-1-4673-0658-4 978-1-4673-0656-0 978-1-4673-0657-7. DOI: 10.1109/RADAR.2012.6212271.

[32] A. Aßmann, A. Izzo, and C. Clemente, "Efficient micro-doppler based pedestrian activity classification for ADAS systems using krawtchouk moments," p. 6, 2016.

[33] C. Pang, Y. Han, H. Hou, S. Liu, and N. Zhang, "Micro-doppler signal time-frequency algorithm based on STFRFT," *Sensors*, vol. 16, no. 10, p. 1559, Sep. 24, 2016, ISSN: 1424-8220. DOI: 10.3390/s16101559.

[34] D. Donnelly, "The fast fourier and hilbert-huang transforms: A comparison," in *The Proceedings of the Multiconference on "Computational Engineering in Systems Applications"*, Beijing, China: IEEE, Oct. 2006, pp. 84–88, ISBN: 978-7-302-13922-5 978-7-900718-14-3. DOI: 10.1109/CESA.2006.4281628.

[35] P. Sathe, D. A, K. P. Ray, S. D, and V. A, "Helicopter main and tail rotor blade parameter extraction using micro-doppler," in *2018 19th International Radar Symposium (IRS)*, Bonn, Germany: IEEE, Jun. 2018, pp. 1–10, ISBN: 978-3-7369-9545-1. DOI: 10.23919/IRS.2018.8448242.

[36] L. Stankovic, M. Dakovic, T. Thayaparan, and V. Popovic-Bugarin, "Inverse radon transform–based micro-doppler analysis from a reduced set of observations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 2, pp. 1155–1169, Apr. 2015, ISSN: 0018-9251. DOI: 10.1109/TAES.2014.140098.

[37] Ming-Kuei Hu, "Visual pattern recognition by moment invariants," *IEEE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, Feb. 1962, ISSN: 0018-9448. DOI: 10.1109/TIT.1962.1057692.

[38] S. Qi, Y. Zhang, C. Wang, J. Zhou, and X. Cao, "A survey of orthogonal moments for image representation: Theory, implementation, and evaluation," 2021, Publisher: arXiv Version Number: 3. DOI: 10.48550/ARXIV.2103.14799.

[39] P. Kaur, H. S. Pannu, and A. K. Malhi, "Comprehensive study of continuous orthogonal moments—a systematic review," *ACM Computing Surveys*, vol. 52, no. 4, pp. 1–30, Jul. 31, 2020, ISSN: 0360-0300, 1557-7341. DOI: 10.1145/3331167.

[40] C.-H. Teh and R. Chin, "On image analysis by the methods of moments," in *Proceedings CVPR '88: The Computer Society Conference on Computer Vision and Pattern Recognition*, Ann Arbor, MI, USA: IEEE Comput. Soc. Press, 1988, pp. 556–561, ISBN: 978-0-8186-0862-9. DOI: 10.1109/CVPR.1988.196290.

[41] L. Wang, Y. Wu, and M. Dai, "Some aspects of gaussian-hermite moments in image analysis," in *Third International Conference on Natural Computation (ICNC 2007)*, Haikou, China: IEEE, 2007, pp. 450–454, ISBN: 978-0-7695-2875-5. DOI: 10.1109/ICNC.2007.676.

[42] R. Mukundan, S. Ong, and P. Lee, "Image analysis by tchebichef moments," *IEEE Transactions on Image Processing*, vol. 10, no. 9, pp. 1357–1364, Sep. 2001, ISSN: 10577149. DOI: 10.1109/83.941859.

[43] A. Chiang, S. Liao, Q. Lu, and M. Pawlak, "Gegenbauer moment-based applications for chinese character recognition," in *IEEE CCECE2002. Canadian Conference on Electrical and Computer Engineering. Conference Proceedings (Cat. No.02CH37373)*, Winnipeg, Man., Canada: IEEE, 2002, pp. 908–911, ISBN: 978-0-7803-7514-7. DOI: 10.1109/CCECE.2002.1013064.

[44] A. Khotanzad and Y. Hong, "Invariant image recognition by zernike moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 489–497, May 1990, ISSN: 01628828. DOI: 10.1109/34.55109.

[45] Z. Ping, R. Wu, and Y. Sheng, "Image description with chebyshev–fourier moments," *Journal of the Optical Society of America A*, vol. 19, no. 9, p. 1748, Sep. 1, 2002, ISSN: 1084-7529, 1520-8532. DOI: 10.1364/JOSAA.19.001748.

[46] G. Amu, S. Hasi, X. Yang, and Z. Ping, "Image analysis by pseudo-jacobi (p = 4, q = 3)–fourier moments," *Applied Optics*, vol. 43, no. 10, p. 2093, Apr. 1, 2004, ISSN: 0003-6935, 1539-4522. DOI: 10.1364/AO.43.002093.

[47] Z. Ping, H. Ren, J. Zou, Y. Sheng, and W. Bo, "Generic orthogonal moments: Jacobi–fourier moments for invariant image description," *Pattern Recognition*, vol. 40, no. 4, pp. 1245–1254, Apr. 2007, ISSN: 00313203. DOI: 10.1016/j.patcog.2006.07.016.

[48] H.-t. Hu, Y.-d. Zhang, C. Shao, and Q. Ju, "Orthogonal moments based on exponent functions: Exponent-fourier moments," *Pattern Recognition*, vol. 47, no. 8, pp. 2596–2606, Aug. 2014, ISSN: 00313203. DOI: 10.1016/j.patcog.2014.02.014.

[49] B. Xiao, J.-F. Ma, and X. Wang, "Image analysis by bessel–fourier moments," *Pattern Recognition*, vol. 43, no. 8, pp. 2620–2629, Aug. 2010, ISSN: 00313203. DOI: 10.1016/j.patcog.2010.03.013.

[50] H. Ren, Z. Ping, W. Bo, W. Wu, and Y. Sheng, "Multidistortion-invariant image recognition with radial harmonic fourier moments," *Journal of the Optical Society of America A*, vol. 20, no. 4, p. 631, Apr. 1, 2003, ISSN: 1084-7529, 1520-8532. DOI: 10.1364/JOSAA.20.000631.

[51] Qian Liu, Hongqing Zhu, and Qian Li, "Object recognition by combined invariants of orthogonal fourier-mellin moments," in *2011 8th International Conference on Information, Communications & Signal Processing*, Singapore: IEEE, Dec. 2011, pp. 1–5, ISBN: 978-1-4577-0031-6 978-1-4577-0029-3 978-1-4577-0030-9. DOI: 10.1109/ICICS.2011.6174265.

[52] Pew-Thian Yap, R. Paramesran, and Seng-Huat Ong, "Image analysis by krawtchouk moments," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1367–1377, Nov. 2003, ISSN: 1057-7149. DOI: 10.1109/TIP.2003.818019.

[53] P.-T. Yap, R. Paramesran, and S.-H. Ong, "Image analysis using hahn moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 2057–2062, Nov. 2007, ISSN: 0162-8828. DOI: 10.1109/TPAMI.2007.70709.

[54] H. Zhu, H. Shu, J. Liang, L. Luo, and J.-L. Coatrieux, "Image analysis by discrete orthogonal racah moments," *Signal Processing*, vol. 87, no. 4, pp. 687–708, Apr. 2007, ISSN: 01651684. DOI: 10.1016/j.sigpro.2006.07.007.

[55] H. Zhu, M. Liu, H. Shu, H. Zhang, and L. Luo, "General form for obtaining discrete orthogonal moments," *IET Image Processing*, vol. 4, no. 5, p. 335, 2010, ISSN: 17519659. DOI: 10.1049/iet-ipr.2009.0195.

[56] Pew-Thian Yap, Xudong Jiang, and A. C. Kot, "Two-dimensional polar harmonic transforms for invariant image representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 7, pp. 1259–1270, Jul. 2010, ISSN: 0162-8828. DOI: 10.1109/TPAMI.2009.119.

[57] N. M. Temme and J. L. Lopez, "The askey scheme for hypergeometric orthogonal polynomials viewed from asymptotic analysis," *Journal of Computational and Applied Mathematics*, vol. 133, no. 1, pp. 623–633, Aug. 2001, ISSN: 03770427. DOI: 10.1016/S0377-0427(00)00683-X. arXiv: math/0109185.

[58] R. Benouini, I. Batioua, K. Zenkouar, A. Zahi, H. E. Fadili, and H. Qjidaa, "Fast and accurate computation of racah moment invariants for image classification," *Pattern Recognition*, vol. 91, pp. 100–110, Jul. 2019, ISSN: 00313203. DOI: 10.1016/j.patcog.2019.02.014.

[59] B. Kaur, S. Singh, and J. Kumar, "Robust iris recognition using moment invariants," *Wireless Personal Communications*, vol. 99, no. 2, pp. 799–828, Mar. 2018, ISSN: 0929-6212, 1572-834X. DOI: 10.1007/s11277-017-5153-8.

[60] G. A. Papakostas, B. G. Mertzios, and D. A. Karras, "Performance of the orthogonal moments in reconstructing biomedical images," in *2009 16th International Conference on Systems, Signals and Image Processing*, Chalkida: IEEE, Jun. 2009, pp. 1–4. DOI: 10.1109/IWSSIP.2009.5367686.

[61] X. Wang, Y. Yang, and K. Huang, "Combining discrete orthogonal moments and DHMMs for off-line handwritten chinese character recognition," in *2006 5th IEEE International Conference on Cognitive Informatics*, Beijing, China: IEEE, Jul. 2006, pp. 788–793, ISBN: 978-1-4244-0475-9. DOI: 10.1109/COGINF.2006.365590.

[62] F. C. Soon, H. Y. Khaw, and J. H. Chuah, "Pattern recognition of vehicle logo using tchebichef and legendre moment," in *2015 IEEE Student Conference on Research and Development (SCOReD)*, Kuala Lumpur, Malaysia: IEEE, Dec. 2015, pp. 82–86, ISBN: 978-1-4673-9572-4. DOI: 10.1109/SCORED.2015.7449438.

[63] S. Machhour, E. Grivel, P. Legrand, V. Corretja, and C. Magnant, "A comparative study of orthogonal moments for micro-doppler classification," in *2018 26th European Signal Processing Conference (EUSIPCO)*, Rome: IEEE, Sep. 2018, pp. 366–370, ISBN: 978-90-827970-1-5. DOI: 10.23919/EUSIPCO.2018.8553069.

[64] C. Clemente, L. Pallotta, C. Ilioudis, F. Fioranelli, G. Giunta, and A. Farina, "Chebychev moments based drone classification, recognition and fingerprinting," in *2021 21st International Radar Symposium (IRS)*, Berlin, Germany: IEEE, Jun. 21, 2021, pp. 1–6, ISBN: 978-3-944976-31-0. DOI: 10.23919/IRS51887.2021.9466211.

[65] T. Jahid, H. Karmouni, A. Hmimid, M. Sayyouri, and H. Qjidaa, "Fast computation of charlier moments and its inverses using clenshaw's recurrence formula for image analysis," *Multimedia Tools and Applications*, vol. 78, no. 9, pp. 12 183–12 201, May 2019, ISSN: 1380-7501, 1573-7721. DOI: 10.1007/s11042-018-6757-z.

[66] M. Yamni, A. Daoui, O. El Ogri, H. Karmouni, M. Sayyouri, and H. Qjidaa, "Accurate 2d and 3d images classification using translation and scale invariants of meixner moments," *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 26 683–26 712, Jul. 2021, ISSN: 1380-7501, 1573-7721. DOI: 10.1007/s11042-020-10311-y.

[67] S. Elshoura and D. Megherbi, "Analysis of noise sensitivity of tchebichef and zernike moments with application to image watermarking," *Journal of Visual Communication and Image Representation*, vol. 24, no. 5, pp. 567–578, Jul. 2013, ISSN: 10473203. DOI: 10.1016/j.jvcir.2013.03.021.

[68] Zi-ping Ma, Bao-sheng Kang, and B. Xiao, "A study of bessel fourier moments invariant for image retrieval and classification," in *2011 International Conference on Image Analysis and Signal Processing*, Wuhan, Hubei, China: IEEE, Oct. 2011, pp. 316–320, ISBN: 978-1-61284-881-5 978-1-61284-879-2. DOI: 10.1109/IASP.2011.6109054.

[69] W. Wang, Y. Li, and S. Liu, "A polar complex exponential transform-based zero-watermarking for multiple medical images with high discrimination," *Security and Communication Networks*, vol. 2021, J. Wang, Ed., pp. 1–13, Mar. 17, 2021, ISSN: 1939-0122, 1939-0114. DOI: 10.1155/2021/6615678.

[70] X. Ma, R. Pan, and L. Wang, "License plate character recognition based on gaussian-hermite moments," in *2010 Second International Workshop on Education Technology and Computer Science*, Wuhan, China: IEEE, Mar. 2010, pp. 11–14, ISBN: 978-1-4244-6388-6 978-1-4244-6389-3. DOI: 10.1109/ETCS.2010.591.

[71] H. T. Le, S. L. Phung, A. Bouzerdoum, and F. H. C. Tivive, "Human motion classification with micro-doppler radar and bayesian-optimized convolutional neural networks," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, AB: IEEE, Apr. 2018, pp. 2961–2965, ISBN: 978-1-5386-4658-8. DOI: 10.1109/ICASSP.2018.8461847.

[72] A. R. Persico, C. Clemente, L. Pallotta, A. De Maio, and J. Soraghan, "Micro-doppler classification of ballistic threats using krawtchouk moments," in *2016 IEEE Radar Conference (RadarConf)*, Philadelphia, PA: IEEE, May 2016, pp. 1–6, ISBN: 978-1-5090-0863-6. DOI: 10.1109/RADAR.2016.7485086.

[73] J. S. Rani and D. Devaraj, "Face recognition using krawtchouk moment," *Sadhana*, vol. 37, no. 4, pp. 441–460, Aug. 2012, ISSN: 0256-2499, 0973-7677. DOI: 10.1007/s12046-012-0090-4.

[74] W. Zhu, Y. Zhu, X. Li, and D. Yuan, "The posture recognition of pigs based on zernike moments and support vector machines," in *2015 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, Taipei, Taiwan: IEEE, Nov. 2015, pp. 480–484, ISBN: 978-1-4673-9323-2. DOI: 10.1109/ISKE.2015.51.

[75] K. Wang, H. Zhang, Z. Ping, and Haiying, "Chinese chess character recognition with radial harmonic fourier moments," in *2011 International Conference on Document Analysis and Recognition*, Beijing: IEEE, Sep. 2011, pp. 1369–1373, ISBN: 978-1-4577-1350-7 978-0-7695-4520-2. DOI: 10.1109/ICDAR.2011.275.

[76] S. P. Priyal and P. K. Bora, "A study on static hand gesture recognition using moments," in *2010 International Conference on Signal Processing and Communications (SPCOM)*, Bangalore, India: IEEE, Jul. 2010, pp. 1–5, ISBN: 978-1-4244-7137-9. DOI: 10.1109/SPCOM.2010.5560535.

[77] B. P. Sari and Y. Jusman, "Classification system for cervical cell images based on hu moment invariants methods and support vector machine," in *2021 International Conference on Intelligent Technologies (CONIT)*, Hubli, India: IEEE, Jun. 25, 2021, pp. 1–5, ISBN: 978-1-72818-583-5. DOI: 10.1109/CONIT51480.2021.9498353.

[78] P. Bolourchi, H. Demirel, and S. Uysal, "Continuous moment-based features for classification of ground vehicle SAR images," in *2016 European Modelling Symposium (EMS)*, Pisa, Italy: IEEE, Nov. 2016, pp. 53–57, ISBN: 978-1-5090-4971-4. DOI: 10.1109/EMS.2016.019.

[79] R. Upneja and C. Singh, "Fast computation of jacobi-fourier moments for invariant image recognition," *Pattern Recognition*, vol. 48, no. 5, pp. 1836–1843, May 2015, ISSN: 00313203. DOI: 10.1016/j.patcog.2014.11.012.

[80] R. Mukundan and K. Ramakrishnan, "Fast computation of legendre and zernike moments," *Pattern Recognition*, vol. 28, no. 9, pp. 1433–1442, Sep. 1995, ISSN: 00313203. DOI: 10.1016/0031-3203(95)00011-N.

[81] H.-y. Yang, S.-r. Qi, P.-p. Niu, and X.-y. Wang, "Color image zero-watermarking based on fast quaternion generic polar complex exponential transform," *Signal Processing: Image Communication*, vol. 82, p. 115 747, Mar. 2020, ISSN: 09235965. DOI: 10.1016/j.image.2019.115747.

[82] H. Cohl and C. MacKenzie, *Generalizations and specializations of generating functions for jacobi, gegenbauer, chebyshev and legendre polynomials with definite integrals*, Jun. 26, 2013. arXiv: 1210.0039[math].

[83] K. M. Hosny, "Image representation using accurate orthogonal gegenbauer moments," *Pattern Recognition Letters*, vol. 32, no. 6, pp. 795–804, Apr. 2011, ISSN: 01678655. DOI: 10.1016/j.patrec.2011.01.006.

[84] C.-W. Chong, P. Raveendran, and R. Mukundan, "Translation and scale invariants of legendre moments," *Pattern Recognition*, vol. 37, no. 1, pp. 119–129, Jan. 2004, ISSN: 00313203. DOI: 10.1016/j.patcog.2003.06.003.

[85] B. Yang and M. Dai, "Image analysis by gaussian–hermite moments," *Signal Processing*, vol. 91, no. 10, pp. 2290–2303, Oct. 2011, ISSN: 01651684. DOI: 10.1016/j.sigpro.2011.04.012.

[86] A. Daoui, H. Karmouni, M. Yamni, M. Sayyouri, and H. Qjidaa, "On computational aspects of high-order dual hahn moments," *Pattern Recognition*, vol. 127, p. 108 596, Jul. 2022, ISSN: 00313203. DOI: 10.1016/j.patcog.2022.108596.

[87] A. Daoui, H. Karmouni, M. Sayyouri, and H. Qjidaa, "Fast and stable computation of higher-order hahn polynomials and hahn moment invariants for signal and image analysis," *Multimedia Tools and Applications*, vol. 80, no. 21, pp. 32 947–32 973, Sep. 2021, ISSN: 1380-7501, 1573-7721. DOI: 10.1007/s11042-021-11206-2.

[88] S. H. Abdulhussain, A. R. Ramli, S. A. R. Al-Haddad, B. M. Mahmmod, and W. A. Jassim, "Fast recursive computation of krawtchouk polynomials," *Journal of Mathematical Imaging and Vision*, vol. 60, no. 3, pp. 285–303, Mar. 2018, ISSN: 0924-9907, 1573-7683. DOI: 10.1007/s10851-017-0758-9.

[89] S. Li, M.-C. Lee, and C.-M. Pun, "Complex zernike moments features for shape-based image retrieval," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 39, no. 1, pp. 227–237, Jan. 2009, ISSN: 1083-4427, 1558-2426. DOI: 10.1109/TSMCA.2008.2007988.

[90] C. Clemente, L. Pallotta, D. Gaglione, A. De Maio, and J. J. Soraghan, "Automatic target recognition of military vehicles with krawtchouk moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 1, pp. 493–500, Feb. 2017, ISSN: 0018-9251, 1557-9603, 2371-9877. DOI: 10.1109/TAES.2017.2649160.

[91] Z. R. Saputra Elsi, D. Stiawan, A. F. Oklilas, *et al.*, "Feature selection using chi square to improve attack detection classification in IoT network: Work in progress," in *2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, Jakarta, Indonesia: IEEE, Oct. 6, 2022, pp. 226–232, ISBN: 9786239213565. DOI: 10.23919/EECSI56542.2022.9946621.

A

RECURSIVE GEGENBAUER MOMENT CALCULATION

A

Algorithm 1 Recusive calcualtion gegenbauer moment method

$n_{max} \leftarrow$ Maximum order of moments
$F(N,N) \leftarrow$ Image

/* Compute Gegebauer polynomial */
for $i = 1\ to\ N$ **do**
    $x \leftarrow (i * 2/N) - 1$
    $P(0,i) \leftarrow 1$
    $P(1,i) \leftarrow 2 * a * x$
    for $i = 2, n_{max}$ **do**
        $P(n,i) \leftarrow (2*(n+a-1)*x*P(n-1,i) - (n+2*a-2)*P(n-2,i))/n$
    end for
end for
/* Compute Gegebauer Moment */
for $n \leftarrow 0\ to\ n_{max}$ **do**
    for $m \leftarrow 0\ to\ n_{max}$ **do**
        $sum \leftarrow 0$
        for $i \leftarrow 1 to N$ **do**
            for $j \leftarrow 1 to N$ **do**
                $sum \leftarrow sum + P(n,i) * P(m,j) * F(i,j) * ((1 - x(j)^2)^{a-0.5}) * ((1 - x(i)^2)^{a-0.5})$
            end for
        end for
        $moments(n,m) \leftarrow sum * \frac{4}{(N-1)^2} * ((n! * (n+a) * (\Gamma(a))^2)/(pi * \Gamma(n+2*a) * (2^{1-2*a}))) * (m! * (m+a) * (\Gamma(a))^2)/(pi * \Gamma(m+2*a) * (2^{1-2*a}))$
    end for
end for

B

GENERAL RESULTS

B.1. INFLUENCE OF THE MOMENT ORDERS

Moment	Max F1 score	Best Classifier	F1 score							
			n=5	n=10	n=15	n=20	n=25	n=30	n=40	n=50
PJFM	0.900	'RBF svm'	0.835	0.900	0.866	0.775	0.575	0.338	0.120	0.0554
ZM	0.921	'RBF svm'	0.756	0.850	0.900	0.921	0.880	0.646	0.327	0.0606
PZM	0.914	'RBF svm'	0.844	0.905	0.914	0.698	0.531	0.438	0.336	0.3293
CHFM	0.883	'RBF svm'	0.850	0.883	0.854	0.757	0.563	0.309	0.112	0.0554
OFMM	0.889	'RBF svm'	0.833	0.889	0.866	0.775	0.575	0.341	0.120	0.0554
EFM	0.901	'RBF svm'	0.801	0.891	0.901	0.893	0.691	0.504	0.176	0.0554
PCEM	0.817	'RBF svm'	0.759	0.817	0.775	0.691	0.571	0.484	0.403	0.3038
BFM	0.846	'RBF svm'	0.835	0.846	0.702	0.535	0.219	0.159	0.084	0.0554
RHFM	0.894	'RBF svm'	0.842	0.894	0.883	0.794	0.578	0.324	0.066	0.0501

Table B.1: Classification results for the continuous moments defined on a square grid using the Glasgow dataset

Moment	Max F1 score	Best Classifier	F1 score							
			n=5	n=10	n=15	n=20	n=25	n=30	n=40	n=50
PJFM	0.775	'RBF svm'	0.529	0.775	0.740	0.420	0.135	0.046	0.019	0.019
ZM	0.805	'Knn, k =1'	0.595	0.629	0.649	0.752	0.805	0.757	0.722	0.619
PZM	0.772	'Knn, k =1'	0.656	0.719	0.725	0.772	0.707	0.272	0.255	0.173
CHFM	0.766	'Knn, k =4'	0.651	0.766	0.609	0.547	0.534	0.487	0.397	0.375
OFMM	0.777	'Knn, k =4'	0.678	0.777	0.635	0.560	0.534	0.473	0.385	0.382
EFM	0.748	'RBF svm'	0.619	0.748	0.690	0.468	0.240	0.102	0.019	0.019
PCEM	0.775	'Knn, k =5'	0.775	0.655	0.562	0.490	0.480	0.440	0.319	0.307
BFM	0.760	'RBF svm'	0.564	0.760	0.249	0.019	0.019	0.019	0.019	0.019
RHFM	0.783	'RBF svm'	0.528	0.783	0.740	0.417	0.124	0.046	0.019	0.019

Table B.2: Classification results for the continuous moments defined on a square grid using the Xetrhu dataset

Moment	Max F1 score	Best Classifier	F1 score									
			n=5	n=10	n=15	n=20	n=25	n=30	n=35	n=40	n=45	n=50
GM	0.895	'RBF svm'	0.526	0.854	0.895	0.884	0.818	0.605	0.531	0.347	0.093	0.050
LM	0.889	'RBF svm'	0.597	0.889	0.871	0.805	0.575	0.362	0.109	0.055	0.055	0.050
CHM-1st	0.891	'Knn, k =5'	0.791	0.891	0.845	0.840	0.807	0.803	0.732	0.674	0.602	0.621
CHM-2nd	0.878	'RBF svm'	0.487	0.697	0.854	0.878	0.866	0.860	0.862	0.842	0.835	0.731
GHM	0.871	'RBF svm'	0.592	0.751	0.818	0.839	0.867	0.871	0.857	0.846	0.755	0.657

Table B.3: Classification results for the continuous moments defined on a square grid using the Glasgow dataset

Moment	Max F1 score	Best Classifier	F1 score									
			n=5	n=10	n=15	n=20	n=25	n=30	n=35	n=40	n=45	n=50
GM	0.794	'RBF svm'	0.460	0.674	0.761	0.794	0.373	0.053	0.019	0.019	0.019	0.019
LM	0.756	'RBF svm'	0.454	0.659	0.756	0.534	0.101	0.019	0.019	0.019	0.019	0.019
CHM-1st	0.772	'RBF svm'	0.482	0.623	0.708	0.764	0.772	0.745	0.577	0.275	0.132	0.054
CHM-2nd	0.808	'RBF svm'	0.511	0.628	0.720	0.786	0.808	0.798	0.696	0.314	0.132	0.083
GHM	0.787	'RBF svm'	0.452	0.681	0.767	0.787	0.787	0.618	0.256	0.110	0.019	0.019

Table B.4: Classification results for the continuous moments defined on a square grid using the Xtehru dataset

Moment	Max F1 score	Best Classifier	F1 score									
			n=5	n=10	n=15	n=20	n=25	n=30	n=35	n=40	n=45	n=50
RM max	0.880	'RBF svm'	0.564	0.821	0.880	0.870	0.857	0.860	0.862	0.855	0.806	0.705
DHM max	0.873	'Knn, k =10'	0.627	0.814	0.853	0.850	0.856	0.873	0.867	0.843	0.824	0.829
HM max	0.876	'RBF svm'	0.524	0.807	0.870	0.876	0.860	0.857	0.856	0.849	0.818	0.699
CHDM	0.883	'RBF svm'	0.602	0.883	0.867	0.767	0.487	0.112	0.050	0.050	0.050	0.050
KM max	0.870	'RBF svm'	0.398	0.700	0.834	0.831	0.870	0.859	0.857	0.837	0.765	0.664
MM max	0.873	'Knn, k =8'	0.340	0.539	0.789	0.838	0.838	0.873	0.847	0.824	0.758	0.758
CM max	0.873	'RBF svm'	0.605	0.773	0.816	0.873	0.867	0.866	0.852	0.858	0.790	0.687

Table B.5: Classification results for the discrete moments defined on a square grid using the Glasgow dataset

Moment	Max F1 score	Best Classifier	F1 score									
			n=5	n=10	n=15	n=20	n=25	n=30	n=35	n=40	n=45	n=50
RM max	0.810	'RBF svm'	0.231	0.460	0.703	0.754	0.774	0.810	0.635	0.296	0.145	0.019
DHM max	0.802	'RBF svm'	0.181	0.681	0.793	0.802	0.748	0.360	0.132	0.072	0.019	0.019
HM max	0.787	'RBF svm'	0.494	0.664	0.781	0.778	0.775	0.787	0.528	0.207	0.121	0.019
CHDM	0.760	'RBF svm'	0.412	0.671	0.760	0.473	0.090	0.019	0.019	0.019	0.019	0.019
KM max	0.794	'RBF svm'	0.460	0.674	0.761	0.794	0.373	0.053	0.019	0.019	0.019	0.019
MM max	0.834	'Knn, k =5'	0.279	0.471	0.642	0.650	0.746	0.817	0.825	0.834	0.735	0.766
CM max	0.811	'Knn, k =2'	0.271	0.456	0.660	0.786	0.811	0.790	0.762	0.730	0.719	0.672

Table B.6: Classification results for the discrete moments defined on a square grid using the Xehtru dataset

B.2. INFLUENCE OF THE CVD

Moments	Spectrogram		CVD	
	Max F1	Best Classifier	Max F1	Best Classifier
PJFM	0.900	'RBF svm'	0.873	'linear svm'
ZM	0.921	'RBF svm'	0.718	'linear svm'
PZM	0.914	'RBF svm'	0.780	'linear svm'
CHFM	0.883	'RBF svm'	0.879	'linear svm'
OFMM	0.889	'RBF svm'	0.874	'linear svm'
EFM	0.901	'RBF svm'	0.888	'RBF svm'
PCEM	0.817	'RBF svm'	0.842	'linear svm'
BFM	0.846	'RBF svm'	0.866	'linear svm'
RHFM	0.894	'RBF svm'	0.863	'linear svm'
GM	0.895	'RBF svm'	0.797	'RBF svm'
LM	0.889	'RBF svm'	0.790	'linear svm'
CHM-1st	0.891	'Knn, k =5'	0.774	'linear svm'
CHM-2nd	0.878	'RBF svm'	0.797	'Knn, k =10'
GHM	0.871	'RBF svm'	0.813	'linear svm'
RM max	0.880	'RBF svm'	0.843	'linear svm'
DHM max	0.873	'Knn, k =10'	0.880	'linear svm'
HM max	0.876	'RBF svm'	0.837	'linear svm'
CHDM	0.883	'RBF svm'	0.798	'RBF svm'
KM max	0.870	'RBF svm'	0.863	'linear svm'
MM max	0.873	'Knn, k =8'	0.788	'Knn, k =5'
CM max	0.873	'RBF svm'	0.815	'linear svm'

Table B.7: Classification results comparison for the spectrogram vs CVD used as pre-processing technique, using the Glasgow dataset

Moments	Spectrogram		CVD	
	Max F1	Best Classifier	Max F1	Best Classifier
PJFM	0.775	'RBF svm'	0.519	'Quadratic svm'
ZM	0.805	'Knn, k =1'	0.493	'Knn, k =1'
PZM	0.772	'Knn, k =1'	0.507	'Knn, k =1'
CHFM	0.766	'Knn, k =4'	0.509	'Knn, k =4'
OFMM	0.777	'Knn, k =4'	0.518	'Quadratic svm'
EFM	0.748	'RBF svm'	0.607	'Knn, k =8'
PCEM	0.775	'Knn, k =5'	0.619	'Knn, k =4'
BFM	0.760	'RBF svm'	0.515	'Quadratic svm'
RHFM	0.783	'RBF svm'	0.537	'Knn, k =4'
GM	0.794	'RBF svm'	0.640	'Quadratic svm'
LM	0.756	'RBF svm'	0.645	'Knn, k =4'
CHM-1st	0.772	'RBF svm'	0.621	'Knn, k =4'
CHM-2nd	0.808	'RBF svm'	0.626	'Knn, k =2'
GHM	0.787	'RBF svm'	0.627	'Quadratic svm'
RM max	0.810	'RBF svm'	0.647	'Quadratic svm'
DHM max	0.802	'RBF svm'	0.641	'Knn, k =1'
HM max	0.787	'RBF svm'	0.634	'Quadratic svm'
CHDM	0.760	'RBF svm'	0.663	'linear svm'
KM max	0.794	'RBF svm'	0.625	'Knn, k =1'
MM max	0.834	'Knn, k =5'	0.667	'Quadratic svm'
CM max	0.811	'Knn, k =2'	0.660	'Quadratic svm'

Table B.8: Classification results comparison for the spectrogram vs CVD used as pre-processing technique, using the Xterhu dataset

B.3. INFLUENCE OF THE CLASSIFIER

Moments	RBF SVM	Quadratic SVM	Linear SVM	KNN-1	KNN-5	KNN-10	Difference
PJFM	0.775	0.626	0.696	0.754	0.770	0.665	0.150
ZM	0.761	0.664	0.705	0.805	0.718	0.695	0.140
PZM	0.761	0.683	0.697	0.772	0.753	0.645	0.127
CHFM	0.759	0.643	0.657	0.758	0.743	0.681	0.116
OFMM	0.752	0.694	0.706	0.773	0.741	0.686	0.087
EFM	0.748	0.653	0.642	0.713	0.705	0.651	0.106
PCEM	0.765	0.743	0.673	0.775	0.775	0.689	0.102
BFM	0.760	0.723	0.652	0.704	0.710	0.611	0.149
RHFM	0.783	0.661	0.710	0.710	0.742	0.615	0.168
GM	0.794	0.696	0.690	0.732	0.739	0.714	0.104
LM	0.756	0.696	0.687	0.753	0.713	0.702	0.069
CHM-1st	0.772	0.699	0.656	0.719	0.727	0.673	0.116
CHM-2nd	0.808	0.652	0.666	0.759	0.738	0.701	0.156
GHM	0.787	0.650	0.693	0.740	0.730	0.694	0.137
RM max	0.810	0.662	0.684	0.790	0.782	0.782	0.148
DHM max	0.802	0.640	0.668	0.751	0.716	0.645	0.162
HM max	0.786	0.708	0.681	0.757	0.721	0.683	0.105
CHDM	0.760	0.705	0.679	0.754	0.723	0.687	0.081
KM max	0.794	0.623	0.705	0.743	0.739	0.715	0.171
MM max	0.780	0.712	0.684	0.789	0.834	0.778	0.150
CM max	0.801	0.703	0.692	0.754	0.786	0.739	0.109

Table B.9: F1-scores for different moments when different classifiers are used to classify the Xterhu dataset. The difference in F1-score between the best-performing and worst-performing classifier for each moment is included.

Moments	RBF SVM	Quadratic SVM	Linear SVM	KNN-1	KNN-5	KNN-10	Difference
PJFM	0.925	0.960	0.852	0.913	0.942	0.913	0.108
ZM	0.931	0.946	0.844	0.906	0.902	0.876	0.102
PZM	0.932	0.954	0.847	0.950	0.931	0.912	0.106
CHFM	0.924	0.959	0.856	0.912	0.932	0.902	0.103
OFMM	0.922	0.950	0.866	0.918	0.941	0.922	0.084
EFM	0.903	0.917	0.805	0.887	0.836	0.769	0.148
PCEM	0.941	0.926	0.823	0.896	0.891	0.884	0.118
BFM	0.891	0.948	0.834	0.931	0.941	0.903	0.114
RHFM	0.920	0.956	0.834	0.934	0.961	0.897	0.127
GM	0.936	0.888	0.821	0.935	0.940	0.905	0.119
LM	0.915	0.891	0.839	0.939	0.939	0.898	0.100
CHM-1st	0.934	0.751	0.831	0.917	0.903	0.903	0.182
CHM-2nd	0.968	0.883	0.837	0.932	0.932	0.916	0.131
GHM	0.971	0.920	0.920	0.909	0.897	0.857	0.114
RM max	0.967	0.905	0.880	0.895	0.871	0.781	0.186
DHM max	0.970	0.919	0.899	0.949	0.938	0.857	0.113
HM max	0.967	0.909	0.836	0.923	0.935	0.887	0.131
CHDM	0.902	0.888	0.829	0.932	0.939	0.898	0.110
KM max	0.971	0.925	0.899	0.937	0.918	0.894	0.077
MM max	0.966	0.883	0.842	0.912	0.903	0.857	0.124
CM max	0.968	0.935	0.880	0.887	0.887	0.811	0.157

Table B.10: F1-scores for different moments when different classifiers are used to classify the Dopnet dataset. The difference in F1-score between the best-performing and worst-performing classifier for each moment is included.