REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

GESTURE RECOGNITION FOR HUMAN-MACHINE INTERACTION IN AVIATION

Tuğba ZEYBEK

MASTER OF SCIENCE THESIS
Department of Avionics Engineering
Avionics Engineering

Supervisor
Assoc. Prof. Ufuk SAKARYA

February, 2022

REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

GESTURE RECOGNITION FOR HUMAN-MACHINE INTERACTION IN AVIATION

A thesis submitted by Tuğba ZEYBEK in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE** is approved by the committee on 07.02.2021 in Department of Avionics Engineering, Avionics Engineering.

Assoc. Prof. Ufuk SAKARYA
Yildiz Technical University
Supervisor

Approved By the Examining Committee

Assoc. Prof. Ufuk SAKARYA, Supervisor
Yildiz Technical University _____

Prof. Şeref Naci ENGİN, Member
Yildiz Technical University _____

Asst. Prof. Ramazan YENİÇERİ, Member
İstanbul Technical University _____

I hereby declare that I have obtained the required legal permissions during data collection and exploitation procedures, that I have made the in-text citations and cited the references properly, that I haven't falsified and/or fabricated research data and results of the study and that I have abided by the principles of the scientific research and ethics during my Thesis Study under the title of Gesture Recognition for Human-Machine Interaction in Aviation supervised by my supervisor, Assoc. Prof. Ufuk SAKARYA. In the case of a discovery of false statement, I am to acknowledge any legal consequence.

Tuğba ZEYBEK

Signature

Dedicated to my family
and my friends

ACKNOWLEDGEMENTS

I would like to thank and express my thankfulness to my valuable advisor Assoc. Prof. Dr Ufuk SAKARYA, who helped me in the realization of this thesis.

I would like to express my gratitude to all my other university professors, especially my esteemed advisor who helped and guided me in the selection of the my thesis subject, for contributing to me during my master education and equipping me with knowledge.

I would also like to appreciation my friends who helped me with the problems I got stuck and motivated me during my thesis. I would like to thank my dear colleagues and esteemed managers for their understanding and support during this process.

And lastly, I would like to thank my family, who have always supported me throughout my thesis and in all areas of my life, and who have been my greatest luck in this life.

Tuğba ZEYBEK

TABLE OF CONTENTS

LIST OF SYMBOLS viii

LIST OF ABBREVIATIONS ix

LIST OF FIGURES x

LIST OF TABLES xi

ABSTRACT xii

ÖZET xiv

1 INTRODUCTION 1
- 1.1 Literature Review . . . 1
- 1.2 Problem Definition . . . 3
- 1.3 Motivation . . . 3

2 Wavelet based Spatial temporal Feature Extraction for Gesture Recognition 5
- 2.1 Wavelet Transform . . . 5
- 2.2 Proposed Method . . . 6
- 2.3 Experimental Results . . . 8
  - 2.3.1 6D Motion Gesture Database . . . 9
  - 2.3.2 Experiments and Determining the Parameter Space . . . 9
  - 2.3.3 Comparative Experiments . . . 11

3 Dimension Reduction and SVM based Gesture Recognition Architecture 15
- 3.1 Dimension Reduction . . . 15
  - 3.1.1 Principal Component Analysis . . . 16
  - 3.1.2 Linear Discriminant Analysis . . . 17
- 3.2 Support Vector Machine . . . 18
- 3.3 FlightGear Flight Simulator . . . 18
- 3.4 Proposed Method . . . 19
- 3.5 Test Setup and Simulation Results . . . 20

4 Dominant Set Clustering in Gesture Recognition 31

4.1 Dominant Set Clustering 31
4.2 Proposed Method 33
4.3 Experiments and Determining the Parameter Space 34

5 CONCLUSION **38**

REFERENCES **40**

A GESTURES CODES **44**

PUBLICATIONS FROM THE THESIS **45**

LIST OF SYMBOLS

cA	Approximation Coefficients
C_i	Decision of each sensor obtained from the classifier
cD	Detail Coefficients
W_i	Discrete wavelet transform's approximation and detail coefficients
S_i	Each sensor
μ	Mean
Z_i	Mean and variance of W_i
ZN_i	Normalization of Z_i
S_B	The inter-class scatter matrix
S_W	The intra-class scatter matrix
t	Time
W	Transformation matrix
σ	Variance

LIST OF ABBREVIATIONS

6DMG	6D Motion Gesture
CNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform
DOF	Degree of Freedom
DTW	Dynamic Time Warping
DWT	Discrete Wavelet Transform
EMG	Electromyography
FG	Flight-Gear
GS	Gesture Set
HMI	Human-Machine Interaction
HMM	Hidden Markov Model
HP	High Pass
kNN	k-Nearest Neighbors
LDA	Linear Discriminant Analysis
LP	Low Pass
MSFS	Microsoft Flight Simulator
NNs	Neural Networks
PCA	Principal Component Analysis
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
UAV	Unmanned Aerial Vehicle

LIST OF FIGURES

Figure 1.1 The architecture of the thesis 4
Figure 2.1 The architecture of the proposed method 7
Figure 2.2 The decomposition of sensor signals into sub-bands by the discrete wavelet transform 8
Figure 3.1 Proposed Method 21
Figure 3.2 Connection of the Tools for the Experiment Setup 26
Figure 3.3 Initial state of the scenario 27
Figure 3.4 Simulation status as a result of the combination of CIR VER CLK and POKE RIGHT commands 27
Figure 3.5 FG state with SWIPE UP command 28
Figure 3.6 FG state with SWIPE DOWN command 28
Figure 3.7 Aircraft behaviour under SWIPE RIGHT command 29
Figure 3.8 Aircraft behaviour under SWIPE LEFT command 29
Figure 3.9 Data from the aircraft throughout the simulation 29
Figure 3.10 The image of the aircraft on the map during the simulation 30
Figure 4.1 Proposed Method 33
Figure 4.2 Expected and dominant-set results 35

LIST OF TABLES

Table 2.1 Sub-band codes according to the experiments 10
Table 2.2 Parameter space research experiments 10
Table 2.3 Step distance and angles for XY and XZ planes 11
Table 2.4 Comparative Results 13
Table 3.1 Differences between PCA and LDA 16
Table 3.2 Flight Simulator Properties 19
Table 3.3 Classifier Selection Experiments 22
Table 3.4 Majority Voting Results 22
Table 3.5 Success Comparison of Gestures 23
Table 3.6 Success for New Gesture Set 23
Table 3.7 Gestures Selected for the Scenario 24
Table 3.8 Success of the Gestures Selected for the Scenario 25
Table 3.9 Success of the Scenario 25
Table 4.1 α parameter selection experiment 35
Table 4.2 The effect of the total number of users on the dominant set clustering 36
Table 4.3 The effect of the total number of users making same gesture on the dominant set clustering 36
Table A.1 Gesture codes for experiments 44

ABSTRACT

Gesture Recognition for Human-Machine Interaction in Aviation

Tuğba ZEYBEK

Department of Avionics Engineering
Master of Science Thesis

Supervisor: Assoc. Prof. Ufuk SAKARYA

With the developing technology, human-machine interaction (HMI) is very significant in aviation, as in every area. Hand gestures are accepted in HMI applications. Because the human hand is quite adept at controlling a device. Gesture recognition problems in HMI also have some problems due to their nature. Not every human performs the same movement in space in the same way. Even the same person cannot perform the same movement in the exact same manner. This complicates the recognition problem.

Due to large in-class differences between users, the efficiency of features derived from different sensors is compared in user-independent situations. The spatio-temporal nature of gesture signals is considered for recognition and a simple statistical feature-based approach is presented that works on transformed signals. With this approach, it is time-independent and allows the transition to the energy domain.

With the wavelet-based statistical feature method, the size of the new data increases considerably. It becomes difficult to analyze high-dimensional data and to find meaningful data in the data set. Linear dimension reduction techniques are applied to solve the new problem. The scenario created in the motion classes is classified after the combination of dimension reduction and classification techniques. These results are linked to the flight simulator tool.

In this thesis study, in cases where the pilots are insufficient for the increasing number of UAVs, the dominant set clustering method is recommended in order to provide UAV control with less trained pilots. This method is flexible as it can be used in both civil

and military fields. In critical situations, more than one intermediately trained pilot may act like a well-trained pilot. In other cases, already moderately trained pilots will suffice.

Keywords: Human-Machine Interaction, gesture recognition, spatio-temporal, Support Vector Machine, dominant set clustering

ÖZET

Havacılıkta İnsan-Makine Etkileşimi için El Hareketi Sezme

Tuğba ZEYBEK

Aviyonik Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Danışman: Assoc. Prof. Ufuk SAKARYA

Gelişen teknoloji ile birlikte insan-makine etkileşimi (HMI) her alanda olduğu gibi havacılıkta da oldukça önemlidir. HMI uygulamalarında el hareketleri kabul edilmektedir. Çünkü insan eli bir cihazı kontrol etmekte oldukça ustadır. HMI'daki jest tanıma sorunları da doğası gereği bazı problemlere sahiptir. Her insan uzayda aynı hareketi aynı şekilde yapmaz. Aynı kişi bile aynı hareketi aynı şekilde yapamaz. Bu, tanıma sorununu karmaşıklaştırır.

Kullanıcılar arasındaki sınıf içi büyük farklılıklar nedeniyle, farklı sensörlerden türetilen özelliklerin verimliliği, kullanıcıdan bağımsız durumlarda karşılaştırılır. Tanıma için jest sinyallerinin uzamsal-zamansal doğası göz önünde bulundurulur ve dönüştürülmüş sinyaller üzerinde çalışan basit bir istatistiksel özellik tabanlı yaklaşım sunulur. Bu yaklaşımla zamandan bağımsızdır ve enerji alanına geçişi sağlar.

Dalgacık tabanlı istatistiksel öznitelik yöntemi ile yeni verinin boyutu önemli ölçüde artar. Yüksek boyutlu verileri analiz etmek ve veri setinde anlamlı veriler bulmak zorlaşmaktadır. Yeni problemi çözmek için doğrusal boyut küçültme teknikleri uygulanır. Hareket sınıflarında oluşturulan senaryo, boyut küçültme ve sınıflandırma tekniklerinin kombinasyonundan sonra sınıflandırılır. Bu sonuçlar, uçuş simülatörü aracıyla bağlantılıdır.

Bu tez çalışmasında, artan İHA sayısı için pilotların yetersiz kaldığı durumlarda, İHA kontrolünün daha az eğitimli pilotlarla sağlanabilmesi için baskın küme kümeleme yöntemi önerilmektedir. Bu yöntem hem sivil hem de askeri alanda kullanılabildiği

için esnektir. Kritik durumlarda, birden fazla orta düzeyde eğitimli pilot, iyi eğitimli bir pilot gibi davranabilir. Diğer durumlarda, zaten orta derecede eğitimli pilotlar yeterli olacaktır.

Anahtar Kelimeler: İnsan-Makine Etkileşimi, jest tanıma, uzamsal-zamansal, vektör destekli makine, baskın küme kümeleme

1
INTRODUCTION

The importance of the gesture recognition problem for human-machine interaction (HMI) is emerging with the development of technology. The gesture recognition problem for HMI contains some inherent difficulties. Different people can perform the same movement with a certain level of difference. Even the movements of the same person at different times can contain differences. This fact means that each movement is unique in the space-time plane. For this reason, the recognition of gestures independently of the person emerges as a striking problem. Mainly, while performing a movement, parameters such as speed and movement magnitude can be different from person to person. This can cause challenges in solving the problem. The problem of gesture recognition for HMI can be considered as a pattern recognition problem from the resulting vectors that data is collected with different types of sensors in the spatial-temporal plane.

1.1 Literature Review

There are research topics related to HMI in many fields, such as the automotive industry and biomedical science [1]. The study reported algorithms for hand gesture recognition in recent years. Moreover, the applied algorithms were compared and these algorithms were evaluated by experimental methods. Hand gestures are accepted for HMI in the application areas. Ten gesture class classification studies were carried out with the sensors placed inside the vehicle. [2] describes a real-time gesture classification method for automobiles. In [3], hand gestures were also preferred in HMI to increase driver efficiency. CNN was used for hand gesture recognition and the success was reported as 87.5% and 97% under two different test conditions.

In [4], linear discriminant analysis (LDA) was preferred in order to reduce the information in the surface electromyogram signals and improve the identification efficiency. Another study on the classification of hand gestures based on EMG can be found in [5]. The remarkable thing about this study is that it proposes the use

of time-frequency representations (for example, Wavelet and Fourier transforms) as a preprocessing step to exploit temporal information in the data as a future plan.

The gesture recognition problem in aviation is much more critical than in the biomedical and automotive sectors. These areas, however, cannot be kept completely separate from one another.With the increase in research in these areas, the use of hand gestures for HMI in aviation has been considered as a research topic.

[6] High-tech systems, such as automatic collision avoidance maneuvers, have been introduced on modern aviation flight decks to support pilots in helping them fly the aircraft. The article sheds light on the current use and future use of HMI on flight decks. In this article, publications that can be a solution to this problem have been examined and have provided guidance for effective HMI methods [7].

Experiments have been conducted showing that automation and auditory offloading can be successful by easing the overall workload. In an experiment, pilots were given some tasks to complete while flying a single UAV. In another experiment, they were asked to fly two UAVs under the same conditions.

With the development of technology, it is predicted that the number of pilots may not be sufficient in the future. With the proposed method, it makes it possible to get a result by looking at the majority of the decisions made by intermediate pilots in critical situations. With this system, the UAV can be used in many civil and military applications. There may be many critical situations that the UAV will face in military situations, but in civil situations such as drug transfer, the UAV's behavior will mostly be the same.

The use of drones in commercial transportation has started in America and Europe. It was also used in some operations in other countries for COVID-19 disease. In [8], a method is presented for the use of drones in some situations where there is no land transportation, for example, in natural disasters such as earthquakes.

Hand movements are quite widespread in UAV control. For example, the hand gesture is received by a controller and transmitted to the ground station. The ground station transforms the movement into the appropriate action for the drone [9]. In another study, gesture recognition offers a different use. With the motion received from the user by a UAV camera, the system sends a warning message. This application is intended to be used in human rescue [10]. This study prioritised the problem of gesture recognition in UAV control.

There are many difficulties in the gesture recognition problem. Here, as the difference

in subjects is a problem, even the same person does not perform exactly the same when he repeats the movement a second time. Many problems arise, such as the differences between right-handers and left-handers. There are approaches in the literature to reach a solution without being affected by different parameters such as speed differences and gesture magnitudes. Dynamic time warping (DTW) [11], neural networks (NNs) [12], hidden Markov model (HMM) [13], [14], [15] and feature-based statistical approaches [16], [17] are the main approaches for analyzing spatial-temporal patterns.

In this study, considering the studies in the literature, a wavelet-based approach has been proposed, and after this approach, a different feature has been proposed by taking statistical information. The proposed method is developed on the basis of statistical approaches.

1.2 Problem Definition

The proliferation of unmanned aerial vehicles (UAVs) and the ability to use gestures in UAV control is a research topic. The main motivation for the proposed study is to develop a novel vector-based feature extraction method for gesture recognition in UAV control. The vector-based approach is preferred because it has some advantages for practical usage. One of them is that the decision part may not be handled outside of UAV. Therefore, the features may be transferred to the UAV and then the decision may be made by itself using these vector-based features and some additional information according to the situation. In other words, not only vector-based features can have low memory sizes, but also low-complexity solutions for vector-based decision processes. Since complexity and energy are critical for UAVs, On the other hand, the proposed method can be applicable in different application areas for gesture recognition too. Therefore, the proposed method is a feature extraction method for gesture recognition without loss of generality in the wide application areas.

1.3 Motivation

Humans have the ability to control many devices around them with their gestures. With the existence of studies where this ability is also used in drone control and with the developing technology, it will be difficult for a single person to follow these controls. It's pretty clear that the number of drones is going to start to outnumber pilots. This study aims to obtain an output from the inputs of more than one experienced intermediate pilot and enable the air system to make the correct decision. Furthermore, because of human nature, controlling with hand movements will reduce

reaction time when compared to other devices.

The main architecture of the thesis is given in the Figure 1.1. The operations up to the normalization process are the identical in all chapters. After the normalization process, the thesis presents 3 different methods in separate chapters.

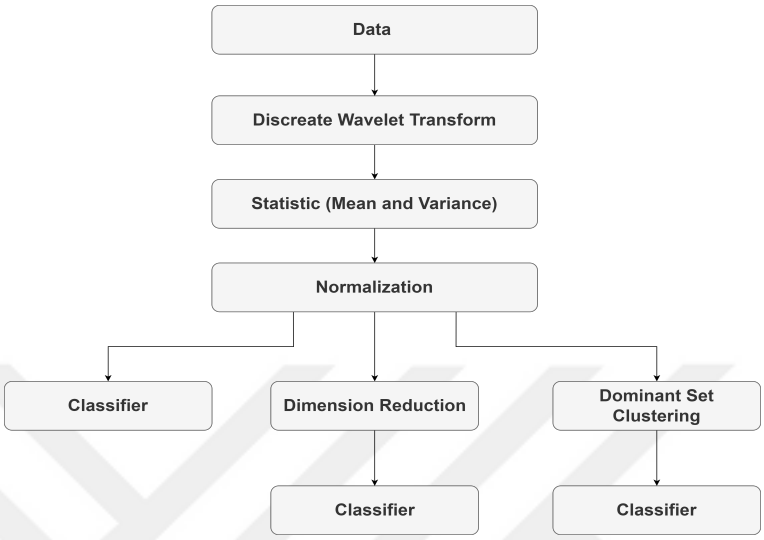


Figure 1.1 The architecture of the thesis

In this study, wavelet-based feature extraction is discussed in Chapter 2. This section describes the success of the proposed method and the corresponding tests. The optimal classification process is not in focus. In Chapter 3, training and test data are separated as in real problems. Afterwards, two different classifiers are discussed. The problem is also discussed before and after the dimension reduction. Also, in this chapter, a scenario was created from successful movements, and this scenario was run in the simulator. Chapter 4 contains the dominant set and its results. It is aimed at controlling multi-pilot with a dominant set method. The conclusion part of the study is given in Chapter 5.

2
Wavelet based Spatial temporal Feature Extraction for Gesture Recognition

In this section, the wavelet transform and its applications in the literature are discussed. Afterwards, the proposed method is given. The proposed method section includes the gesture data-set used and the determination of parameters and experiments. At the end of the chapter, there is a discussion and conclusion part.

2.1 Wavelet Transform

[18] The wavelet transform is basically of two types: continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The continuous wavelet transform is an alternative to the short-time Fourier transform (STFT) and was developed to overcome the resolution problem. The analysis is quite similar to the STFT analysis in that the signal is multiplied by a function. The discrete wavelet transform is the sampled version of the CWT. DWT is sufficient for analyzing the signal. Additionally, it is computationally more convenient than CWT.

With the discrete wavelet transform, filters are used to analyze a signal in different scales and frequency domains [18]. There are various investigations into this topic. The discrete wavelet transform was used as a feature extraction in electromyography (EMG) signals and then classified using a Support Vector Machine (SVM) [19]. In another similar study, a 2D wavelet transform was used as a feature extraction after image processing from hand movements. Using the 7th level approximation vector in the wavelet transform, it is classified using a multi-class SVM classifier [20]. In addition, the wavelet transform technique has been used to recognize short duration articulated object motions [21]. Multi-scale stationary wavelet analysis has been used to generate image-based multi-scale spatial temporal representations of human behavior in [22]. The representation proposed in this paper exploits the ability of the 3D stationary wavelet transform to combine highlighted spatio-temporal information at different scales and orientations.

In this study, the use of the discrete wavelet transform for the problem of recognizing person-independent hand movements using multi-sensor data is proposed. In addition, the proposed method is based on statistical feature extraction from this transformed signal. Rubine [16] presented the extraction of statistically based features in the spatial-temporal space. The proposed method is different from Rubine's method. Instead of working in the spatial-temporal space, sensor signals are transferred to the scale-frequency space using the discrete wavelet transform. It tries to extract a robust feature against the movement differences of different people or the same people at different times.

In this new data space, 1st and 2nd order statistical features (mean and variance) have been applied. The proposed method is also different from the previously applied discrete wavelet transform approaches in the gesture recognition problem for HMI. In the proposed study, the discrete wavelet transform is applied to multi-sensor data in spatial and temporal space. The problem of multi-sensor data coming from each user in different dimensions is overcome by extracting statistical features from signals that have undergone the discrete wavelet transform. Gestures are recognized by a classifier with the last obtained features. This paper does not focus on finding the optimum classifier for the proposed feature. It is another research topic. However, in this study, it is concentrated on the demonstration of the success of the proposed feature according to comparative analysis. In order to do this fairly, a widely used non-parametric approach was preferred. The nearest neighbor classifier (NN) has been used [23]. Finding the most suitable classifier for the proposed feature space is the next chapter research topic. A non-parametric approach is selected because when the comparative experiments are realized, selecting suitable parameters for the classifier is another problem for achieving fair comparison.

2.2 Proposed Method

The proposed wavelet-based spatial-temporal feature extraction method for gesture recognition is presented. The proposed method uses multi-sensor data. The discrete wavelet transform is used on multi-sensor data collected in both the temporal plane and the three-dimensional spatial plane.

The system architecture is presented in Figure 2.1. S_i refers to sensors (i = 1 to N). Each S_i sensor consists of a B_i size vector that it receives at each sampling time. It takes a total of M_j samples within the total time T_j where each j indexed motion starts and ends. Thus, S_i sensor data for j indexed gesture movement consists of a matrix of $B_i x M_j$ size.

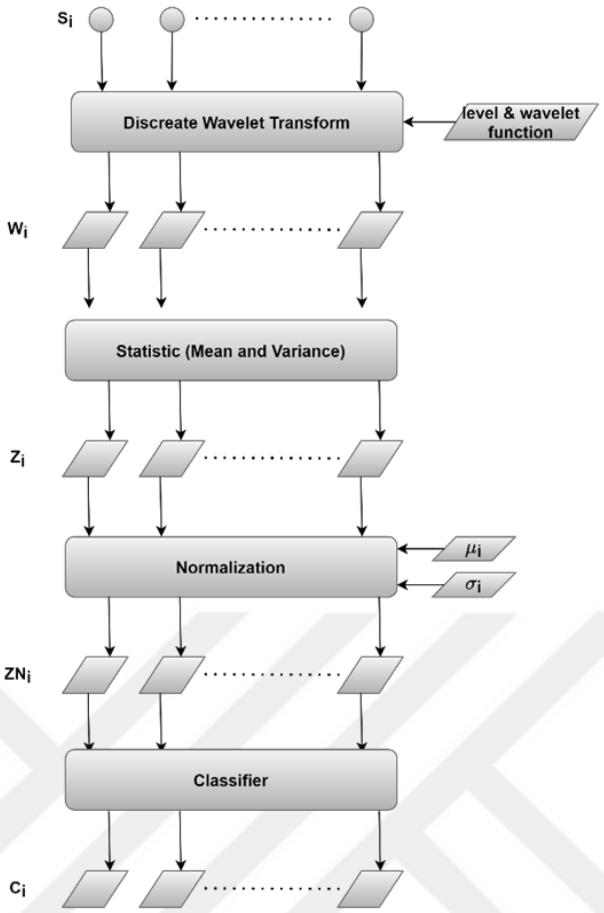


Figure 2.1 The architecture of the proposed method

Total T_j time of each j-indexed movement may differ from each other. If even the same person performs the same movement at different speeds, different T_j times can be obtained. For this reason, although the spatial vector dimension B_j of each sensor is the same, the temporal vector dimension M_j may not be the same. Thus, a problem arises that it is difficult to compare in the vector space. In addition, the spatial or temporal differences between different or the same people make the gesture recognition problem difficult while performing the same gesture. As mentioned in the previous section, statistical approaches have been used to overcome this problem, especially in the temporal domain. However, these approaches are statistical approaches applied directly to sensor data in the temporal domain. In the proposed method, using the discrete wavelet transform before applying statistical approaches, it is aimed to make the sensor temporal data independent from time and less affected by the individual differences that make the movement.

In the proposed method, the time-dependent signals received from each S_i sensor are transferred to the scale-frequency domain by the discrete wavelet transform. The reason why it is preferred to use the wavelet transform instead of the Fourier transform is that while the Fourier transform tells which frequencies are in the signal as a whole,

the wavelet transform tells which frequencies are on what scale and in which region. The region is important to represent for temporal behavior because different gestures can have different frequency characteristics in temporal slices according to time scales. As a result, W_i refers to the discrete wavelet transform's approximation and detail coefficients.

The signals (s[n]) from each sensor are passed through low-pass (LP) and high-pass (HP) filters. Signals passed through LP express approximation coefficients (cA), signals passed through HP express detail coefficients (cD). For example, cD1: Level 1 detail coefficient, cD2: Level 2 detail coefficient. Fig. 2.2 shows the wavelet transformation tree up to Level 3. The red areas on the left, shown in the , indicate that the cD1 signal is also decomposed up to Level 2. Here, cA21 refers to the Level 2 approximation coefficient of the cD1 signal.

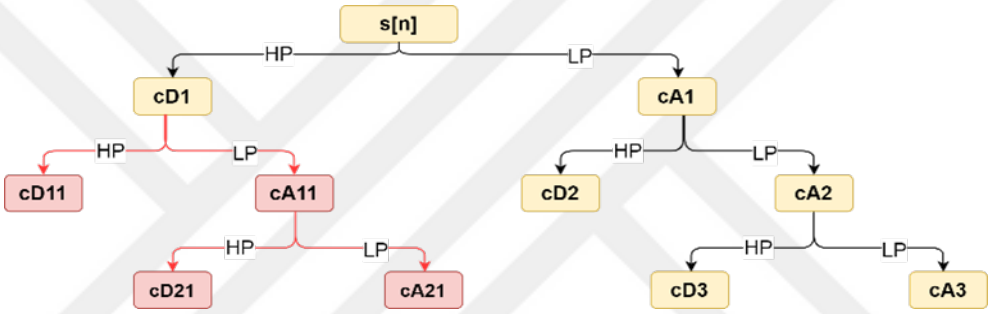


Figure 2.2 The decomposition of sensor signals into sub-bands by the discrete wavelet transform

Since the characteristics of each individual sensor may be different, it may be necessary to determine the appropriate discrete wavelet transform parameters for each sensor. There are two main parameters in the proposed method. The first one is the level of the discrete wavelet transform. The second parameter is the importance of high-frequency or low-frequency sub-band components. Since these two parameter spaces may change according to the data, the approach presented on this subject will be explained in detail in the experimental studies section.

In the next step, the first and second order statistics Z_i of these coefficients are taken. After the normalization process of the obtained data, the normalized data is given to the classifier. The decision C_i of each sensor obtained from the classifier is mostly decided on the final class LC_j using the majority voting algorithm.

2.3 Experimental Results

It is important to understand the data set before performing gesture recognition. For this reason, in this section, the data set and the implementation details are given. All

experiments were realized in the MATLAB2019b [24] tool.

Generally, the reported recognition results are above 90% in the literature. These results have been taken in different data sets and in various experimental settings, and it is not meaningful to directly compare the performance achievements of these techniques [15]. For this reason, in this study, it is tried to show the success of the proposed method through a comparative testing use of the same data set with a feature-based reference method under the same conditions.

Sensors in our experiment are position, orientation, acceleration, angular speed, and velocity in the form of i = 1,2,3,4,5, respectively. The experiments have been conducted in the experimental results in order to see in which frequency region and at which level the signals in the data set are more effective.

2.3.1 6D Motion Gesture Database

6D Motion Gesture (6DMG) data-set [15], [25] includes the time information, position, acceleration, orientation, and angular speed of common gesture movements received from different users. Speed information in the data set is created by deriving from time and position. The more precisely the motion can be tracked, the more successful the pattern recognition will be. While 2D tracking was sufficient for graphical user interfaces in the past, now the 6DOF (Degree of Freedom, DOF) tracking feature comes to the fore. This data-set is preferred because the hand is a resourceful object with more than 20DOF. Due to dependencies between joints and fingers, DOF may be reduced; however, studies have shown that at least six dimensions are required [25].

For registration, 21 right-handed and 7 left-handed (22 of them are male and 6 of them are female and range from ages 15 to 33) were taken from the participants. Each participant repeated each movement 10 times, and these repetitions were recorded. The 6DMG consists of 20 gestures and 5600 motion patterns captured by 28 participants [15], [25].

2.3.2 Experiments and Determining the Parameter Space

Using the 6DMG data set, new signals are obtained by performing the discrete wavelet transform as described in the proposed method. Nine experiments have been conducted in order to see which wavelet scale and frequency region is suitable for the data set. The level of wavelet was chosen to be 2 in experiments 1, 2 and 3, as well as 3 in experiments 4, 5 and 6, and as 4 in experiments 7, 8 and 9. The representation of the

importance of high frequencies or low frequencies sub-band components is realized by 3 experiments at each level. The wavelet sub-band codes are given according to experiments in Table 2.1. In each experiment here, first order statistics of the signals obtained from the wavelet transform are used.

Table 2.1 Sub-band codes according to the experiments

Experiment Numbers	Sub-band codes
1	cD1, cD2, cA2
2	cD1, cD2, cA2, cA1
3	cD1, cD2, cA2, cD11
4	cD1, cD2, cD3, cA3
5	cD1, cD2, cD3, cA3, cA1, cA2
6	cD1, cD2, cD3, cA3, cD11, cD21
7	cD1, cD2, cD3, cD4, cA4
8	cD1, cD2, cD3, cD4, cA4, cA1, cA2, cA3
9	cD1, cD2, cD3, cD4, cA4, cD11, cD21, cD31

In the experiments, 50 random samples from each of the 20 classes were taken and tested. The samples taken were given to the NN classifier, and this process was repeated 10 times. The values in Table 2.2 are the average of 10 repetitions performed.

As seen in Table 2.2, the performance in experiment 8 is the most successful result in the experiments. After selecting this parameter set, in addition to the first statistic, the second statistic features are also extracted. After repeating the same processes for this parameter set, the results are given in Table 2.4 as the proposed method. When the variance is added to this data set, performance differences of between 7-11% are observed in each sensor.

Table 2.2 Parameter space research experiments

Experiments	Position	Orientation	Acceleration	Angular Angular	Velocity
1	79.42	81.73	55.37	58.39	75.43
2	77.19	80.73	56.74	61.67	76.81
3	73.94	77.75	50.74	55.64	74.31
4	80.08	82.30	57.04	55.93	74.41
5	77.43	81.14	58.87	62.15	77.09
6	74.09	76.43	50.55	52.90	72.33
7	79.75	81.47	58.01	56.23	73.29
8	78.66	81.76	61.02	63.96	78.38
9	74.36	76.21	51.42	52.67	70.24

2.3.3 Comparative Experiments

Statistical feature-based methods are advantageous due to their fast and easy implementation. In this paper, the Rubine feature set [16] is chosen to compare the 6DMG data set as the reference model from the literature. The Rubine feature set is proposed for a 2D trajectory. Hoffman adapted the Rubine feature set to the 3D orbit [16], [17]. In order to make the comparative experiment fair, the proposed method is under examination with the same random data set and the same parameters. In addition, the Rubine feature set, originally for a 2D trajectory and only for position sensors, is adapted to 3D and 5 sensors using the following formulations: Before defining the features, step distance and angles for the XY and XZ planes were first defined in Table 2.3:

Table 2.3 Step distance and angles for XY and XZ planes

$\Delta x_k = x_k - x_{k-1}$	$\Delta y_k = y_k - y_{k-1}$	$\Delta z_k = z_k - z_{k-1}$
$\Theta_k = \arctan\left(\frac{\Delta x_k \Delta y_{k+1} - \Delta x_{k+1} \Delta y_k}{\Delta x_k \Delta x_{k+1} - \Delta y_k \Delta y_{k+1}}\right)$		
$\gamma_k = \arctan\left(\frac{\Delta x_k \Delta z_{k+1} - \Delta x_{k+1} \Delta z_k}{\Delta x_k \Delta x_{k+1} - \Delta z_k \Delta z_{k+1}}\right)$		

The f_{1-3} features are the sine and cosine of the motion in the XY and XZ planes of the initial angles. The f_{4-6} features are sine and cosine in the XY and XZ plane for the first and last points of the motion. The f_7 is the distance between the maximum and minimum points. The f_{8-9} are the diagonal angle of the bounded region. The f_{10} gives the distance between the first and last points of the motion, and the f_{11} gives the total length of the motion. The f_{12-13} refer to the total angle traveled in the XY and XZ planes, respectively. The f_{14-15} are the absolute value of the angle at each point in motion. The f_{16-17} are the sum of the squares of the angles. The f_{18} is the maximum speed of the motion, and finally, the f_{19} is the duration of the motion. Similar features are extracted for orientation, acceleration, angular speed, and velocity properties. The performance of the reference method created by using these Rubine features in the 6DMG data set is given in Table 2.4 together with the results of the proposed method. The Rubine features are normalized before the classification process because of the fairly comparison.

$$f_1 = \frac{x_3 - x_1}{\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2}} \tag{2.1}$$

$$f_2 = \frac{y_3 - y_1}{\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2}} \tag{2.2}$$

$$f_3 = \frac{z_3 - z_1}{\sqrt{(x_3 - x_1)^2 + (z_3 - z_1)^2}} \tag{2.3}$$

$$f_4 = \frac{x_N - x_1}{\sqrt{(x_N - x_1)^2 + (y_N - y_1)^2}} \tag{2.4}$$

$$f_5 = \frac{y_N - y_1}{\sqrt{(x_N - x_1)^2 + (y_N - y_1)^2}} \tag{2.5}$$

$$f_6 = \frac{z_N - z_1}{\sqrt{(x_N - x_1)^2 + (z_N - z_1)^2}} \tag{2.6}$$

$$f_7 = \sqrt{(x_{max} - x_{min})^2 + (y_{max} - y_{min})^2 + (z_{max} - z_{min})^2} \tag{2.7}$$

$$f_8 = \arctan\left(\frac{y_{max} - y_{min}}{x_{max} - x_{min}}\right) \tag{2.8}$$

$$f_9 = \arctan\left(\frac{z_{max} - z_{min}}{x_{max} - x_{min}}\right) \tag{2.9}$$

$$f_{10} = \sqrt{(x_N - x_1)^2 + (y_N - y_1)^2 + (z_N - z_1)^2} \tag{2.10}$$

$$f_{11} = \sum_{k=2}^{N-1} \sqrt{\Delta x_k^2 + \Delta y_k^2 + \Delta z_k^2} \tag{2.11}$$

$$f_{12} = \sum_{k=2}^{N-1} \Theta_k \tag{2.12}$$

$$f_{13} = \sum_{k=2}^{N-1} \gamma_k \tag{2.13}$$

$$f_{14} = \sum_{k=2}^{N-1} |\Theta_k| \tag{2.14}$$

$$f_{15} = \sum_{k=2}^{N-1} |\gamma_k| \tag{2.15}$$

$$f_{16} = \sum_{k=2}^{N-1} \Theta_k^2 \tag{2.16}$$

$$f_{17} = \sum_{k=2}^{N-1} \gamma_k^2 \tag{2.17}$$

$$f_{18} = max_{k=2}^{N-1} \left(\frac{\Delta x_k^2 + \Delta y_k^2 + \Delta z_k^2}{\Delta t_k^2} \right) \tag{2.18}$$

$$f_{19} = t_{k-1} - t_0 \tag{2.19}$$

In Table 2.4, the performance of the proposed method is clearly seen in all the sensor data. According to Table 2.4, the proposed method has significantly better performance than the reference method for all sensors. The reference method and the proposed method are based on some statistical features. However, the main difference between the proposed method and the reference method is the use of the wavelet transform for changing the data domain from a spatial-temporal space to a scale-frequency space. According to experimental results, this change is a result of the significant performance characteristic of the sensor data. Therefore, according to experimental results, it can be said that the proposed wavelet-based features are more robust against the movement differences of different people or the same people at different times than the reference method.

Table 2.4 Comparative Results

	Position	**Orientation**	**Acceleration**	**Angular Angular**	**Velocity**
Reference Method	81.56	75.29	56.57	52.98	55.26
Proposed Method	88.04	89.66	72.27	76.11	85.85

In the reference method results given in Table 2.4, all the features in Table 2.4 were extracted separately for each sensor data. Since the statistical features in the reference method are more suitable for the position sensor, the performance rate is 81.56%, but the performance rate in other sensor data is very low. Although the reference method achieves the best result in the position sensor, the proposed method takes precedence over the reference method. The difference between the best and the worst

sensor performance in the reference method is 28.58%. In the proposed method, this difference is 17.39%. This means that the proposed method provides a certain performance independent of the sensor. It does not include features compatible with a single sensor as in the reference method. The proposed wavelet-based features have achieved a certain success in all sensors.

In addition, when the majority voting algorithm is added to the design, performance rises to 92.65% in a model in which the majority of the results of position, orientation, acceleration, angular speed, and velocity are the ultimate class.

This chapter does not focus on finding the optimal classifier for the proposed feature set. Chapter 3 deals with this issue. Chapter 3 discusses the success of the SVM classifier. It also provides a number of methods to reduce the size of the data-set to increase the success of SVM.

3
Dimension Reduction and SVM based Gesture Recognition Architecture

In this section, dimensional reduction techniques and the SVM classifier used in the study are explained. It also includes the proposed method and the preparation of the test setup for this method. A scenario is created from the decisions made after the test results are known. With this scenario, the outputs of the movements classified in the MATLAB tool are input to the Flight-Gear Simulator. There are two important stages in this section. One is dimension reduction, and the other is classification. Unlike Chapter 2, the data-set here is divided into two parts: Training Data and Test Data.

The effect of the dimension reduction operation on the success of the gesture data-set is investigated in this section. The kNN classifier was used in the first stage to reveal the value of the wavelet-based feature with ease of implementation. However, its success in the large data space is limited. The motivation here is to examine the success of SVM for data and the effect of dimension reduction in combination with SVM.

3.1 Dimension Reduction

Analysis of high dimensional data sets is mathematically difficult. One of the problems with high-dimensional data sets is that not every measured value in the data set is "significant". It is preferable to reduce the size of the data-set rather than modelling of it [26]. There are many techniques for dimension reduction in the literature. [23] Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) methods were applied in this study because linear dimension reduction techniques are simpler and easier to apply than non-linear techniques. The comparison of PCA and LDA is summarized in the Table 3.1.

Dimension reduction is the projection of a data-set with a large number of features into a less dimensional sub-space. This will also cause the computational cost [27]. It offers the opportunity to work in a more optimum feature space.

Table 3.1 Differences between PCA and LDA

PCA	**LDA**
linear	linear
unsupervised	supervised
component axes that maximize the variance	maximizing the component axes for class separation

As explained in Chapter 2, the dimension of the new data-set after the feature extraction step is 48. Our goal is to reduce the dimension of the data-set to a more appropriate value using these two dimension reduction techniques.

3.1.1 Principal Component Analysis

Principal Component Analysis (PCA) is an unsupervised and linear dimension reduction technique [28], [29], [27]. In addition, PCA is a widely used dimension reduction method. PCA is a statistical method that converts a group of correlated variables into a group of uncorrelated variables. PCA finds several orthogonal linear combinations (PCs) of the original variables with the largest variance and tries to reduce the size of the data using that technique. PCA can be used in data analysis to study the relationships between a group of variables.

For example, to reduce an n-dimensional data set $(x^{(1)}, x^{(2)}, ..., x^{(m)})$ to k dimensions, following equations should be applied sequentially [27].

$$x_j^i = \frac{x_j^i - \mu_j}{\sigma_j}, \quad \forall j \tag{3.1}$$

$$\sum = \frac{1}{m}\sum_i^m (x_i)(x_i)^T, \quad \sum \in R^{n*n} \tag{3.2}$$

$$U = \begin{bmatrix} | & | & | \\ u_1 & u_2 ... & u_n \\ | & | & | \end{bmatrix}, \quad u_i \in R^n \tag{3.3}$$

$$u^T \sum = \lambda\mu \tag{3.4}$$

$$x_i^{new} = \begin{bmatrix} u_1^T x^i \\ u_2^T x^i \\ ::::: \\ u_k^T x^i \end{bmatrix}, \quad \in R^k \tag{3.5}$$

First of all, it is necessary to standardize the data set and, for this, calculate the mean and standard deviation for each feature in (3.1). The equation (3.2) gives the calculation of the co-variance matrix for the entire data set. Co-variance shows the relationship between two features. The eigenvalue and eigenvector of the co-variance matrix given in (3.3) are calculated as in (3.4). The eigenvalues are ordered, and the top eigenvector k is selected. Calculation of new data is also given in the equation (3.5). Thus, new k-dimension data is obtained from n-dimensional data.

3.1.2 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a commonly used dimension reduction method for the preprocessing phase in machine learning applications. The LDA dimension reduction technique projects the data into a lower-dimensional vector space by maximizing the ratio of the variance between classes to the variance within classes. LDA and PCA are very similar approaches in that they maximize data variance. LDA, on the other hand, is supervised, whereas PCA is unsupervised and ignores class distinction. LDA aims to find a feature sub-space that maximizes multiple class separation [27], [30], [31].

The following steps should be followed for LDA [32] :

- First, the means of each data-set class and the entire data-set are calculated. From here, the d-dimensional mean vector is obtained.
- Intra-class and inter-class scatter matrices are calculated.
- The eigenvectors and their respective eigenvalues are calculated for the scatter matrices 3.6.
- Eigenvectors are ordered by decreasing the eigenvalues. Eigenvectors k with maximum eigenvalues are chosen. From here, a W matrix in dxk form is obtained. Here, each column of the W matrix gives an eigenvector.
- Using this d×k eigenvector matrix, the samples are projected into the new subspace.

$$S_W^{-1} S_B W = \lambda W \tag{3.6}$$

In the equation 3.6, S_W refers to the intra-class scatter matrix and S_B refers to the inter-class scatter matrix. W is transformation matrix and λ is also eigenvalues [33].

3.2 Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning algorithm used in classification and regression problems. SVM, which is an algorithm proposed by Vapnik in 1995, is popular in machine learning studies such as image recognition, speech recognition, data mining etc. [34].

SVM basically aims to separate data between two classes with many linear planes. This decision boundary created between the two classes is known as the hyper-plane. The hyper-plane is oriented at the maximum distance from the nearest data point of each class, [35].The reason for choosing the maximum margin is to make a better classification. In other words, even if there is a small error in the position of the hyperplane, the maximum margin is chosen so that it does not cause mis-classification [36].

SVM is originally a linear classification method. However, there is also the kernel method as an alternative use. Many problems in nature are non-linear. These problems are solved with the nonlinear model of SVM. An additional dimension is added to the actual data with the kernel function. And so the problem in the higher dimension becomes linear.

The 6DMG data-set we used in this study is not a linear data-set. However, considering the time and cost of the calculation mathematically, linear SVM was used. The success of linear SVM was compared with the tests performed before and after dimension reduction.

3.3 FlightGear Flight Simulator

FlightGear is a free and open source flight simulator. It can run on macOS, Linux, Windows, etc. platforms. With the FlightGear project, it was aimed to create an open source flight simulator to be used in various environments, such as academic or engineering industries. In this way, a flight simulator environment has been prepared that can be used by anyone who wants to contribute. Features of the simulator [37];

- Freedom: it is open-source project. It means that developed source codes are easily accessible.
- Flight Dynamics Models: In addition to the flight dynamics models presented, new models can be added.
- Extensive World Scenery Database
- Detailed Sky Model
- Flexible and Open Aircraft Modeling System

FlightGear is compared with similar platforms in the Table 3.6. In this study, FlightGear was chosen because of the possibilities and convenience it provides for use. Also, the 2018.3 version of FlightGear is used.

Table 3.2 Flight Simulator Properties

	Platforms	**Free Option**	**Creating Your Own Add-Ons**	**Open Source**
MSFS [38]	DOS, Windows, Classic Mac OS, Xbox One	NO	YES	NO
Prepar3D [39]	Windows 7,8,10	NO	YES	NO
X-Plane [40]	Mac OS, Linux, Windows	NO	YES	NO
FlightGear [37]	Windows, Mac, Linux, etc.	YES	YES	YES

FlightGear is a popular flight simulator for a variety of purposes in academic studies [41], [42], [43]. Some studies describe how the FlightGear simulator enables data exchange with external programs. Some studies use FlightGear to simulate flight control algorithms. Data can be received from FlightGear or sent to FlightGear. In this study, certain commands are sent to FlightGear with the information obtained from MATLAB.

3.4 Proposed Method

The proposed method is given in the Figure 3.1. ZN_i is formed by normalizing the data obtained from the features suggested in chapter 2. Unlike the previous chapter,

it aims to create a real test environment by separating data into train and test data.

The training data consists of data of 5 randomly selected 21 right hand users in the 6DMG data-set. This means that approximately 18% of all the data-set is reserved for training purposes. The test data is created by a person selected from among 7 left-handed users and 16 right-handed users who are not included in the training data. A random trial of the relevant movements is taken, according to the predefined gesture set of this person.

Due to the large size of the data-set, two different linear dimension reduction techniques have been applied. The new dimensional data-set obtained was classified with the linear SVM.

In the Figure 3.1, Tr_i represents train data and Te_i represents test data. Here i refers to sensors ($i = 1, ..., N$), as in chapter 2. Here the size of Tr_i is mxd and the size of Te_i is nxd. Here m is equal to the number of gestures selected for the train data-set. It is around 1000 levels. n is equal to the number of movements in the action scenario defined for each prepared test. d depends on the 16 features obtained from the wavelet theory and the size of the associated sensors. For example, d for the position sensor is 48, while d for the orientation sensor is 64. These values are quite challenging for the classifier.

It is aimed to transform these d-dimensional data into k-dimensions with the dimension reduction techniques ($k << d$). The k value is equal for each sensor and its value is 10. The size of $Tr_i^{'}$ obtained after dimension reduction was mxk, while the size of $Te_i^{'}$ was nxk. Thus, the reduction in size makes the calculation more efficient. After dimension reduction process, linear SVM classifier is trained with train data $Tr_i^{'}$. The test data $Te_i^{'}$ is also given to the trained model and the prediction results are obtained.

3.5 Test Setup and Simulation Results

The train data set is generated from 5 randomly selected right-hand users. The test data is selected with the following steps:

- A gesture set is prepared for the test.
- A random individual is selected from the set of 16 right- and left-handed individuals not used in the training data.
- Random trial of the selected person is selected according to the gesture set.

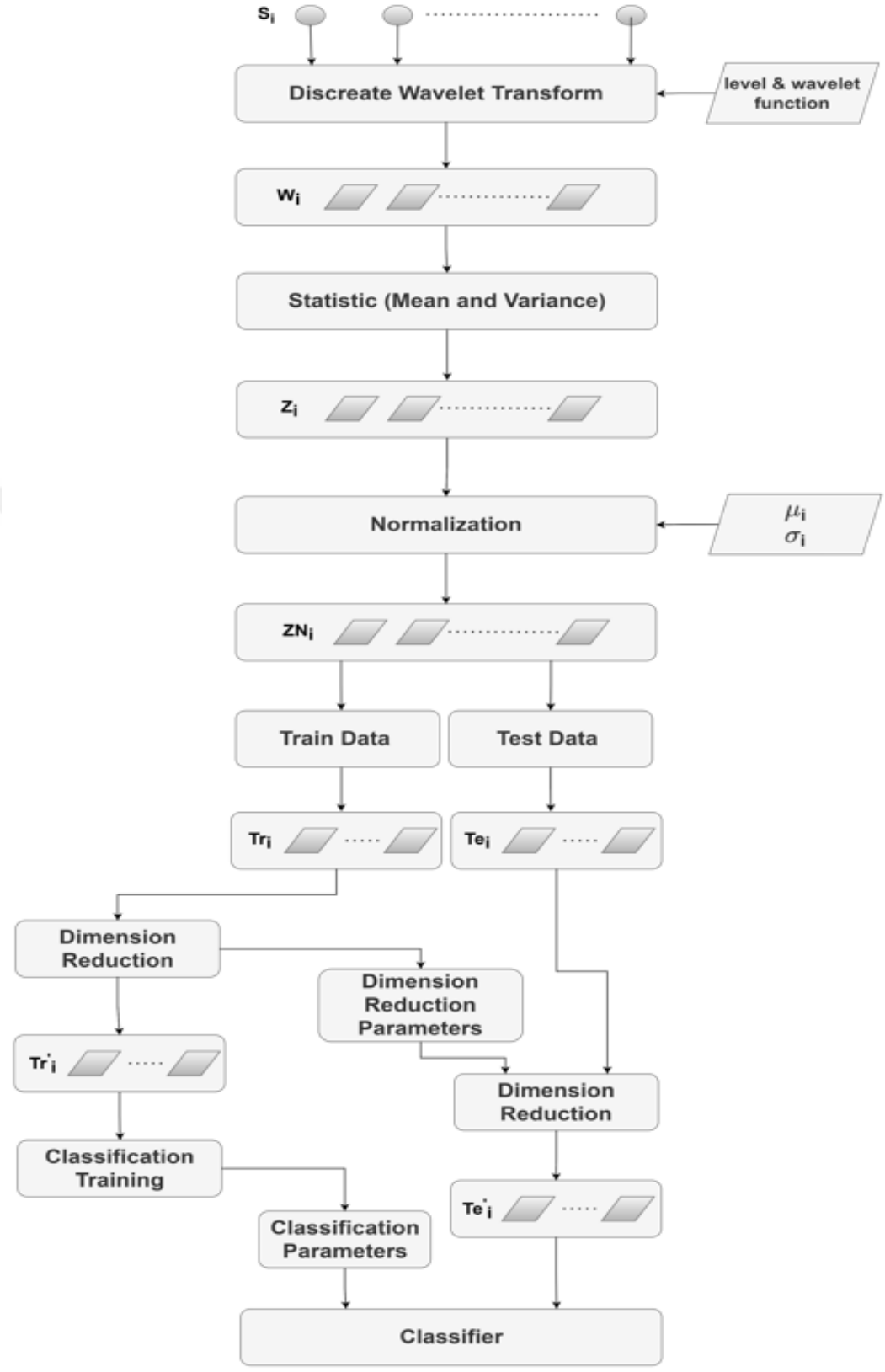


Figure 3.1 Proposed Method

These experiments are repeated 100 times and the success rate is calculated. In the first experiment, a movement set with all the movements was created. Gesture set (GS) is GS = [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20]. Each element in the GS represents the index of a gesture. The indexes corresponding to the movements are given in Appendix A. This experiment was conducted to decide which method and which gesture was more successful for this data set. In Table 3.3, the performance of the classifiers is given for each sensor. The majority voting results of 5 sensors are given in the Table 3.4. Table 3.5 shows the success of all gestures in 4 different methods. From the 3.3, 3.4 and 3.5 tables obtained in this experiment, it is seen that LDA+SVM is more successful than the other methods.

Table 3.3 Classifier Selection Experiments

	Position	**Orientation**	**Acceleration**	**Angular Speed**	**Velocity**
K-NN	48.05	59.40	49.15	59.90	24.50
SVM	61.45	66.35	55.65	65.90	28.20
PCA+SVM	60.75	65.70	54.20	66.50	27.30
LDA+SVM	68.35	73.25	57.25	65.85	60.15

Table 3.4 Majority Voting Results

	Accuracy(%)
K-NN	64.80
SVM	76
PCA+SVM	75.75
LDA+SVM	82.20

We selected the gestures that had more than 80% success and removed the others. In this test, new gesture set is GS = [1 2 3 4 7 8 9 11 13 14 17 18 19 20]. After selecting the LDA+SVM success from the previous experiment, this experiment was performed only for LDA+SVM. LDA+SVM success of the experiment shows in Table 3.6.

Among the new gesture eight gestures that will be equivalent in Flight Simulator have been selected. While choosing these gestures, it was tried to choose meaningful movements that matched the movement of the aircraft that could be easily understood by the users. For example, the "swipe up" gesture was chosen to move the aircraft up. The gestures to be used in the scenario and the corresponding aircraft movement are given in the Table 3.7 in detail. The success of GS= [1 2 3 4 9 11 17 18], in which these 8 movements are given sequentially, is given in the Table 3.8.

A meaningful scenario was prepared with the gestures in Table 3.7. Since it is complex to give commands such as raising and lowering the aircraft to the simulator, the aircraft is launched at a certain altitude at the initial time. In the specified scenario, it is

Table 3.5 Success Comparison of Gestures

Gestures	K-NN	SVM	PCA+SVM	LDA+SVM
1	72	85	88	93
2	62	83	81	80
3	96	94	94	95
4	82	88	86	90
5	67	72	71	78
6	49	66	68	69
7	57	80	79	81
8	51	78	78	84
9	74	76	78	84
10	53	70	68	80
11	71	82	81	87
12	37	57	58	73
13	60	72	72	89
14	47	68	67	83
15	60	58	58	54
16	50	54	54	73
17	74	86	86	87
18	65	82	81	84
19	86	87	85	96
20	83	82	82	84

Table 3.6 Success for New Gesture Set

	Position	Orientation	Accelera-tion	Angular Speed	Velocity	Majority Voting
LDA+SVM	69.50	77.64	60.57	71.21	62.93	86.71

assumed to be launched after the aircraft has completed its take-off. This scenario is as follows:

- Landing gear is closed.
- Engine is accelerating.
- Swipe up and down movements are performed at a certain time interval, respectively.
- Left and right turns are also performed at a certain time interval sequentially. However, after each of these movements, a swipe down movement should be made. These movements in the simulator must be performed with real precision.
- Swipe up and down movements are performed again at a certain time interval, respectively.

Table 3.7 Gestures Selected for the Scenario

Gesture Figure	Gesture Definition	Equivalent in Simulator
	SWIPE UP	Positive movement of the aircraft on the lateral axis (Pitch+)
	SWIPE DOWN	Negative movement of the aircraft on the lateral axis (Pitch-)
	SWIPE RIGHT	a combination of positive movement of the aircraft in the longitudinal axis and positive movement in the lateral axis (Roll+, Pitch+)
	SWIPE LEFT	a combination of negative movement of the aircraft in the longitudinal axis and positive movement in the lateral axis (Roll-, Pitch+)
	POKE RIGHT	Engine power is increased.
	POKE UP	Engine power is reduced.
	CIR VER CLK	Landing gear is opened
	CIR VER CCLK	Landing gear is closed

Table 3.8 Success of the Gestures Selected for the Scenario

	Position	**Orientation**	**Accelera-tion**	**Angular Speed**	**Velocity**	**Majority Voting**
LDA+SVM	71.5	79.13	64.0	71.0	62.37	88.87

- Engine slowing down
- Landing gear is opened.

The resulting motion set selected is {CIR VER CCLK, POKE RIGHT, SWIPE UP, SWIPE DOWN, SWIPE RIGHT, SWIPE DOWN, SWIPE LEFT, SWIPE DOWN, SWIPE UP, SWIPE DOWN, POKE UP, CIR VER CLK}. The success of this scenario GS= [18 9 3 4 1 4 2 4 3 4 11 17] is given in the Table 3.9.

Table 3.9 Success of the Scenario

	Position	**Orientation**	**Accelera-tion**	**Angular Speed**	**Velocity**	**Majority Voting**
LDA+SVM	70.50	72.25	57.16	68.91	59.00	84.91

The setup flow prepared for the experiments is given in the Figure 3.2. First, the database is imported into the MATLAB environment. Here, chapter 2 and all the work including the classification process are done in this chapter. Then the estimate of the movements is written to a text file. Then the estimate of the movements is written to a text file. This file is run with Python [44] code, and the gestures associated with the gestures are sent to the FlightGear.

Images of the scenario are given below, respectively. In Figure 3.3, the image of the plane is given before the scenario starts. The landing gear which is open, are closed with CIR VER CLK command and the engine speed is increased in Figure 3.4. In Figure 3.5, the plane moves upwards with the swipe up gesture. A swipe down gesture is applied to flat the aircraft rising with a swipe up movement, and the state of the aircraft after this gesture is as seen in Figure 3.6. Then the swipe right gesture is applied to the plane. This gesture consists of two stages. First, the ailerons move in the positive direction and as a result of this gesture, the plane is positioned as in Figure 3.7.a. Then, as with the swipe up gesture, the positive movement of the elevators is performed in Figure 3.7.b. The swipe left gesture is similar to the swipe right movement. Only the ailerons move in the negative direction. Afterwards, this movement is followed by the positive movement of the elevators. This movement is also given in Figure 3.8. After the scenario is completed, the aircraft returns to its initial position given in Figure 3.3, the engine speed decreases, and the landing gear opens.

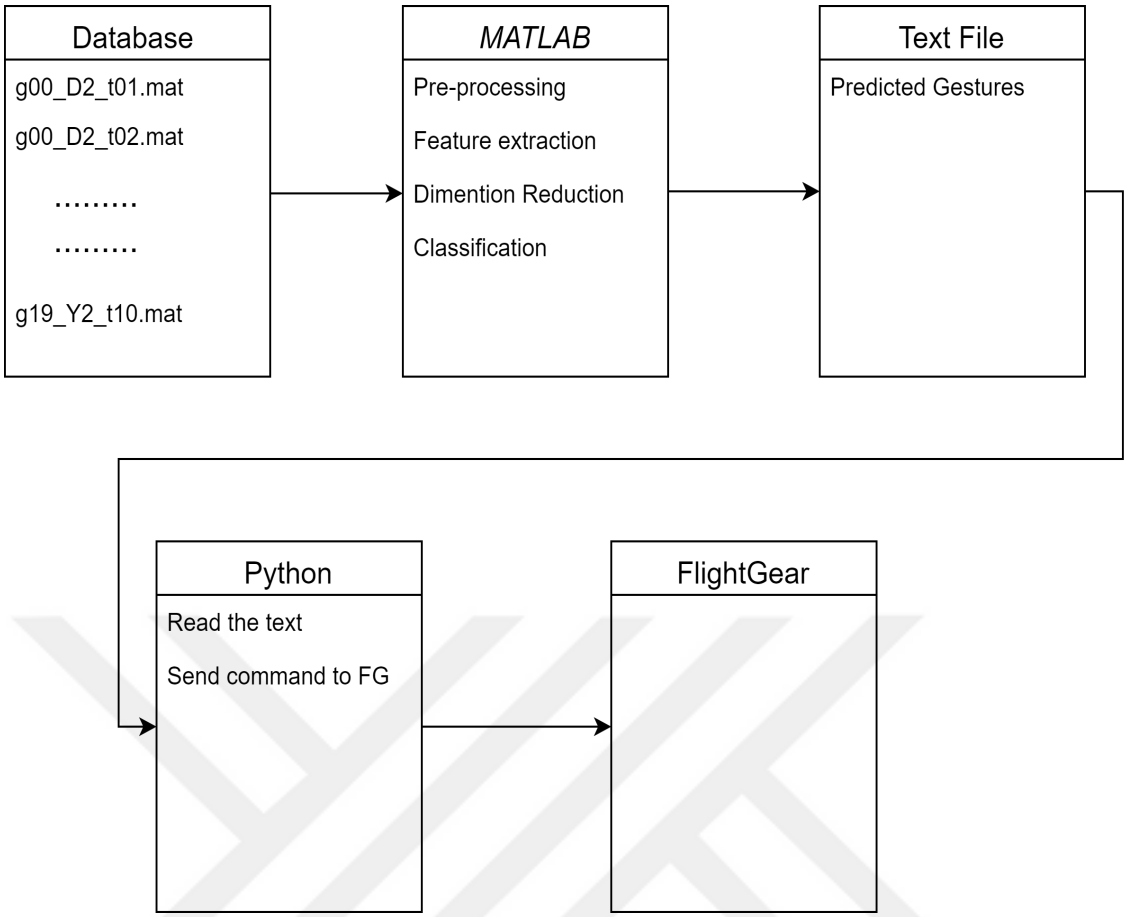


Figure 3.2 Connection of the Tools for the Experiment Setup

Figure 3.9 shows instantaneous data from the web browser while the flight simulator is running. The movement of the aircraft in its axis from one wing tip to the other wing tip can be seen in Figure 3.9.a. This movement is also called pitching. In Figure 3.9.b, the movement of the aircraft on the axis extending from the nose to the tail is given. From this graph, it can be understood that the roll angle is positive and then negative, since this movement is made in the simulation in right and left turns.

The image of the aircraft on the map during the simulation period is given in Figure 3.10. After a certain period of straight movement, right and left turn movement can also be seen on the map.

In summary, in this chapter, test data and training data are separated from each other as in real applications. Then, the size of the data was reduced with the dimension reduction techniques, and the studied space was optimized. This is also reflected in the results obtained. Significant differences show up for each type of data between the success before and after the application of dimension reduction techniques. The last process is the classification step. This step is followed by the implementation of the majority voting algorithm. As in the previous chapter, the multi-sensor results of the

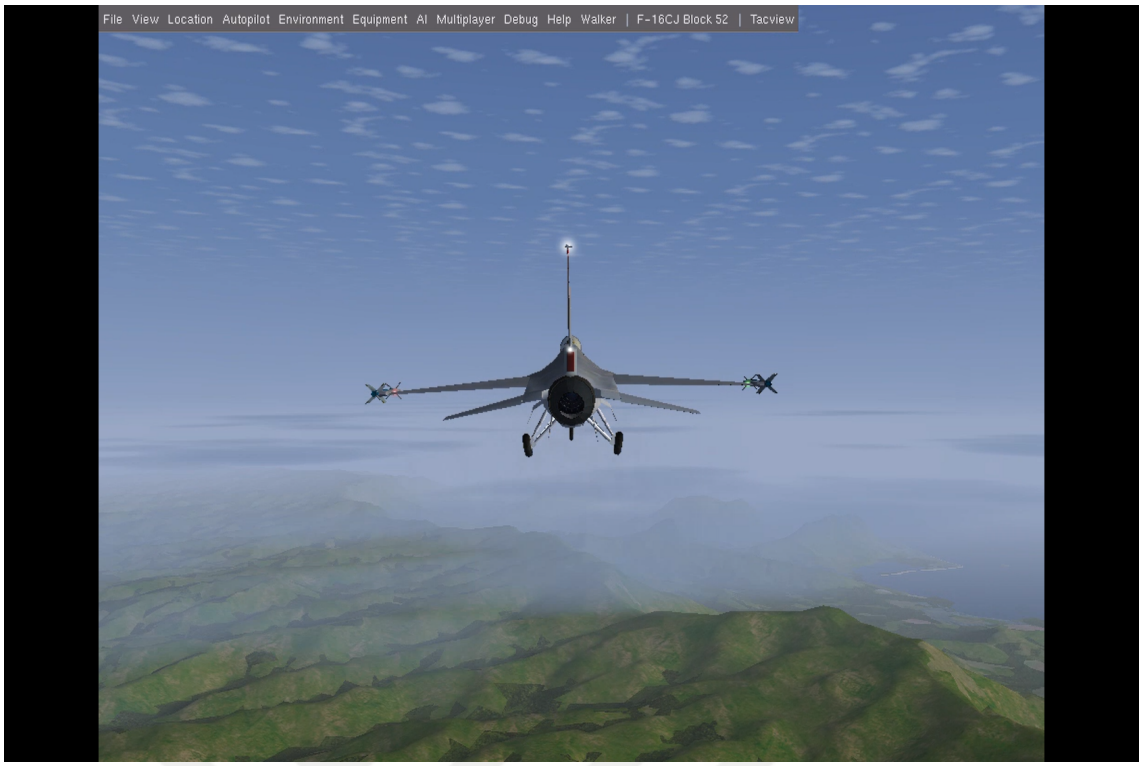


Figure 3.3 Initial state of the scenario

Figure 3.4 Simulation status as a result of the combination of CIR VER CLK and POKE RIGHT commands

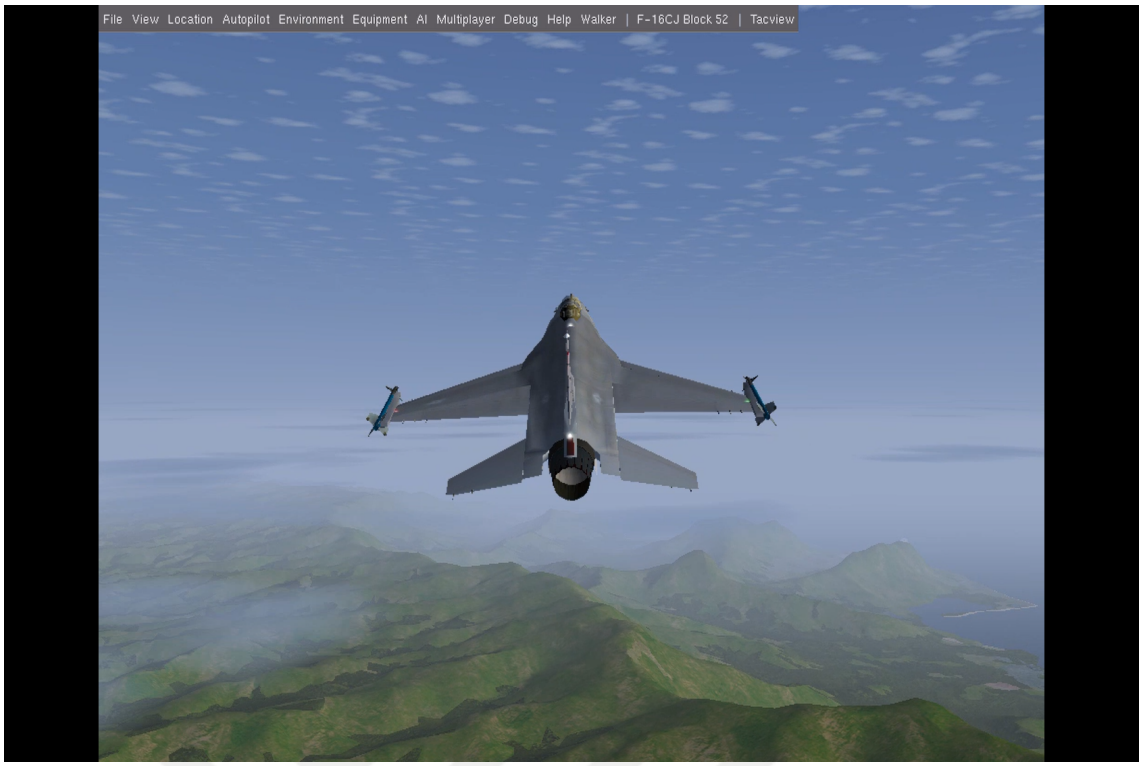
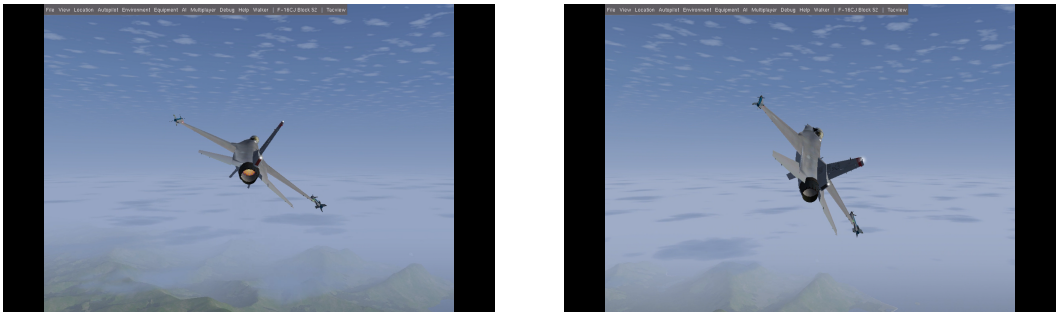


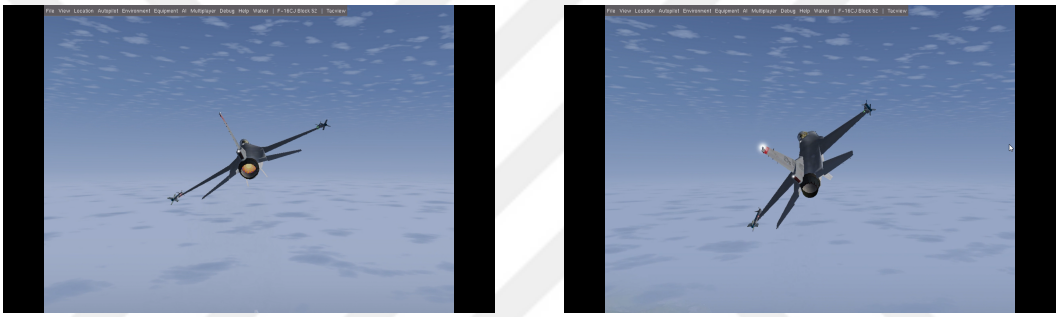
Figure 3.5 FG state with SWIPE UP command

Figure 3.6 FG state with SWIPE DOWN command



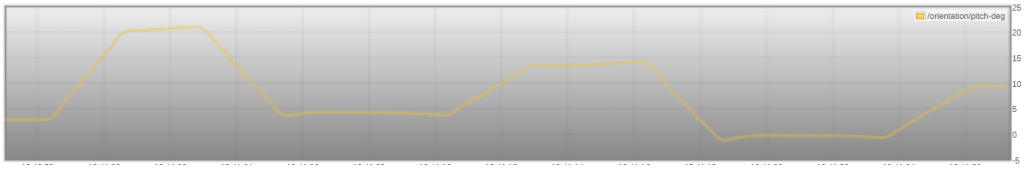
(a) aileron movement **(b)** elevator movement

Figure 3.7 Aircraft behaviour under SWIPE RIGHT command



(a) aileron movement **(b)** elevator movement

Figure 3.8 Aircraft behaviour under SWIPE LEFT command



(a) pitch degree

(b) roll degree

Figure 3.9 Data from the aircraft throughout the simulation

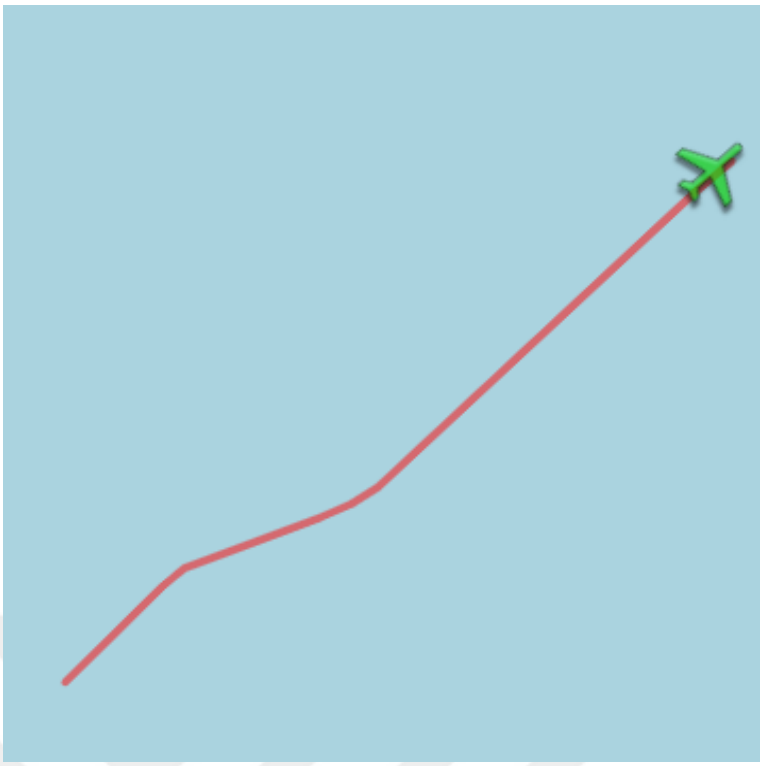


Figure 3.10 The image of the aircraft on the map during the simulation

data received from the multi-sensors are examined. Whatever the majority of these sensors decide is the final result.

4
Dominant Set Clustering in Gesture Recognition

In this chapter, the dominant set clustering approach is used in gesture recognition. This chapter includes;

- Dominant set clustering and literature studies
- Proposed method which is presented in this study
- The experimental results obtained in MATLAB

The motivation in this section is to achieve well-trained pilot success with less trained pilots. In civil and military applications, under-trained pilots may be inadequate in critical situations, but an optimum solution can be achieved with the decision of more than one of them.

An example is the case of flying drones between skyscrapers for civil applications. In the introduction, it was mentioned that UAVs have started to be used in the field of transportation. In this area, UAVs, whose behavior in the city is mostly the same, can remain in critical situations among skyscrapers in developed cities. In addition, as this service industry increases, the control of UAVs will become more difficult.

In military applications, more critical situations are encountered than in the civilian field. Advances in technology in unmanned aerial systems will become equal to or superior to manned warplanes. It is very difficult for unmanned combat aerial vehicles (UCAV) to make fully autonomous decisions [45]. The dominant set method provides an infrastructure that can be applied to UAVs with human-machine cooperation. In a dogfight, less-trained pilots can make decisions like a trained pilot.

4.1 Dominant Set Clustering

[46] The clustering problem has been discussed in a variety of fields, including machine learning, data mining, image processing, and computer vision. There is no

ecumenically accepted definition for the word "cluster". All objects that have a higher similarity than other objects should be in a cluster.

Pavan and Pelillo proposed the graph-based dominant cluster method for data clustering [47]. In the proposed method, a set definition based on graph theory is made and a structure is established upon meeting two conditions:

- All elements in a set must be similar to each other.
- All elements outside of a cluster must not resemble elements in that cluster.

Pavan and Pelillo [46] established a weighted and undirected graph in their studies and established the structure that satisfies these two conditions. However, the solution of the proposed method may not be suitable due to the complexity of the process. To overcome this problem, the use of a simplistic and continuous optimization method called replicator dynamics, presented in Equation 4.1 has been proposed in the related publication [46].

$$x_i(t+1) = x_i(t)\frac{(Wx)_i}{x(t)^T W x(t)} \tag{4.1}$$

where $i = 0, ..., n$ is the discrete-time version of first-order replicator equations. W is the similarity matrix and it is symmetrical.

First, the initial value of the vector x is assigned. Over time, the vector x is updated using Equation 4.1 . When the difference between the two iterations is small enough, the process is terminated. The elements are divided into two clusters according to the x values obtained. When the value of an element in this vector is greater than 0, that element is considered to be in the dominant set.

The dominant set approach has been made applicable to large grouping problems such as segmentation of high resolution images [48]. The band clustering performance evaluation method is applied to a relatively small portion of it, which contains the relevant classes, rather than to the entire high-dimensional hyper-spectral data. Therefore, the computational cost of the proposed method for determining the performance of spectral bands is low. It is assumed that this method can be used as a pre-processing step that can be applied before feature selection in hyper-spectral images [49].

4.2 Proposed Method

The proposed method with Dominant Set Clustering is shown in Figure 4.1. Here, the training data is generated as in Chapter 3. The test data, however, are quite different. The main idea in this chapter is that a certain number of people are assumed to have made the same gesture. For example, it is guaranteed that 4 people out of 9 random people do the same gesture, and the test data is generated by taking random gestures from the other 5 people. Then, the dominant cluster is obtained from this test data with the dominant set clustering method. This obtained cluster is classified with the kNN classifier.

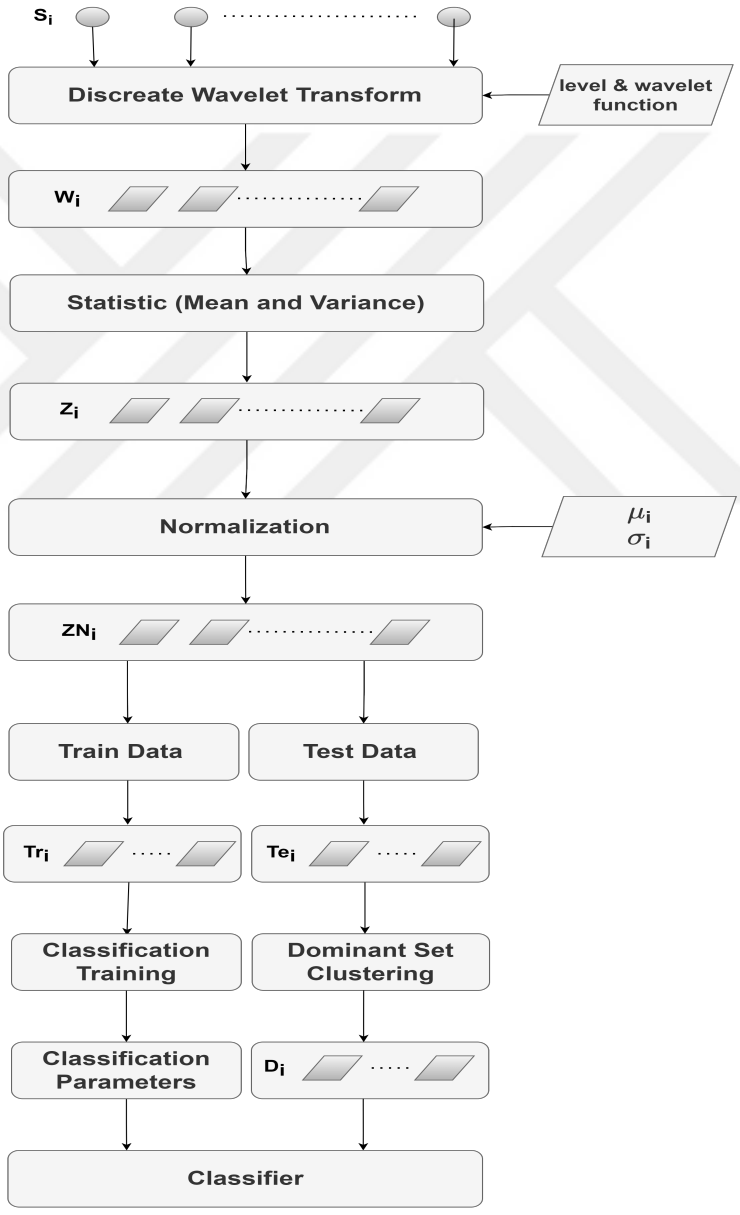


Figure 4.1 Proposed Method

Assuming that it is kxd in size of Te_i in the Figure 4.1, k is proportional to the number of people selected. D_i is the data-set with dominant clusters from this test data (Te_i) and its size is lxd (l<k). i refers to the number of sensors as in other chapters and i = 1 can be up to the N.

The similarity between two people (a and b) is calculated as in the equation 4.2. Here, the α parameter is critical for the dominant set. In the equation, $d(a,b)$ gives the Euclidean distance between person a and b, while α is the decay parameter.

$$s(a,b) = exp(-\frac{d(a,b)}{\alpha}) \tag{4.2}$$

This equation 4.2 for two persons is applied among all persons selected for test data. Thus, the similarity matrix between each selected person and other people is created. This similarity matrix corresponds to the weight matrix (W) matrix in the replicator dynamics equation 4.1.

4.3 Experiments and Determining the Parameter Space

In this section, various experiments related to the dominant-set have been made. As mentioned in the previous section, the α parameter is important in the dominant set method. Thus, firstly, an experiment was created to find the optimum α parameter. In this experiment, 5 right hand users were determined, and the remaining 9 people were randomly selected. Random gestures were selected so that 4 people out of these 9 would do exactly the same gesture. The prepared data were given to the dominant-set by finding similarity matrices in the relevant α values. People in the dominant cluster are assigned a value of one, while those in the other are assigned a value of zero. At the same alpha value, these experiments are performed in 1000 loops and recorded in the relevant α value in the Table 4.1. In the Table 4.1, D1 goes as far as D9, denoting the 1st person. Ideally, this table is expected to be D = [1.0 1.0 1.0 1.0 0 0 0 0 0]. With this experiment, the closest α value was tried to be selected.

As seen in the Table 4.1, 3 was chosen as the most appropriate alpha value. Here, the values up to D1 and D4 are pretty close to 1. In the elements between D5 and D9, the closest values were obtained compared to the others.

In this experiment, 4 people are guaranteed to do the same gesture, but there is a situation like this. Among the other five people, there may be those who make this gesture. While α 3 is selected in the Figure 4.2, the expected and dominant set of results for these 9 people are given.

Table 4.1 α parameter selection experiment

α	D1	D2	D3	D4	D5	D6	D7	D8	D9
0.001	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
0.01	0.95	0.95	0.95	0.95	0.93	0.93	0.93	0.92	0.92
0.03	0.50	0.52	0.47	0.50	0.17	0.18	0.17	0.18	0.18
0.07	0.54	0.55	0.54	0.54	0.20	0.19	0.17	0.20	0.19
0.1	0.58	0.56	0.56	0.58	0.23	0.24	0.23	0.23	0.23
0.3	0. 77	0.75	0.75	0.75	0.28	0.28	0.29	0.30	0.29
0.7	0. 77	0.78	0.77	0.77	0.28	0.31	0.32	0.29	0.30
1	0.77	0.77	0.79	0.77	0.25	0.24	0.25	0.24	0.26
3	0.91	0.90	0.89	0.90	0.31	0.33	0.32	0.34	0.32
7	0.94	0.94	0.94	0.95	0.42	0.42	0.42	0.43	0.42
10	0.95	0.96	0.95	0.95	0.45	0.47	0.45	0.45	0.45
30	0.99	0.99	0.99	0.99	0.75	0.77	0.76	0.76	0.77

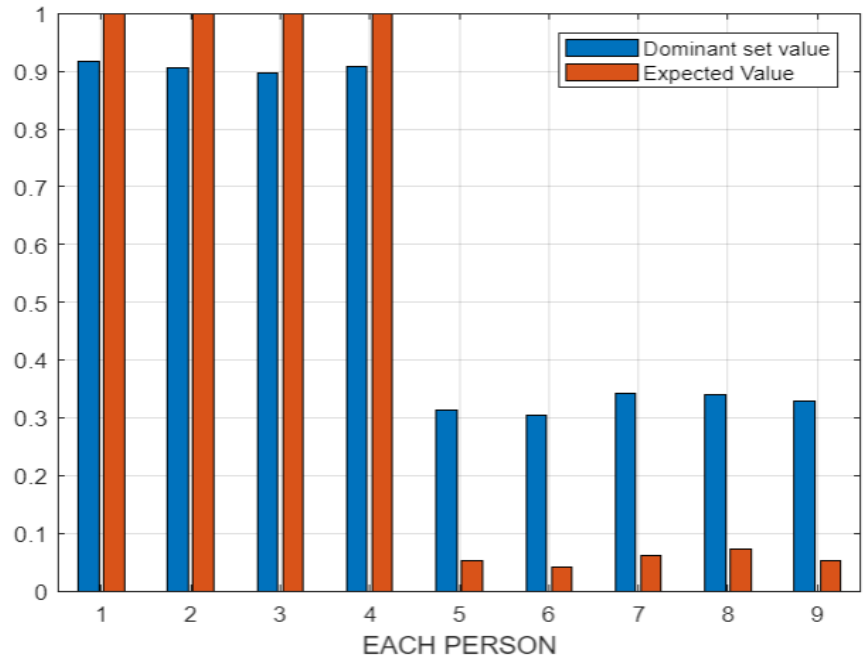


Figure 4.2 Expected and dominant-set results

After selecting the alpha value of 3, an experimental environment was prepared to measure the dominant set's success. Here, three different methods are compared.

- One of the selected persons who performs the correct gesture is given to the classifier. So a person's decision is given to the system.
- All of the selected persons are given to the classifier and the majority decision is chosen to the system.
- The dominant cluster is removed from the selected person and the result is given

to the classifier. The decision of the majority is chosen.

The total number of people in the Table 4.2 was changed and the effect of this was examined. It was ensured that at least four people made the same movement in each experiment. The results in the Table were performed at 20000 cycles.

Table 4.2 The effect of the total number of users on the dominant set clustering

Total Users	Number of users making the same gesture	A person	Majority Voting	Dominant Set and Majority Voting
9	4	0.754	0.834	0.815
10	4	0.752	0.818	0.798
11	4	0.752	0.801	0.780
12	4	0.755	0.792	0.766
13	4	0.755	0.783	0.749

As can be seen in the Table 4.2, the success of a person selected from the total user base is stable in any case. After finding the dominant cluster, when we do majority voting, it is expected to be more successful. However, in the experiments, the majority voting result is higher among the dominant sets.

In Table 4.3, the total number of people remained constant. In each experiment, the number of people performing the same movement was increased. Here, as in the table 4.2, the success of a single person remained around 75% in all experiments. This can be called the success of the kNN (k=1) classifier.

Table 4.3 The effect of the total number of users making same gesture on the dominant set clustering

Total Users	Number of users making the same gesture	A person	Majority Voting	Dominant Set and Majority Voting
12	3	0.757	0.637	0.597
12	4	0.755	0.792	0.766
12	5	0.753	0.863	0.847
12	6	0.750	0.899	0.885
12	7	0.756	0.919	0.907
12	8	0.750	0.934	0.921
12	9	0.760	0.942	0.926

In the summary of this section, we aimed to investigate whether the dominant set clustering method can be used for UAVs. With this study, it is predicted that UAV control can be provided with less trained pilots in cases where there are insufficient pilots. It is a fact that dominant-set clustering can be more successful, by finding

the appropriate parameters. In the future study, it is aimed to find the optimum parameters by coding the dominant-set more effectively.

5
CONCLUSION

The user-independent situation is difficult due to the large intra-class differences between users. The efficiency of features derived from different sensors is compared in user-independent situations. Considering the spatial-temporal nature of gesture signals for recognition, an approach with a simple statistical feature-based non-parametric classifier worked on wavelet transformed signals is presented. While the success rate is 81% with only statistical features in the spatial-temporal domain, the performance of the position sensor reaches 89% in the Chapter 2 proposed method obtained by the wavelet transform. Considering that this value is obtained only from the first and second statistics, it clearly reveals the success of the proposed method. According to experimental results, it is demonstrated that the proposed wavelet-based features are more robust against the movement differences of different people or the same people at different times than the reference method.

In Chapter 3, LDA as a dimension reduction technique and SVM as a classifier were proposed. As in real applications, in this section the training data is separated from the test data and the classifier is trained with a training rate of 17.8%. Under the scenarios created for FG, the success of LDA+SVM was recorded at 84.91% with majority voting. The success of each gesture is different. The system was established with moves that had a high success rate. It is a problem where success will be limited by some difficulties brought by data. The data is open source and general hand gesture data. For the application used, it would be more appropriate to produce the data accordingly. Also, the success rate of the proposed method will increase if more data is used to training process. This method can be used in entertainment applications and in many risk-free areas such as the use of drones in non-urban areas. For critical use (inner city areas), it can be used by training the model with more data and making new studies.

One of the important studies in this thesis is dominant set clustering. It is the recommended method for the control of increased UAVs in various civil and military fields.A solution has been presented for the number of pilots anticipated to be

insufficient. According to the decision made by most of the less trained pilots, the direction of the UAV is determined. In the experiments, the dominant set was left behind by about 2%. It is very difficult to find the relevant parameters for the dominant set. Various data sets and parameter spaces could be used to obtain better results in the dominant set clustering method.

For future studies, the data-set will be specially produced for the application, and the parameter tuning in the dominant set will be detailed. In addition, success under different classifiers needs to be examined. In this thesis, an important method for future UAV control is proposed.

REFERENCES

[1] P. Pisharady, M. Saerbeck, "Recent methods and databases in vision-based hand gesture recognition: A review," *Computer Vision and Image Understanding*, vol. 141, pp. 152–165, Dec. 2015. DOI: 10.1016/j.cviu.2015.08.004.

[2] T. Kopinski, S. Geisler, L.-C. Caron, A. Gepperth, U. Handmann, "A real-time applicable 3d gesture recognition system for automobile hmi," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2014, pp. 2616–2622. DOI: 10.1109/ITSC.2014.6958109.

[3] S. Tateno, Y. Zhu, F. Meng, "Hand gesture recognition system for in-car device control based on infrared array sensor," in *2019 58th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, 2019, pp. 701–706. DOI: 10.23919/SICE.2019.8859832.

[4] J. Qi, G. Jiang, G. Li, Y. Sun, B. Tao, "Intelligent human-computer interaction based on surface emg gesture recognition," *IEEE Access*, vol. 7, pp. 61 378–61 387, 2019. DOI: 10.1109/ACCESS.2019.2914728.

[5] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, A. N. Skodras, "Deep learning in emg-based gesture recognition," in *PhyCS*, 2018.

[6] M. Carroll, N. Dahlstrom, "Human computer interaction on the modern flight deck," *International Journal of Human–Computer Interaction*, vol. 37, no. 7, pp. 585–587, 2021. DOI: 10.1080/10447318.2021.1890495. eprint: https://doi.org/10.1080/10447318.2021.1890495. [Online]. Available: https://doi.org/10.1080/10447318.2021.1890495.

[7] S. R. Dixon, C. D. Wickens, "Control of multiple-uavs: A workload analysis," ILLINOIS UNIV AT URBANA-CHAMPAIGN SAVOY AVIATION HUMAN FACTORS DIVISION, Tech. Rep., 2003.

[8] K. Yakushiji, H. Fujita, M. Murata, N. Hiroi, Y. Hamabe, F. Yakushiji, "Short-range transportation using unmanned aerial vehicles (uavs) during disasters in japan," *Drones*, vol. 4, no. 4, p. 68, 2020.

[9] A. Sarkar, K. A. Patel, R. Ganesh Ram, G. K. Capoor, "Gesture control of drone using a motion controller," in *2016 International Conference on Industrial Informatics and Computer Systems (CIICS)*, 2016, pp. 1–5. DOI: 10.1109/ICCSII.2016.7462401.

[10] C. Liu, T. Szirányi, "Gesture recognition for uav-based rescue operation based on deep learning.," in *Improve*, 2021, pp. 180–187.

[11] J. Liu, L. Zhong, J. Wickramasuriya, V. Vasudevan, "Uwave: Accelerometer-based personalized gesture recognition and its applications," *Pervasive and Mobile Computing*, vol. 5, pp. 657–675, Dec. 2009. DOI: 10.1016/j.pmcj.2009.07.007.

[12] S. Mitra, T. Acharya, "Gesture recognition: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 3, pp. 311–324, 2007. DOI: 10.1109/TSMCC.2007.893280.

[13] H.-K. Lee, J. Kim, "An hmm-based threshold model approach for gesture recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 961–973, 1999. DOI: 10.1109/34.799904.

[14] C. Amma, D. Gehrig, T. Schultz, "Airwriting recognition using wearable motion sensors," in *Proceedings of the 1st Augmented Human International Conference*, ser. AH '10, Megève, France: Association for Computing Machinery, 2010, ISBN: 9781605588254. DOI: 10.1145/1785455.1785465. [Online]. Available: https://doi.org/10.1145/1785455.1785465.

[15] M. Chen, G. Alregib, B.-H. Juang, "Feature processing and modeling for 6d motion gesture recognition," *Multimedia, IEEE Transactions on*, vol. 15, pp. 561–571, Apr. 2013. DOI: 10.1109/TMM.2012.2237024.

[16] D. Rubine, "Specifying gestures by example," in *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '91, New York, NY, USA: Association for Computing Machinery, 1991, pp. 329–337, ISBN: 0897914368. DOI: 10.1145/122718.122753. [Online]. Available: https://doi.org/10.1145/122718.122753.

[17] M. Hoffman, P. Varcholik, J. J. LaViola, "Breaking the status quo: Improving 3d gesture recognition with spatially convenient input devices," in *2010 IEEE Virtual Reality Conference (VR)*, 2010, pp. 59–66. DOI: 10.1109/VR.2010.5444813.

[18] R. Polikar, *The wavelet tutorial*. [Online]. Available: http://users.rowan.edu/~polikar/WTtutorial.html (visited on 02/15/2021).

[19] S. Orguc, H. S. Khurana, K. M. Stankovic, H. Leel, A. Chandrakasan, "Emg-based real time facial gesture recognition for stress monitoring," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 2651–2654. DOI: 10.1109/EMBC.2018.8512781.

[20] R. Agarwal, B. Raman, A. Mittal, "Hand gesture recognition using discrete wavelet transform and support vector machine," in *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, 2015, pp. 489–493. DOI: 10.1109/SPIN.2015.7095326.

[21] A. Sharma, D. K. Kumar, "Moments and wavelets for classification of human gestures represented by spatio-temporal templates," in *AI 2004: Advances in Artificial Intelligence, 17th Australian Joint Conference on Artificial Intelligence, Cairns, Australia, December 4-6, 2004, Proceedings*, G. I. Webb, X. Yu, Eds., ser. Lecture Notes in Computer Science, vol. 3339, Springer, 2004, pp. 215–226, ISBN: 3-540-24059-4. [Online]. Available: http://springerlink.metapress.com/openurl.asp?genre=article&issn=0302-9743&volume=3339&spage=215.

[22] M. N. Al-Berry, H. M. Ebied, A. S. Hussein, M. F. Tolba, "Human action recognition via multi-scale 3d stationary wavelet analysis," in *2014 14th International Conference on Hybrid Intelligent Systems*, 2014, pp. 254–259. DOI: 10.1109/HIS.2014.7086208.

[23] R. O. Duda, P. E. Hart, D. G. Stork, C. R. O. Duda, P. E. Hart, D. G. Stork, *Pattern classification, 2nd ed*, 2001.

[24] *Mathworks*. [Online]. Available: https://www.mathworks.com/company/newsroom/mathworks-announces-release-r2019b-of-matlab-and-simulink.html (visited on 02/20/2021).

[25] M. Chen, G. Alregib, B.-H. Juang, "6dmg: A new 6d motion gesture database," Feb. 2012, pp. 83–88. DOI: 10.1145/2155555.2155569.

[26] C. O. S. Sorzano, J. Vargas, A. P. Montano, "A survey of dimensionality reduction techniques," *arXiv preprint arXiv:1403.2877*, 2014.

[27] G. T. Reddy, M. P. K. Reddy, K. Lakshmanna, R. Kaluri, D. S. Rajput, G. Srivastava, T. Baker, "Analysis of dimensionality reduction techniques on big data," *IEEE Access*, vol. 8, pp. 54 776–54 788, 2020. DOI: 10.1109/ACCESS.2020.2980942.

[28] I. Fodor, "A survey of dimension reduction techniques," Tech. Rep., 2002.

[29] M. Á. Carreira-perpiñán, "A review of dimension reduction techniques," Tech. Rep., 1997.

[30] J. Ye, R. Janardan, Q. Li, "Two-dimensional linear discriminant analysis," in *Advances in Neural Information Processing Systems*, L. Saul, Y. Weiss, L. Bottou, Eds., vol. 17, MIT Press, 2005. [Online]. Available: https://proceedings.neurips.cc/paper/2004/file/86ecfcbc1e9f1ae5ee2d71910877da36-Paper.pdf.

[31] S. Balakrishnama, A. Ganapathiraju, "Linear discriminant analysis - a brief tutorial," 1995.

[32] *Using linear discriminant analysis (lda) for data explore: Step by step*. [Online]. Available: https://www.apsl.net/blog/2017/07/18/using-linear-discriminant-analysis-lda-data-explore-step-step/ (visited on 11/20/2021).

[33] U. Sakarya, "Hyperspectral dimension reduction using global and local information based linear discriminant analysis.," *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 2, no. 7, 2014.

[34] V. N. Vapnik, *The nature of statistical learning theory*. Springer-Verlag New York, Inc., 1995.

[35] S. HUANG, N. CAI, P. P. PACHECO, S. NARRANDES, Y. WANG, W. XU, "Applications of support vector machine (svm) learning in cancer genomics," *Cancer Genomics & Proteomics*, vol. 15, no. 1, pp. 41–51, 2018, ISSN: 1109-6535. eprint: https://cgp.iiarjournals.org/content/15/1/41.full.pdf. [Online]. Available: https://cgp.iiarjournals.org/content/15/1/41.

[36] V. Jakkular, "Tutorial on support vector machine (svm)," School of EECS, Washington State University, Tech. Rep. [Online]. Available: https://course.ccs.neu.edu/cs5100f11/resources/jakkula.pdf.

[37] *Flightgear*. [Online]. Available: https://www.flightgear.org/ (visited on 04/12/2021).

[38] *Mfs*. [Online]. Available: https://en.wikipedia.org/wiki/Microsoft_Flight_Simulator (visited on 01/01/2022).

[39] *Prepar3d*. [Online]. Available: https://www.prepar3d.com/ (visited on 01/01/2022).

[40] *X-plane*. [Online]. Available: https://en.wikipedia.org/wiki/X-Plane_(simulator) (visited on 01/01/2022).

[41] P. Cao, X. Hu, G. Zhang, "Interface research and flight control based on flightgear," in *2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2017, pp. 397–402. DOI: 10.1109/ICIEA.2017.8282877.

[42] J. Ying, H. Luc, J. Dai, H. Pan, "Visual flight simulation system based on matlab/flightgear," in *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2017, pp. 2360–2363. DOI: 10.1109/IAEAC.2017.8054444.

[43] V. Kumar, H. Yong, D. Min, E. Choi, "Auto landing control for small scale unmanned helicopter with flight gear and hils," in *5th International Conference on Computer Sciences and Convergence Information Technology*, 2010, pp. 676–681. DOI: 10.1109/ICCIT.2010.5711140.

[44] *Python*. [Online]. Available: https://www.python.org/ (visited on 08/01/2022).

[45] J. Jordan, "The future of unmanned combat aerial vehicles: An analysis using the three horizons framework," *Futures*, vol. 134, p. 102 848, 2021, ISSN: 0016-3287. DOI: https://doi.org/10.1016/j.futures.2021.102848. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0016328721001579.

[46] M. Pavan, M. Pelillo, "Dominant sets and pairwise clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 167–172, 2007. DOI: 10.1109/TPAMI.2007.250608.

[47] M. Pavan, M. Pelillo, "A new graph-theoretic approach to clustering and segmentation," vol. Vol. 1, Jul. 2003, pp. I–145, ISBN: 0-7695-1900-8. DOI: 10.1109/CVPR.2003.1211348.

[48] M. Pavan, M. Pelillo, "Efficient out-of-sample extension of dominant-set clusters," in *Advances in Neural Information Processing Systems*, Citeseer, 2005, pp. 1057–1064.

[49] O. Haliloğlu, U. Sakarya, B. U. Töreyin, "Evaluation of clustering performance of hyperspectral bands," in *2015 23nd Signal Processing and Communications Applications Conference (SIU)*, 2015, pp. 2497–2500. DOI: 10.1109/SIU.2015.7130391.

A
GESTURES CODES

Table A.1 Gesture codes for experiments

Gesture Index	Gesture Name
1	SWIPE RIGHT
2	SWIPE LEFT
3	SWIPE UP
4	SWIPE DOWN
5	SWIPE UPRIGHT
6	SWIPE UPLEFT
7	SWIPE DOWNRIGHT
8	SWIPE DOWNLEFT
9	POKE RIGHT
10	POKE LEFT
11	POKE UP
12	POKE DOWN
13	V SHAPE
14	X SHAPE
15	CIR HOR CLK
16	CIR HOR CCLK
17	CIR VER CLK
18	CIR VER CCLK
19	TWIST CLK
20	TWITS CCLK

PUBLICATIONS FROM THE THESIS

Conference Papers

1. T. Zeybek and U. Sakarya, “Wavelet-based Spatial-temporal Feature Extraction for Gesture Recognition”, 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2021, pp. 1-5, doi: 10.1109/HORA52670.2021.9461268.