



HÖGSKOLAN
Dalarna

Identification of cause of impairment in spiral drawings, using non-stationary feature extraction approach

Muhammad Usman Yaseen

2012

**Master Thesis
Computer
Engineering
Nr:E4179D**



DEGREE PROJECT

Computer Engineering

Programme Masters Programme in Computer Engineering - Applied Artificial Intelligence	Reg number E4179D	Extent 15 ECTS
Name of student Muhammad Usman Yaseen	Year-Month-Day 2012-02-06	
Supervisor Mevludin Memedi	Examiner Mark Dougherty	
Company/Department Computer Science	Supervisor at the Company/Department Mevludin Memedi	
Title Identification of cause of impairment in spiral drawings, using non-stationary feature extraction approach		
Keywords Parkinson's Disease, Hilbert Huang Transform, Principle Component Analysis		

Abstract

Parkinson's disease is a clinical syndrome manifesting with slowness and instability. As it is a progressive disease with varying symptoms, repeated assessments are necessary to determine the outcome of treatment changes in the patient. In the recent past, a computer-based method was developed to rate impairment in spiral drawings. The downside of this method is that it cannot separate the bradykinetic and dyskinetic spiral drawings. This work intends to construct the computer method which can overcome this weakness by using the Hilbert-Huang Transform (HHT) of tangential velocity. The work is done under supervised learning, so a target class is used which is acquired from a neurologist using a web interface. After reducing the dimension of HHT features by using PCA, classification is performed. C4.5 classifier is used to perform the classification. Results of the classification are close to random guessing which shows that the computer method is unsuccessful in assessing the cause of drawing impairment in spirals when evaluated against human ratings. One promising reason is that there is no difference between the two classes of spiral drawings. Displaying patients self ratings along with the spirals in the web application is another possible reason for this, as the neurologist may have relied too much on this in his own ratings.

ACKNOWLEDGMENT

I would like to thank Almighty ALLAH, Who granted me enough strength and courage to do this thesis work. I am thankful to my supervisor Mr. Mevludin Memedi, whose advice, help and supervision was invaluable. This is because of his guidance and long-term, praiseworthy support that I have now completed the thesis work.

I am deeply thankful to Mr. Dag Nyholm who played an important role in my thesis work for rating the spiral drawings.

Also I would like to thank the entire team of Artificial Intelligence program for their support and assistance during my study in Dalarna University.

At the end I am thankful to my dear family members especially my mother and father for their kind support and courage they provided me during my stay in Sweden.

Muhammad Usman Yaseen

TABLE OF CONTENTS

1. INTRODUCTION	1
1.1- Parkinson's disease.....	7
1.2- Symptoms and stages.....	8
1.2.1- Motor Symptoms.....	8
1.2.2- Non-Motor Symptoms.....	9
1.2.3- Stages.....	10
1.3- Treatment.....	11
1.3.1- Drug Treatment.....	11
1.3.2- Surgical Treatment.....	11
1.3.3- Therapy Treatment.....	11
1.4- Problem Description	12
1.5- Objectives and Work Flow.....	13
2. THEORETICAL BACKGROUND	13
2.1- Related work and Ongoing Research.....	13
2.2- Spiral Drawings.....	13
2.3- Kinematic Measures.....	14
2.4- Hilbert-Huang Transform.....	14
2.5- Principal Component Analysis.....	15
2.6- Classification Method.....	16
2.7- Specificity and Sensitivity.....	17
2.8- ROC curve.....	17
3. MATERIALS AND METHODS	18
3.1- Data Processing.....	18
3.1.1 Radial Velocity.....	19
3.1.2 Angular Velocity.....	19
3.1.3 Tangential Velocity.....	19
3.2- Web Application.....	20
3.2.1 LoginScreen.....	20
3.2.2 Rating Page.....	20
3.2.3 Start Rating.....	21
3.2.4 Rating History.....	22
3.3- Hilbert-Huang Transform.....	22
3.3.1- Theoretical background of HHT.....	22
3.3.2- The emperical mode decomposition.....	23
3.3.3- IMF Selection.....	24
3.3.4- Hilbert Transform	25
3.4- Dimension Reduction.....	26
3.5- Classification.....	27
4. RESULTS AND DISCUSSION	28
4.1- Naive Bayes.....	28
4.2- C4.5	39
4.3- Sensitivity and Specificity	30
4.4- ROC curve	31
4.5- Discussion.....	32
CONCLUSIONS	33
REFERENCES	34

LIST OF FIGURES

Figure.3.2.1 user login.....	10
Figure.3.2.2. wellcome page.....	10
Figure.3.2.3.Rate spirals.....	11
Figure.3.2.4 History (a).....	11
Figure.3.2.4 History (b).....	12
Figure.3.3.1 Velocity vs Time.....	15
Figure 3.3.2 IMFs.....	19
Figure 4.4 ROC curve.....	25

LIST OF TABLES

Table 1.1- Stages of Parkinson’s disease	4
Table 3.1- Comparison of HHT and traditional methods.....	17
Table 4.1- Performance Vector1.....	29
Table 4.2- Performance Vector2.....	30
Table 4.3- Performance Measures.....	31

1- INTRODUCTION

1.1- Computer Technology in medicine

The thought of using computer science in medicine is growing rapidly with the passage of time. In 1950s computers were used mostly for dental projects. Later on different programming languages were used and a number of applications were written for different clinical applications. One of the major contributions of the computer technology in the medical field is that it is providing the services to facilitate the patient at home. If a patient is suffering from a progressive disease, which requires continues attention of a medical expert then he should be present in the hospital which can cause a lot of expense. Computer technology can prove to be very helpful in such a situation. It can provide different ways to the patient to interact with the medical experts by staying at home. This will reduce not only the cost but will also be helpful to reduce the diagnostic faults.

Ongoing research [1] and a number of studies related to this area reveal that the computer based systems for healthcare can not only improve the overall competence but will also decrease the errors which can occur while treating the disease. This can be achieved by using the recent advances in numerous fields including wireless communication and data mining. Advances in web technology can also be used to observe and monitor the patients distantly. Especially collection of patient's data and sending it to a central station can be made easy through this technology.

1.1.1- IT-based Methods

A number of IT- based methods are in use today. These methods are based on the fusion of statistical pattern recognition techniques and machine learning methods. These methods are supportive in predicting the therapeutic outcome of the patients. A systematic way is followed to use the clinical information and then this prediction is obtained. Experts can use this information in order to ease their decision making process.

The selection of the method is entirely dependent on the required outcome. Two types of outcomes are usually desired, one is numeric and the other is nominal outcome. For numeric outcome, numeric prediction methods can be useful. And for the nominal outcome, classification methods can be used. The performance of both the methods can be measured by observing its accuracy and errors. Also, it is very important to evaluate these methods for their metrics such as reliability and validity.

1.2- Parkinson's disease

The research on Parkinson's disease is increasing from the last few years. According to a study [2] Parkinson's disease (PD) is the second most common neurodegenerative disorder. The major portion of the brain that is affected by this disease is the substantia nigra. This portion of brain holds a dedicated set of neurons that transmits signals in the form of a neurotransmitter called dopamine. The neurotransmitters pass through to the striatum with the help of extended fibers known as axons. The movement of this pathway manages the normal activities or actions of the body. The problem occurs when the neurons in the substantia nigra are reduced. This results in the loss of dopamine. Due to the loss of dopamine, nerve cells of the striatum fire markedly. Because of this phenomenon it becomes impossible for people to control their movements, leading to the primary motor symptoms of PD.

With the passage of time the interference of this disease with the motion of the body increases and it becomes more devastating and complex. As the complexity of disease increases, many other problems and diseases hit the patient which can result in to serious consequences. The other diseases and problems can be with swallowing and chewing, speech problems, excessive sweating and difficulties with sleep etc. However the functional activities that are directly affected by the Parkinson's includes balance, speech, fastening buttons, handwriting, walking typing, driving, and many other activities that are directly controlled by dopamine and the basal ganglia.

1.2.1- Dyskinesia

Dyskinesia can be defined as abrupt, unmanageable and disordered movements of different parts of the body. The causes of dyskinesia include extent of nigral cell loss, combined interplay of impairment of storage machinery, and damage from modality of drug administration. The patient may suffer from grimacing, grinding of teeth and rapid blinking as these are the main symptoms of dyskinesia. In order to manage dyskinesia, occasionally a decrease in levodopa and Deep brain stimulation (DBS) is used.

1.2.2- Bradykinesia

Bradykinesia refers to tardiness of movement and is the most distinctive clinical attribute of Parkinson's disease. The tardiness of movement becomes more clear when starting and executing actions. The causes of bradykinesia include multiple systems atrophy, progressive supranuclear palsy and medication e.g. anti-psychotic and anti-seizure drugs. Bradykinesia develops steadily over time and is not usually visible at the starting stages of Parkinson's disease. There is tardiness in starting or repeating movements and trouble with rapid fine

movements. Sometimes the patient may solidify in mid stride with incapacity to proceed further. Executing simple routine tasks such as dressing or eating also turns out to be a problem as a result of bradykinesia. Levodopa is tremendously effectual in recovering bradykinesia in Parkinson's disease. Amantadine is useful in the early stages. Surgery can be helpful in some cases. Physiotherapy also plays a definite role in the treatment.

1.2.3- Treatment

A patient having Parkinson's disease can be treated by different ways. However most of the studies [3][2] have classified the treatment of this disease into three categories, i.e. Drug treatment, surgical treatment and therapy treatment. The purpose of the drug treatment is to raise the level of dopamine directly or indirectly in the brain. It is not possible for the patients to simply take dopamine pills. The reason is that the dopamine does not go through the blood vessels into the brain easily. Usually the medicines which are used for PD patients are dopamine precursors – substances such as levodopa that cross the blood-brain barrier and are then changed into dopamine. Some more drugs which are used for PD patients affect other neurotransmitters present in the body in order to relieve some of the symptoms of the disease. An example is anticholinergic drugs. These drugs decrease the movement of the neurotransmitter acetylcholine. Such drugs are of great importance if tremors and muscle stiffness is to reduce.

The second treatment for PD patients is surgical treatment. Surgical treatment mainly includes brain surgery in which surgeons removes those parts of the brain that were “misfiring.” In this way most of the symptoms of PD patients can be lightened. Now-a-days clinicians have put much effort to improve these techniques and because of these efforts, techniques are much safer now, but the problem of irreversible is still there. There is a number of harmonizing and encouraging therapies which can be used for the treatment of PD. Exercise can also help to resist against disease progression, it can also help patients to increase strength and improvement in mobility. Muscles can also be strengthened by physical therapy and muscle-strengthening exercises.

1.3- Motivation

It is often difficult to determine the treatment outcome of patients suffering from advanced Parkinson's disease which is explained briefly in the previous chapter. In clinical practice, doctors often use the historical information of the patients to assess motor disability. But using video recordings for this task is a very time consuming procedure. To save time Self-assessments in diaries can be used but it also has a problem that sometime paper patient

diaries are not filed in at the requested time-point (Stone et al., 2003). This can lead to doubts in the correctness of the reported symptomatology. To remove this drawback electronic diary can be used where the information is demanded at definite time-points and the e-diary notices when the information was provided by the patient[4]. In 2005 liu et al [5] used the spiral drawings for the assessment of unconscious movements of the patients. Even in 2008 and 2009 Saunders-Pullman et al. [6] and Banaszekiewicz et al. [7] assessed bradykinesia and dyskinesia from spiral drawings by using different calculated spiral “indices” and spiral drawing completion time.

In [8] researchers have used a PDA with touch screen to gather symptoms from almost 65 patients. This is done by tracing pre-drawn spirals on the screen. Each spiral drawn by the patient has almost 200 x and y pixel coordinates, plus the time. In order to rate the spiral drawings impairments a computer-based method was developed by [9]. This method had a drawback that it lacks the means to differentiate between bradykinetic spirals (possibly caused by low medicine levels) and dyskinetic spirals (possibly caused by high medicine levels). In this work I have focused in finding such a method.

1.4- Objectives

The above mentioned spiral drawing method had a serious drawback that It only identify disability but it is not able to differentiate between bradykinesia and dyskinesia. The objective of this thesis is to work on resolving this issue. So a novel computer method for the Classification of cause of impairment in spiral drawings will be developed. This computer method will aim to classify bradykinesia(possibly caused by low medicine levels) and dyskinesia(possibly caused by high medicine levels) in to two distinct classes. A dataset that will be used is gathered from a series of tests consisting of self assessment and motor tests. This includes spiral drawings obtained through a mobile device test battery. To achieve the features from the spirals Hilbert Huang Transform will be used. As the dimension of the features is too high therefore Principle Component Analysis can be applied to reduce the dimension. In order to get the right direction in the feature space PCA has been applied in a specific way. To apply the PCA, firstly a subset of features on basis of rated spirals will be selected and then PCA will be applied on this subset. Secondly, coefficients of the principal components will be used in order to calculate scores for all the dataset. The training and testing set will be separated using cross validation (10 folds). As the work will be done in context of the supervised learning paradigm therefore a target class or desired output will be needed. This desired output class has been taken from an expert neurologist by using a web

interface. A predictive model will be used to find relationships between extracted features and the manual ratings.

1.5- Thesis Outline

Organization of the thesis is as follows. In the next section, a theoretical background is given, in which a description of all the techniques which are used is provided. In the methodology section, a detailed explanation of the method which is developed is specified. The obtained results are summarized in the results section followed by the discussion part. Finally, the report ends up with the conclusion section.

Chapter 2

Theoretical Background

In this section, a brief theoretical overview of the concepts is being presented. Firstly, the work that has been done so far in this field is reviewed, and then the basic concept of different techniques that has been used in this work is thrashed out.

2.1- Related work and Ongoing Research

A comprehensive work has already been done before in the domain of Parkinson's disease. In order to assess the effects of treatment in patients repeated clinical ratings of motor function on a response scale [10] is used. Also home diaries, both on paper [11] and hand computers [10, 12, 13], have been used by clinical experts. As described earlier that the use of paper based diaries was not good [14] therefore e-diaries were used [12]. Spiral drawings on tablet PC touch screens have been used by the researchers for quantification of unconscious movements such as dyskinesias [15] and tremor [16, 17]. In [8] a test battery containing the self assessment and motor tests for the assessment of patients is introduced. For this purpose a PDA with touch screen has been used to gather symptoms from 65 patients.

The test battery contained disease associated questions as well as motor test. Motor test includes tapping test and spiral drawing test. The reason behind using the spiral drawing task was used to judge the drawing impairment originated because of involuntary movements, like tremor, dyskinesias and bradykinesia. The patient had to track a pre-drawn spiral on a touch screen. Among these spirals some hundred spirals have been rated on clinical scale for drawing impairment and associated cause by two clinical experts [8].

2.2- Spiral Drawings

A spiral is a curve in the plane or space which originates from a central point, and by revolving around the centre point it progressively moves farther away in a special way. The most common spiral is the archimedean spiral which comes under the category of 2D spirals. It has a constant separation distance at successive turnings. In this work spiral drawings are used to assess the state of a patient with advance parkinson's disease. The assessment is based on the drawing impairment caused by instinctive movements, such as dyskinesias and bradykinesia.

The spiral drawing data consists of around 256 x and y coordinates position and time. The x and y coordinates constitutes a spiral and there are almost 400 spirals of different patients. The coordinates position and time for each spiral was not fixed in the given data. Some spirals

had less than 256 data points while some had more than 256, so the technique of zero padding is applied in order to keep them fixed for all spirals. It helped make the data processing more comprehensible and simple.

2.3- Kinematic Measures

Kinematics can be used if there is a need to deal with the objects in motion. And it is very important to specify the position to describe the objects in motion. To denote the position of a body variable x and y are usually used [18]. These two variables are required if the object lies on a plane otherwise in case of three dimensional splendour a third variable z can be introduced. In my study also the data set of spirals is consisted of x and y coordinated which are used to describe the position of spiral. These position describing variables has dimension of length therefore it is necessary to identify the coordinate system prior to using them.

Three things are used to define a rectangular coordinate system: the origin, direction of axes and the distance scale. In one dimensional system, a single variable can denote a spatial coordinate but two variables x and y are needed to denote spatial coordinates in two dimensional system. Both are perpendicular to each other. On the other side an alternative to the rectangular coordinate system is polar coordinate system. In this system instead of using x and y , a point is represented by r and θ . r represents the length of a straight line from the origin to the point and θ represents the angle which the line makes with the x -axis.

2.4- Hilbert-Huang Transform

In non-stationary processes variations in the signals occurs significantly. To describe non-linear and non-stationary distorted waves in detail, Hilbert-Huang transform can be used. The Hilbert-Huang transform is a pretty new method for data analysis and is dissimilar from the traditional data analysis methods. The traditional data analysis methods are typically used for stationary and linear processes while HHT can be used for non stationary and non linear processes as well [22].

In most real life situations either natural or even man-made ones, the data can be both non-linear or non-stationary in nature and traditional data analysis methods like wavelet transform cannot handle them. The reason is that in wavelet transform the selection of wavelet base (or mother wavelet) is a critical problem. Also if different mother wavelets are used to analyze same problem, it will produce different results. Hence the results taken from wavelet analysis are inadequate by mother wavelet, and wavelet components are significant only to the chosen mother wavelet. Also an essential requirement to characterize non-linear and non-stationary data is to have an adaptive basis, and an apriori defined function (no matter how sophisticated

it is) cannot be relied on as a basis. Hence HHT [23]–[25] is the way to solution for such problems because this technique uses a posteriori-defined basis.

2.5- Principal Component Analysis

Principal component analysis is a statistical technique and is commonly used in modern data analysis and compression. Due to its simplicity and efficient dimension reduction quality, it is being used now-a-days in different fields of science. It can extract relevant information from a huge data set nicely and provides an approach to identify patterns in dataset. This identification of patterns can help to squeeze the data by reducing the number of dimensions, and interestingly this reduction in dimension is done without much loss of information.

PCA converts the dataset into a set of principal components by eliminating the redundant data from it. These principal components can be equal to or less than the original variables. The procedure is performed in such a manner that the first principal component contributes much in the dataset, i.e. it has as high a variance as feasible. High variance means that it accounts for as much of the inconsistency in the data as possible. Similarly the second principal component has the second most variance (important information), as compared to the previous one. All the remaining principal components exhibit decreasing variance and are uncorrelated with all other principal components. [27]

2.6- C4.5 Classifier

Weka comes up with a variety of classification tools including some tree algorithms. Most of the tree algorithms work recursively. C4.5 also comes in the category of algorithms that work recursively. It starts by selecting an attribute as the parent node or root node and then splits the set of samples into subsets. To generate an efficient decision tree, it is necessary to make sure that the selected root node splits the data effectively. For this reason information gain is used. The node with the highest information gain is selected as the root node.

When the classifier comes across a training set it focuses on the attribute that differentiates the variety of instances most clearly. This attribute which is capable to inform us, most about the data instances will have the highest information gain. The split is made by taking this attribute with highest information gain as root node. Now with the values of this attribute, if there exists any value for which the data instances lying inside its category have the same value for the target variable, then branch is ceased and the target value is assigned to it. For the remaining cases, another attribute is selected on the same previous criteria of the highest information content. The same procedure is repeated again until a decision is made about the combination of attributes that gives particular target value. If there are no more attributes left

then a target value is assigned to this branch that the majority of the items under this branch possess [28].

2.7- Naive Bayes Classifier

Naive Bayes classifier is a plain probabilistic classifier. It is based on Bayesian theorem and assumes attribute independence. Attribute independence means that it assumes that all the predictive attribute are conditionally independent given the class. The ultimate objective of the classifier is to predict the output class of test instances, and this has to be done as accurately as possible. Naive Bayes classifier stands on two basic assumptions: firstly, which is also explained earlier that the predictive attributes are conditionally independent given the class. And the second is that, the values of numeric attributes are normally distributed within each class.

The function of Naive Bayes is different for discrete and continuous attributes. In case of discrete attribute, the probability that the attribute X will take on the particular x when the class is c is modeled by a single real number between 0 and 1 []. On the other hand, every continuous attribute is structured by some continuous probability distribution over a range of that attribute's values.

Bayesian theorem is used in Naive Bayes which computes the probability by counting the frequency of values and combinations of values in the past data. Bayesian Theorem calculates the probability of an occurrence given the probability of another occurrence which has been occurred before. If B corresponds to the dependent occurrence and A corresponds to the preceding occurrence, Bayes' theorem can be given as;

$$\text{Prob}(B \text{ given } A) = \text{Prob}(A \text{ and } B) / \text{Prob}(A)$$

This equation means that in order to find out the probability of B given A, the algorithm counts the number of places where A and B took place jointly and divides it by the number of places where A happens to be alone. Naive Bayes classifier necessitates very less training data to approximate the parameters required for classification. This is one of the advantages of Naïve Bayes. Another advantage is that it can be used for both binary and multi class classification problems. [29]

2.9- Specificity and Sensitivity

These are the two statistical measures which are used for the performance of a classification test. In statistics, it is also known as classification function. Sensitivity which is also termed as recall rate measures the proportion of actual positives which are correctly identified as such (e.g. the percentage of sick people who are correctly identified as having the

condition). On the other hand, Specificity measures the proportion of negatives which are correctly identified (e.g. the percentage of healthy people who are correctly identified as not having the condition). In theoretical terms, one can say that the optimal prediction intend to attain 100% sensitivity (i.e. predict all people from the sick group as sick) and 100% specificity (i.e. not predict anyone from the healthy group as sick).

2.10- ROC curve

The receiver operating characteristic (ROC) curve, is a graph between the sensitivity, or true positive rate, and false positive rate ($1 - \text{specificity}$), for a binary classifier system. In a binary classifier system, there are only two output classes. The ROC curve can also be drawn homogeneously by plotting the fraction of true positives out of the positives ($\text{TPR} = \text{true positive rate}$) vs. the fraction of false positives out of the negatives ($\text{FPR} = \text{false positive rate}$).

A ROC space is defined by False Positive Rate and True Positive Rate as x and y axes respectively. As True Positive Rate is the same as Sensitivity and False Positive Rate is the same as $1 - \text{specificity}$, the ROC graph is also termed as the sensitivity vs. ($1 - \text{specificity}$) plot. Each prediction result or one instance of a confusion matrix symbolizes one point in the ROC space.

If there is a method that can give most excellent possible prediction then it would yield a point in the upper left corner of the ROC space i.e. coordinate $(0,1)$, demonstrating 100% sensitivity (no false negatives) and 100% specificity (no false positives). This point at upper left corner is also known as a perfect classification. An absolutely random guess would yield a point towards the diagonal line (the so-called line of no-discrimination) from the left bottom to the top right corners. This diagonal line divides the ROC space. The area above the diagonal shows good classification results while the area below the line depicts poor results.

Chapter 3

Materials and Methods

In this section, the overall methodology of the system has been explained step by step.

3.1- Data Processing

As described earlier, the dataset which has been used in this study consists of 400 spirals. It was necessary to use some kinematic measures of each spiral so that signal processing technique could be applied on it. The reasonable thing was to use velocity as a kinematic measure as the x and y coordinates along with the time were available. Hence using [19] rectangular coordinates are converted to polar coordinates. The relationship between the rectangular coordinates and polar coordinates can be seen by the following equations:

$$\text{i) } x = r \cos \theta$$

$$\text{ii) } y = r \sin \theta$$

$$\text{iii) } r^2 = x^2 + y^2$$

$$\text{iv) } \tan \theta = y/x, x \neq 0$$

3.1.1 - Radial velocity

Radial velocity can be defined as the component of velocity away from or toward the origin.

Mathematically it can be expressed as :

$$v_r = \Delta r / \Delta t$$

where

v_r = radial velocity

r = radius

t = time (s)

3.1.2- Angular velocity

The angular velocity is a vector quantity which can be defined as the rate of change of angular displacement with respect to time. It indicates the angular speed of a body and the axis about which the body is revolving and is measured in radians per second. Mathematically it can be expressed as [20].

$$\omega = \Delta \theta / \Delta t$$

where

ω = angular velocity (rad/s)

θ = angular displacement (rad)

t = time (s)

3.1.3- Tangential velocity

When a body revolves around in a circle, the linear velocity at a point r meters away from the center of the circle is the tangential velocity [21].

Following equation is used to solve for tangential velocity. This equation describes that tangential velocity is the product of radius of circle and angular velocity.

Where

ω = angular velocity

r = radius of circle

By using the above equations, I have calculated the tangential velocity for every coordinate position with respect to time.

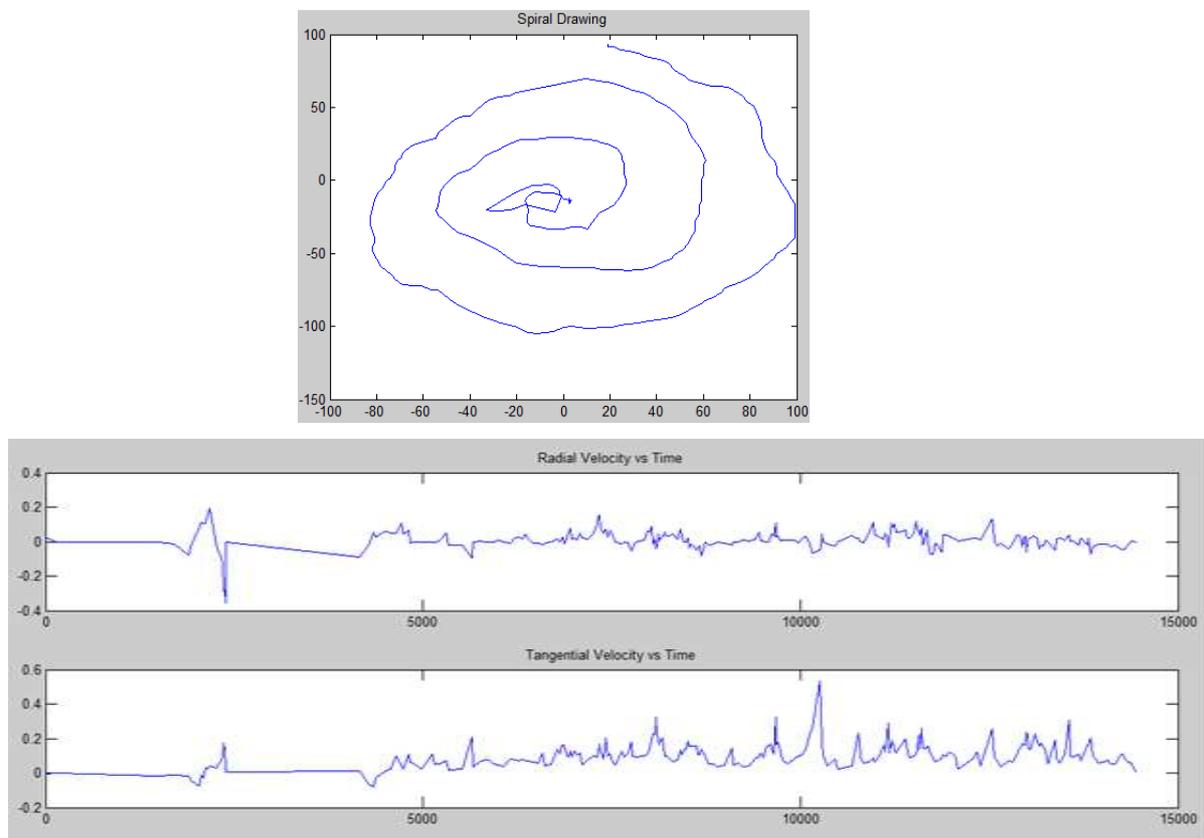


Figure 1: Velocity vs Time graph

3.1.4- Schematic Diagram

A schematic diagram of the overall procedure is shown below; further explanation of each module is described in the following sections.

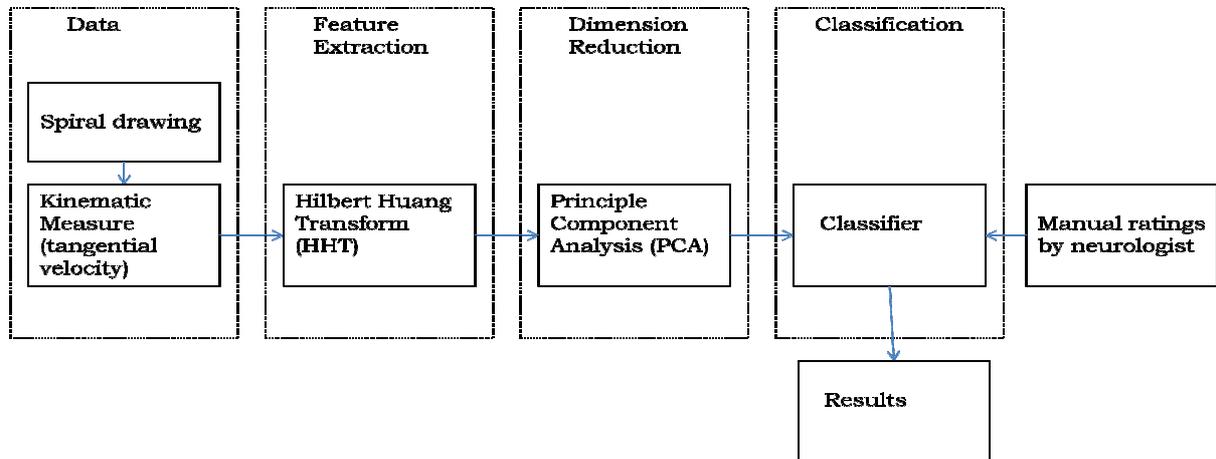


Figure 2: HHT is used to extract the features from spiral drawings. After reducing the dimension of features, classification is performed.

3.2- Web Application

In order to show the spiral drawings from the spiral test and to permit users (i.e. Parkinson's disease specialists) to rate spiral drawing mutilation, a web interface has been used. The rated spirals can then be used as the target class of the classifier. This web interface is completely user friendly and no technical skills are required for it. It has been designed by using PHP and MYSQL. MYSQL served the purpose of back-end database and php is used for web services. To display spiral drawings it retrieves paired x and y coordinates from the database that contains the spiral information. The GD extension drawing functions are used to produce an image and to sketch the retrieved pixels from database on the display. PHP has the ability to draw a variety of geometric shapes, including lines, rectangles and polygons etc. Some other information such as drawing completion time and patient's self-assessment of motor state at the time of the particular test occasion were also retrieved and displayed on the web page. Following are few screen shots and their detail which explains the web interface and its usage comprehensively.

3.2.1- Start Rating

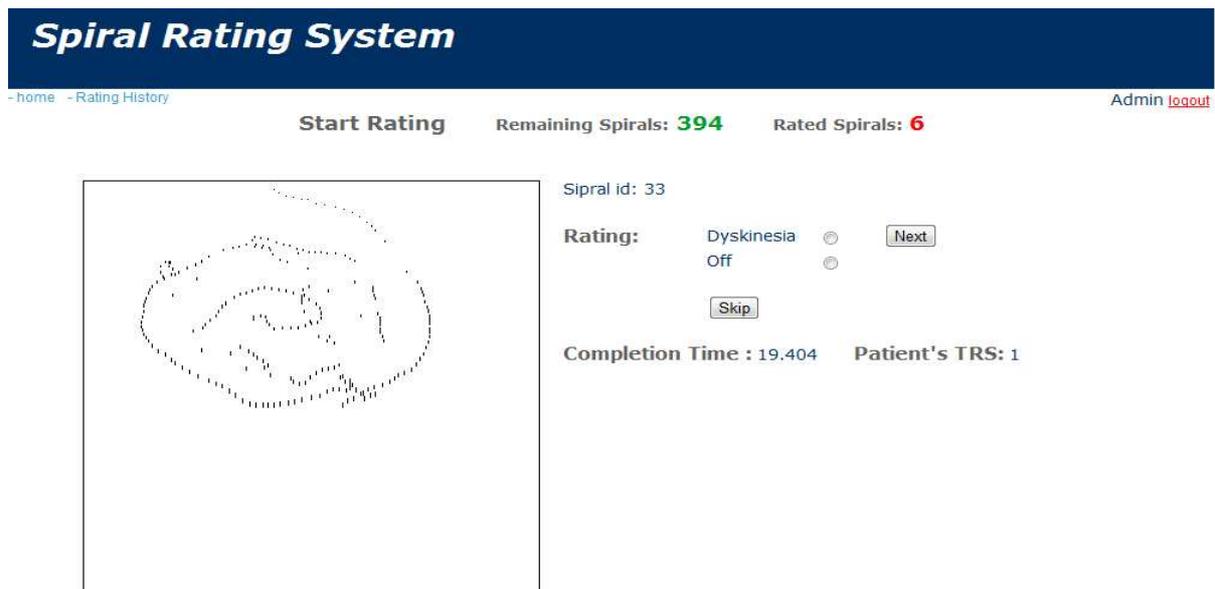


Figure 2: Interface for rating spirals; each spiral can be rated either as dyskenetic or Off

The start rating page enables the user to rate the spirals. This page displays the spirals randomly one at a time. A spiral can be rated as Dyskinesia or Bradykinesia. Also if a user is not confirmed about the class of the spiral, he can skip that one. The spiral which is skipped can appear again later on. Once a spiral is rated it will not appear in the start rating section again. This helps the user to rate all the spirals successfully.

3.2.2- Rating History



Figure 3: showing the history of rated spirals

As the name shows rating history page shows the history of a user. It can be used to see the ids as well as the completion time of all the rated spirals. Also it provides the option to shift a rated spiral from one class to another. By using the shift button a spiral can be transferred to other class as shown below:

Spiral Rating System

[home](#)

[logout](#)

Rating History

Dyskinesia (10)					<input type="button" value="Show"/>	Bradykinesia (11)				
QuestionID	TestDataID	Time	AnsVal	Shift	QuestionID	TestDataID	Time	AnsVal	shift	
14	2611	19.106	3	shift	15	870	10.98	2	shift	
14	2726	31.793	1	shift	14	1364	8.446	-2	shift	
15	2835	15.226	1	shift	15	1769	22.09	-1	shift	
14	3977	5.134	0	shift	13	2004	13.808	-1	shift	
14	5129	4.479	1	shift	13	2290	32.914	-1	shift	
15	5715	8.527	1	shift	14	4300	10.244	0	shift	
13	5786	10.697	1	shift	14	5064	7.195	1	shift	
13	8142	15.221	-3	shift	14	5367	5.339	-1	shift	
14	8534	35.444	1	shift	14	5539	18.493	1	shift	
13	9042	12.693	-2	shift	13	9173	29.126	1	shift	
					15	9769	14.851	1	shift	

copyright @ Usman

Figure 4: Spirals History; one spiral is shifted from dyskinesia to off

3.3- Hilbert Huang Transform

The Hilbert-Huang transform is quite a new method for data analysis and is different from the state of the art data analysis methods. The traditional data analysis methods are typically used for stationary and linear processes while HHT can be used for non stationary and non linear processes as well. A comprehensive detail of this technique is mentioned below:

3.3.1- Theoretical Background of HHT

In order to mine appropriate information from data, different Fourier-based methods had been used by many researchers. These methods employ stationary sines and cosines as basis functions to decompose data. But as described earlier that in real life situations data come from natural phenomena. In natural phenomenon there is usually no idea that when a natural process will begin and when it will end. Because of which the ensuing signal will be non-stationary and there will be no information on whether some standard wavelengths are going to repeat or not. Hence the procedure to analyze data from such processes should be adapted from the data itself devoid of using a priori basis sets. In simple words we can say that the data should make known its own solutions, rather to impose the solutions on the data.

An entirely adaptive method that has been developed by Huang et al. (1998) is the HHT (Hilbert–Huang transform). It is different from the traditional state of the art methods as in this method the basis sets are derived from the intrinsic time-scales of the data through a sifting process and is completely adaptive. The table below shows a comparison of HHT and traditional methods.

	Fourier	Wavelet	Hilbert
Basis	a priori	a priori	adaptive
Nonlinear	no	no	yes
Nonstationary	no	yes	yes
Feature Extraction	No	discrete: no continuous: yes	yes

TABLE 1: Comparison of HHT and traditional methods

In HHT the basis sets are plagiaristic from the intrinsic time-scales of the data through a sifting process known as empirical mode decomposition (EMD). The basis achieved from this method is termed as intrinsic mode functions (IMFs) and they form a complete set. By adding all the IMFs present in the set, original signal can be obtained.

3.3.2- The empirical mode decomposition method (the sifting process)

The empirical mode decomposition method is a spontaneous, straight, and adaptive process which is designed to deal with data obtained from non-stationary and non-linear processes. It is an intuitive method with an a posteriori-defined basis derived from the data, It employs a decomposition process which is based on the straightforward postulation that every data consists of different simple intrinsic modes of oscillations. The decompositions process proceeds as follows:

First of all, the local maxima and local minima of the signal are calculated. Then all the local maxima are connected to form an upper envelop, similarly all the local minima are connected together to form a lower envelop. The upper and lower envelops contains all the data within themselves. After producing the upper and lower envelops, there mean is calculated. If the mean of upper and lower envelop is represented by m_1 , then the first component will be the difference between the data x_1 and mean m_1 .

Mathematically it can be written as;

$$h_1 = x(t) - m_1$$

$$m_1 = L+U/2$$

where L and U are the local minima and local maxima respectively.

If no error has been introduced during the process then h_1 can be considered as the first IMF, but generally error occurs and because of which the sifting process has to be repeated many times. There are two main functions performed by this process; one is the removal of small waves that gives the impression to ride big waves and the second is that it helps the signal to become more symmetric about the local zero mean line.

During the second phase of the sifting process, h_1 which was calculated in the previous phase is treated as the data and a new mean is calculated by the same way as it was done before. If we represent the new mean with m_{11} , then mathematically it can be given as;

$$h_{11} = h_1 - m_{11}$$

The same process is repeated up to 'k'times and h_{1k} turns out to be an IMF. It can be expressed as;

$$h_{1k} = h_{1(k-1)} - m_{1k}$$

Now let $h_{1k} = c_1$, where c_1 consists of shortest period component of data. If c_1 is subtracted from the original data then the resulting data will contain the longer period component only.

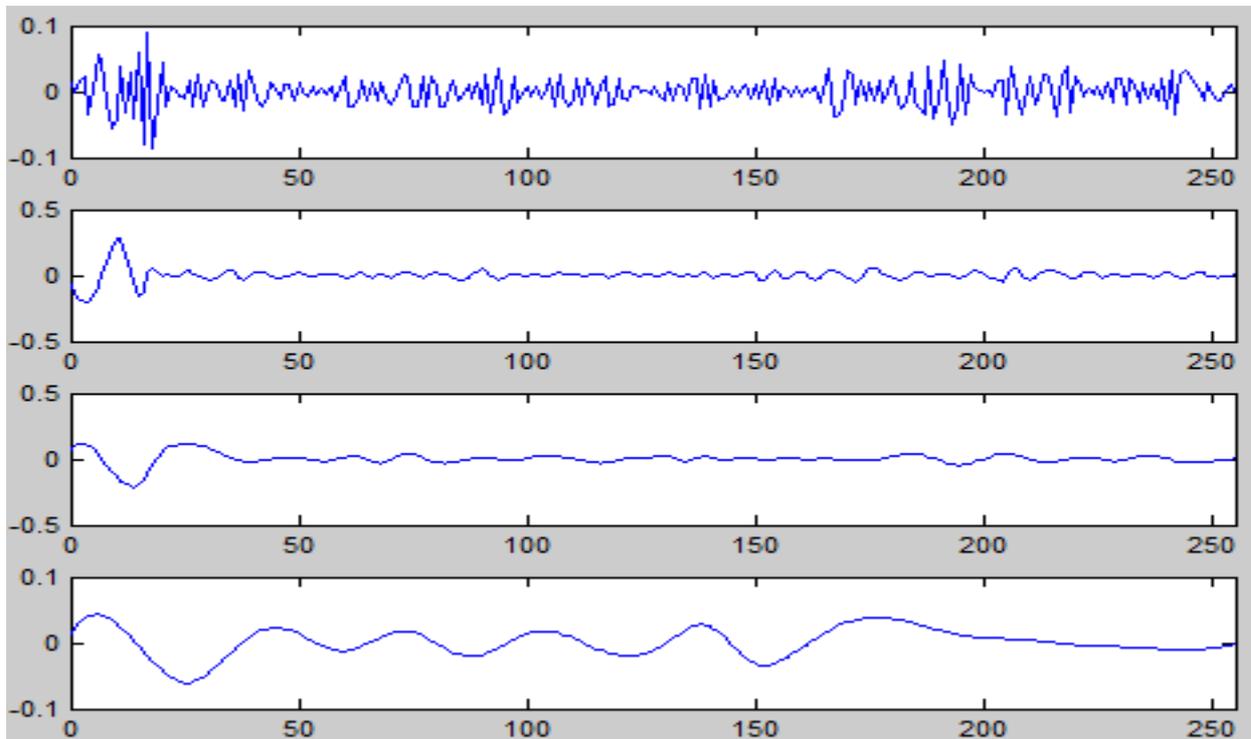
$$x(t) - c_1 = r_1$$

r_1 is the residue and is treated as the new data. It contains the long period components only. This residue then undergoes to the same sifting process . The process is repeated for all the subsequent r_j 's as shown below:

$$r_1 - c_2 = r_2 \dots$$

$$r_{n-1} - c_n = r_n$$

Also there is a stopping criteria to stop the sifting process because if the process is allowed to continue further it will remove the important signal variations and features. Huang et al proposed a stopping criteria to limit the value of the sum of the difference (SD), calculated from two successive sifting results. The value of SD between 0.2 and 0.3 is typically preferred based on experimental results presented by Huang et al. (1998).



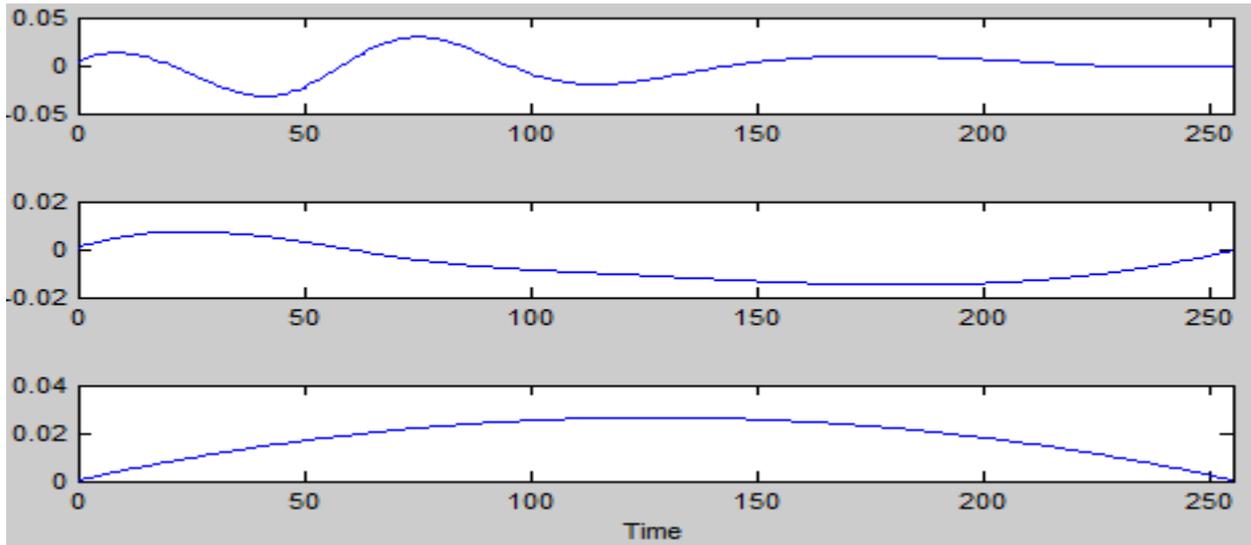


Figure 5: A decomposed spiral; each graph represents an IMF

3.3.3- IMF SELECTION

Once the IMFs have been achieved, it is possible to draw necessary information from them. But before that it is also necessary to get rid from those IMFs which are not relevant to the decomposition process or irrelevant. Irrelevant IMFs can be the result of numerical errors as EMD is a numerical procedure. It means that there is a need to discriminate between the relevant IMFs and irrelevant IMFs as only relevant IMFs carry the essential knowledge required to analyze the underlying system.

In this thesis work a stringent threshold has been used that was proposed by ALBERT et al. 2010 [26] to discriminate between relevant and irrelevant IMFs. The authors used a method in which they find out the correlation coefficients between each IMF and the original signal. On the basis of these correlation coefficients they determined more rigorous threshold coefficients. This threshold coefficient can be used to distinguish between related and unrelated IMFs, particularly taking into consideration the signals that have noise. After testing a number of assessment signals, a threshold is carried out as a function of maximum correlation coefficient. The IMFs having the correlation coefficients greater than the prearranged threshold are retained. The expression of the threshold is given as;

$$\mu_{TH} = \max(\mu_i) / 10 \times \max(\mu_i) - 3, \quad i = 1, 2, \dots, n$$

Here;

μ_{TH} represents the threshold; μ_i represents correlation coefficient of the i th IMF with the original signal,

n is for total number of IMFs and $\max(\mu_i)$ is the maximum correlation coefficient observed.

This threshold has eliminated almost all the spurious IMFs which were generated during EMD process. It has been observed by using the proposed equation that long-period, low-

frequency components that put a part to create the last few IMFs are mainly left out as irrelevant. It shows that the IMFs that were left out as irrelevant did not come from any of the original components used to create the original signal. Instead they were due to the numerical errors produced during the EMD process.

3.3.4- HILBERT TRANSFORM

“Hilbert–Huang transform” (HHT) is the combination of empirical mode decomposition and the Hilbert transform. For a continuous signal $x(t)$, the Hilbert transform can be given as [24];

$$Z(t) = X(t) + iY(t)$$

I have used the built-in function of matlab to implement the hilbert transform. $Z(t)$ can be expressed as;

$$\begin{aligned} \text{Where} \quad Z(t) &= a(t)e^{i\theta(t)} \\ \text{and} \quad a(t) &= \sqrt{x^2(t) + y^2(t)} \\ \theta(t) &= \arctan Y(t)/X(t) \end{aligned}$$

Now, the instantaneous frequency can be defined as;

$$\omega(t) = d \theta(t)/dt$$

3.4- Dimension reduction via PCA

PCA converts the dataset into a set of principal components by removing the redundant data from it. These principal components can be equal to or less than the original variables. The procedure is performed in such a manner that the first principal component contributes much in the dataset, i.e. it has as high a variance as feasible.

PCA ends up by returning a square coefficient matrix. The size of this coefficient matrix belongs to the size of feature vectors. Each column of this coefficient matrix contains coefficients for one principal component. The data is represented in terms of eigenvectors and eigen values. The eigenvectors are sorted by eigen values from the highest to the lowest. This is done to get the components in the order of contribution to variance. Now the first principal component represents the direction in feature space corresponding to maximum variance. In this thesis work, PCA was first applied on a subset of spiral data i.e. 60 rated spirals. These 60 rated spirals were assumed to be accurately classified by the neurologist. After applying PCA on these 60 selected spirals a matrix of coefficients is obtained. Then the coefficients of the principal components have been used in order to calculate scores for all the dataset. This calculated score along with the spiral ratings is then used for classification.

3.5- Classification

In order to test the validity of the method and to perform the classification task, an open source tool known as Rapid-Miner has been used. It comes up with different classification and regressions algorithms which can be used for different purposes according to the need of the problem. Here it is used to classify the data consisting of a set of features obtained using HHT along with the target output. The set of features are first normalized using z-score normalization and then it is used for the classification. The results of the classifiers are discussed in the following section.

Chapter 4

Results and Discussion

In this section of report, the results obtained after applying classifiers on the data generated from Hilbert Huang Transform are presented and then a comprehensive analysis of the results is given. The results of the classifiers will not only help us to understand the performance of the method but will also enable us to judge the weaknesses present in the method. As the classifiers used for classification are based on supervised learning therefore a target class or desired output is used that was taken from the neurologist using the web interface.

4.1- Naive Bayes

A total of 232 spiral drawings are used in this method. The input of the naive bayes classifier is the PCA applied HHT generated features of these spiral drawings. The use of PCA facilitated to reduce the dimension of the features. To apply the PCA, firstly a subset of features on basis of rated spirals is selected and then PCA is applied on this subset. Secondly, coefficients of the principal components are used in order to calculate scores for all the dataset. The training and testing set is separated using cross validation (10 folds). The target class which can be either 0 or 1 is attached along with the input features to make it capable for supervised learning. 1 represents 'Dyskenesia' and 0 represents 'Off' in the target class. Here is the outcome of the classifier:

Correctly Classified Instances	148	63.7931%
Incorrectly Classified Instances	84	36.2069%
23	65	
19	125	

Table 2- Performance Vector

The above mentioned results displays the classification of naive bayes classifier for all the 232 spiral drawings. We can see that a total of 148 spiral drawings have been classified correctly by the classifier giving a correctly classified percentage of almost 64% while the percentage for incorrectly classified instance is 36%. Now if we have a close look on the confusion matrix we can see that Dyskenesia spirals which are represented by class 1 here are classified pretty good here, 125 spirals out of 144 are classified correctly. Only 19 spirals out of 144 are miss classified. Which means that percentage drop in the overall result is not because of Dyskenisea spirals but it is because of Off spirals. It can also be seen in the confusion matrix as only 23 spirals are classified correctly, rest are miss classified by the classifier. This point of

view will be further cleared if we have a look on the output of another classifier which is C4.5 classifier.

4.2- C4.5 Classifier

For C4.5 classifier, the same number of spiral drawings have been used as they were used for the experiment with Naive Bayes classifier. The usage of same number of spirals for each classifier will help to compare and understand the results correctly. The input of the classifier is again the PCA applied HHT generated features of these spiral drawings. The same target class has been used again as the number of spiral drawings were the same. 1 represents 'Dyskenesia' and 0 represents 'Off' in the target class. Here is the outcome of the classifier:

Correctly Classified Instances	133	57.3276%
Incorrectly Classified Instances	99	42.6724%
13	75	
24	120	

Table 3- Performance Vector

The figure above shows the classification of C4.5 decision tree for all the 232 spiral drawings. Again it can be noted that the decision tree is performing well for Dyskenisea spirals as most of the spirals are classified correctly. It can be seen that a total of 133 spiral drawings have been classified correctly by the classifier giving a correctly classified percentage of almost 58% and from these correctly classified spirals 120 are dyskenisea spirals, while on the other hand the percentage for incorrectly classified instance is 42%, and the major contribution in this percentage is of Off spirals. It means that again Off spirals are doing not that much good.

4.3- Sensitivity and Specificity

Sensitivity and Specificity are the two statistical measures which are used for the performance of a classification test.

Sensitivity-

It refers to the ability of the test to recognize positive results.

$$\text{sensitivity} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}}$$

Specificity-

It refers to the ability of the test to recognize negative results.

$$\text{specificity} = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false positives}}$$

Sensitivity / Specificity Table of Naive Bayes and C4.5		
Parameters	Naive Bayes	C4.5
TP	23	13
TN	125	120
FP	19	24
FN	65	75
Sensitivity	0.26	0.14
Specificity	0.86	0.83

Table 4- Performance Measures

It can be seen from the above sensitivity and specificity parameters that the sensitivity is low while specificity is comparatively high for both the classifiers. Here sensitivity refers to the measure of proportion of bradykenisea spirals which are correctly identified, and low sensitivity shows that method is not performing well for such spirals. Also as sensitivity is low, therefore the number of false negative is high. On the other hand, specificity refers to the measure of proportion of dyskenisea spirals which are correctly identified, and high specificity shows that method is actually performing much better for these spirals. But only good specificity percentage is not enough for a method to be practically useful, sensitivity percentage should also be high. As it is not in this method therefore the method cannot be said to be useful in practice.

4.4- ROC curve

ROC curve is defined by False Positive Rate and True Positive Rate as x and y axes respectively. True Positive Rate is the same as Sensitivity or it can be defined as the fraction of true positives out of the positives (TPR = true positive rate) and False Positive Rate is the same as $1 - \text{specificity}$ or it can be defined as the fraction of false positives out of the negatives (FPR = false positive rate). The figure below shows the ROC curve of the system:

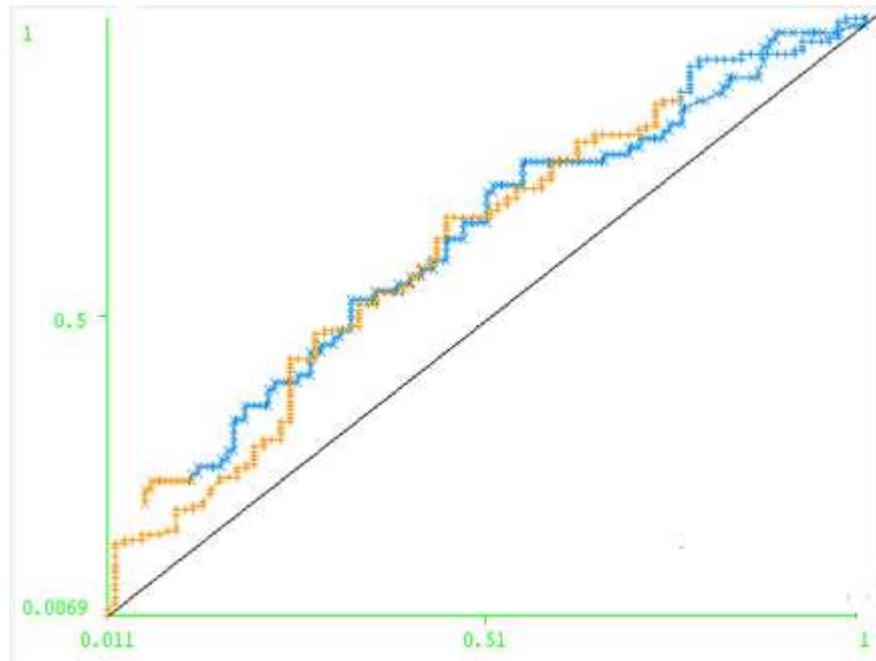


Figure 6: ROC curve; x = false positive rate; y = true positive rate; depicting bad performance

The figure above, depicts the performance of the classifier. It can be seen clearly that the graph is not in the region of perfect classification. If a graph lies close to the y-axis, i.e. close to the True Positive Rate, it depicts good classification and a graph close to the x-axis, i.e. False Positive Rate depicts bad classification. As in this case, the resultant ROC is not in the perfect classification region, therefore the system cannot be said to be feasible for practical implementation in real life scenarios. The possible reasons for such a classification by the classifiers is discussed in the next section.

4.5- Discussion:

The results explained above depicts that the overall percentage of successful classification is not good. A promising reason is the manual ratings of the spirals which form the desired output class. The desired output class required for supervised learning is not very well determined by the neurologist. It is difficult from the spiral drawings to determine whether the patient was "off" or dyskinetic at the time of the test. Even the neurologist was provided with the option to skip the spiral for which he is not sure that from which class the spirals belongs but even then it was very thorny for him to rate the spirals only by looking at its drawing without any other information. In his second attempt for rating spirals a whole lot had been skipped out but still the discrimination of the spirals was not guaranteed. So it will be better to try some other method, more precise and accurate, to obtain better spiral ratings from the neurologist. For-example a clinical study can be designed in which the patients will use the

test battery (the PDA) and at the same time the doctor will rate their motor state by using clinical rating scales. In this way more accurate ratings can be achieved.

Also it would be better to classify the spirals in more than two classes. Instead of considering only the two stages, the intermediate stages should also be taken into account. A patient can also be “very off”, “moderately off”, “a little off” and “a little dyskinetic”, “moderately dyskinetic” and “very dyskinetic”. If the neurologist is provided with enough information, so that he can rate the spirals in each of these categories, possibly a better classification can be achieved. Also if we have a look for the classification of Bradykinesia spirals, the outcome of the method is disappointing. The reason is that the dataset is consisted only on spiral data, tapping test data is not included in it. If we go through to the publication of westin et al [9] it can be seen there that bradykinesia was even not successfully classified by using DWT and tapping test data was required for it. Same is the case is here with HHT. Additional data may be required for the correct classification of such spirals.

Another reason of bad performance is the visibility of patient’s self assesment alongwith the spirals on the web interface. It was a mistake to show neurologist the patients' self-rating. He is often using this in his own rating when he in fact cannot see a pattern himself. It means that he is relying too much on the self assesment of the patient for making his own rating. This is also an explanation for the bad performance. Much better performance is achieved when i too used self-rating as input, correct classification rate increased to 85 %. These results are much more better than the previous one and shows that self rating is playing an important role in the classification. It means that the neurologist has relied much on the self rating instead of focussing on the spiral drawing. It might not be possible for the neurologist to classify only by looking at the spiral drawing. That is why he is focusing much on patient’s self rating. There should be some other way instead of still spiral drawing which can help the neurologist to make better classification. A possible techniqe which can prove to be useful is provided in future work.

Future Work:

Instead of using still spiral images, it may be better to display spiral animations on the web page. The spiral animation will provide much more information to the neurologist and he will be able to rate the spiral more accurately even if there is no pattern available in the image. The playing time of animation will be the same as the completion time which is already available in the database and the same x y coordinates can be used to draw the animation. There must

also be a replay option which will further facilitate the neurologist and will help to acquire better manual ratings.

Also as far as the methodology is concerned, another technique named 'improved hilbert huang transform' is available and is used in many research articles by the researchers. This technique is supposed to overcome the weaknesses present in the standard HHT method. One can use this technique to further improve the classification results.

Conclusion

A novel computer method for the Classification of cause of impairment in spiral drawings has been proposed in this thesis work. The methodology aims to classify bradykinesia(possibly caused by low medicine levels) and dyskinesia(possibly caused by high medicine levels) in to two separate classes. A dataset of 400 spiral drawings has been used in this work. The spiral drawings undergoes through a transformation known as HHT to decompose it into a set of features. Furthermore PCA has been applied on it to reduce the dimentionality of the features. To apply the PCA, firstly a subset of features on basis of rated spirals is selected and then PCA is applied on this subset. Secondly, coefficients of the principal components are used in order to calculate scores for all the dataset. The training and testing set is separated using cross validation (10 folds). As the work has been done in context of the supervised learning paradigm therefore a target class or desired output was needed. This desired output class has been taken from an expert neurologist by using a web interface. This output class is then used alongwith the HHT features in weka to test the performance of the method. Then lastly a predictive model is created to find relationships between extracted features and the manual ratings.

In conclusion, the method does not respond as accurately as it was expected to do. The overall percentage of successfull classification is not good. Although it is performing much better for Dyskinesia spirals but in case of Bradykinesia spirals the method responds very badly. One possible reason behind this is the manual ratings of the spirals which form the desired output class. The desired output class required for supervised learning is not very well determined by the neurologist. It is difficult from the spiral drawings to determine whether the patient was "off" or dyskinetic at the time of the test. To overcome this drawback, a clinical study can be designed in which the patients will use the test battery (the PDA) and at the same time the doctor will rate their motor state by using clinical rating scales. In this way more accurate ratings can be achieved. Another reason is that there is no difference between the two classes of spiral drawings.

References:

- [1] *Parkinson's disease, Challenges, Progress, and Promise*. National Institute of Neurological Disorders and Stroke National Institutes of Health. November 2004
- [2] *The professional guide to parkinsons disease*. Parkinson's Disease Society , uk, april 2010.
- [3] *Epidemiology of Parkinson's disease by Lonneke M L de Lau, Monique M B Breteler, Lancet Neurol 2006; 5: 525–35, Vol 5 June 2006*
- [4] *Wireless real-time electronic data capture for self-assessment of motor function and quality of life in Parkinson's disease by Nyholm D, Kowalski J, Aquilonius SM. MovDisord 2004;19(4):446–51.*
- [5] *Quantifying drug-induced dyskinesias in the arms using digitised spiral-drawing tasks. Liu X, Carroll CB, Wang SY, Zajicek J, Bain PG. J Neurosci Methods 2005;144(1):47–52.*
- [6] *Validity of spiral analysis in early Parkinson's disease. Saunders-Pullman R, Derby C, Stanley K, Floyd A, Bressman S, Lipton RB, et al. Mov Disord 2008;23(March (4)):531–7.*
- [7] *Quantitative tremor measurement with the computerized analysis of spiral drawing. Rudzinska M, Izvorski A, Banaszkiwicz K, Bukowczan S, Marona M, Szczudlik A. Pol J Neurol Neurosurg 2007;41(6):510–6.*
- [8] *A home environment test battery for status assessment in patients with advanced Parkinson's disease. Westin J, Dougherty M, Nyholm D, Groth T. Comput Methods Programs Biomed. 2010 Apr;98(1):27-35.*
- [9] *A new computer method for assessing drawing impairment in Parkinson's disease. Westin J, Ghiamati S, Memedi M, Nyholm D, Johansson A, Dougherty M, Groth T. J Neurosci Methods. 2010 Jun 30;190(1):143-8.*
- [10] *Duodenal levodopa infusion monotherapy vs. Oral polypharmacy in advanced Parkinson disease, D. Nyholm, A.I. Nilsson Remahl, N. Dizdar, Askmark, Neurology 64 (2) (2005) 216–223.*
- [11] *Parkinson's disease home diary: further validation and implications for clinical trials, Movement Disorders, J R.A. Hauser, F. Deckers, P. Leher, 19 (12) (2004) 1409–1413.*
- [12] *Wireless real-time electronic data capture for self-assessment of motor function and quality of life in Parkinson's disease, D. Nyholm, J. Kowalski, S.M. Aquilonius, 19 (4) (2004) 446–451.*
- [13] *Electronic motor function diary for patients with Parkinson's disease: a feasibility study, K.E. Lyons, R. Pahwa, Parkinsonism and Related Disorders 13 (5) (2007) 304–307.*
- [14] *Patient compliance with paper and electronic diaries, J A.A. Stone, S. Shiffman, J.E. Schwartz, J.E. Broderick, M.R. Hufford, Controlled Clinical Trials 24 (2) (2003) 182–199.*
- [15] *Quantifying drug-induced dyskinesias in the arms using digitised spiral-drawing tasks, X. Liu, C.B. Carroll, S.Y. Wang, J. Zajicek, P.G. Bain, Journal of Neuroscience Methods 144 (1) (2005) 47–52.*
- [16] *Quantitative tremor measurement with the computerized analysis of spiral drawing, M. Rudzinska, A. Izvorski, K. Banaszkiwicz, S. Bukowczan, M. Marona, A. Szczudlik, Polish Journal of Neurology and Neurosurgery 41 (6) (2007) 510–516.*
- [17] *A novel computer-based technique for the assessment of tremor in Parkinson's disease, N.M. Aly, J.R. Playfer, S.L. Smith, D.M. Halliday, Age and Ageing 36 (4) (2007) 395–399.*
- [18] <http://www.rwc.uc.edu/koehler/biophys/2a.html>
- [19] http://scimath.unl.edu/MIM/files/MATEXamFiles/Favinger_MATpaper_Final_EDITED.pdf
- [20] http://www.engineeringtoolbox.com/motion-formulas-d_941.html
- [21] http://science.dirbix.com/physics/tangential_velocity

[22] *Hilbert-Huang transform and wavelet analysis of time history signal** SH! Chun-xiang LUO Qi-feng
Research Institute of Structural Engineering and Disaster Reduction, Tongji University, Shanghai 200092,
China

[23] "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series
analysis," N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H.
Liu, Mar. 1998.

[24] *Hilbert–Huang Transform and Its Applications*, N. E. Huang and S. S. P. Shen, Eds. Hackensack, NJ:
World Scientific, 2005.

[25] "The application of Hilbert–Huang transforms to meteorological datasets," *J. Atmos. Ocean. Apr.* 2004.

[26] *A criterion for selecting relevant intrinsic mode functions in empirical mode decomposition by albert ayenu
prah, Advances in Adaptive Data Analysis Vol. 2, No. 1 (2010) 1–24*

[27] *A Tutorial on Principal Component Analysis*, Jonathon Shlens, Systems Neurobiology Laboratory, Salk
Institute for Biological Studies CA 92093-0402 (Dated: December 10, 2005; Version 2)

[28] *A Comparison of Three Different Methods for Classification of Breast Cancer Data*, Soria, D.; Garibaldi,
J.M.; Biganzoli, E.; Ellis, I.O.; Sch. of Comput. Sci., Univ. of Nottingham, Nottingham, UK, *Machine
Learning and Applications*, 2008. ICMLA '08. Seventh International Conference

[29] *Diagnosis of Heart Disease for Diabetic Patients using Naive Bayes Method*
International Journal of Computer Applications (0975 – 8887), Volume 24– No.3, June 2011