

Article

Selection of Entropy Based Features for Automatic Analysis of Essential Tremor[†]

Karmele López-de-Ipiña^{1,*}, Jordi Solé-Casals², Marcos Faundez-Zanuy³, Pilar M. Calvo¹, Enric Sesa³, Unai Martínez de Lizarduy¹, Patricia De La Riva⁴, Jose F. Marti-Masso⁴, Blanca Beitia¹ and Alberto Bergareche⁴

¹ Systems Engineering and Automation Department, University of the Basque Country UPV/EHU, Donostia 20018, Spain; pilarmaria.calvo@ehu.eus (P.M.C.); unai.martinezdelizarduy@ehu.eus (U.M.-d.-L.); mariablanca.beitia@ehu.eus (B.B.)

² Data and Signal Processing Research Group, University of Vic—Central University of Catalonia, Vic, Catalonia 08500, Spain; jordi.sole@uvic.cat

³ Escola Superior Politècnica Tecnocampus (UPF), Mataró, Catalonia 08302, Spain; faundez@tecnocampus.cat (M.F.-Z.); sesa@tecnocampus.cat (E.S.)

⁴ BioDonostia Health Institute, Neurology Department Hospital Donostia, Donostia 20014, Spain; patricia.delarivajuez@osakidetza.net (P.D.L.R.); josefelix.martimasso@osakidetza.net (J.F.M.-M.); jesusalberto.bergarecheyarza@osakidetza.net (A.B.)

* Correspondence: karmele.ipina@ehu.es; Tel.: +34-943-01-86-67; Fax: +34-943-01-72-30

† This paper is an extended version of one paper published in the 4th IEEE International Work Conference on Bioinspired Intelligence, Donostia, Spain, 9–12 June 2015.

Academic Editors: Carlos M. Travieso-González and Jesús B. Alonso-Hernández

Received: 8 March 2016; Accepted: 9 May 2016; Published: 16 May 2016

Abstract: Biomedical systems produce biosignals that arise from interaction mechanisms. In a general form, those mechanisms occur across multiple scales, both spatial and temporal, and contain linear and non-linear information. In this framework, entropy measures are good candidates in order provide useful evidence about disorder in the system, lack of information in time-series and/or irregularity of the signals. The most common movement disorder is essential tremor (ET), which occurs 20 times more than Parkinson's disease. Interestingly, about 50%–70% of the cases of ET have a genetic origin. One of the most used standard tests for clinical diagnosis of ET is Archimedes' spiral drawing. This work focuses on the selection of non-linear biomarkers from such drawings and handwriting, and it is part of a wider cross study on the diagnosis of essential tremor, where our piece of research presents the selection of entropy features for early ET diagnosis. Classic entropy features are compared with features based on permutation entropy. Automatic analysis system settled on several Machine Learning paradigms is performed, while automatic features selection is implemented by means of ANOVA (analysis of variance) test. The obtained results for early detection are promising and appear applicable to real environments.

Keywords: permutation entropy; essential tremor; automatic drawing analysis; Archimedes' spiral; non-linear features; automatic feature selection

1. Introduction

Biomedical systems produce biosignals that arise from interaction mechanisms. In a general form, those mechanisms occur across multiple scales, both spatial and temporal, and contain linear and non-linear information. Complex fluctuations are habitually present in the output variables of real systems. These fluctuations are due to noise but also contain information about the dynamics of the system. Linear methods can capture global aspects of the dynamics, but the different approaches are not able to discern all the relevant physical details [1,2]. In this framework, the measurement

of non-linear features (for example the entropy of the system) is an essential and useful tool to investigate the state of the system. This analysis provides the information encoded in the system and the probability distributions of all its possible states [1]. Particular characteristics of data affect the applicability of entropy-based methodologies. In this sense, stationarity property, noise level, length of the time series, *etc.*, are of great relevance, since important information may be present in the temporal dynamics. Habitually, all those aspects are not taken into account [1,3].

Biological and biomedical systems generate time series generated that contain deterministic and stochastic components [4]. Classic methods of signal and noise analysis can deal with part of the interesting features, but they only model linear components without yielding any information about non-linearities, irregularities or stochastic components. When analyzing inconspicuous changes this complex information could be essential. Massimiliano Zanin *et al.* [1] present a review based on biomedical applications, which includes the analysis of heart rhythms, anesthesia, electroencephalography (EEG) or cognitive neuroscience. These last ones, related to neurological diseases, are challenging due to their variability and the tremendous impact they exert on the society.

Essential tremor (ET) affects individuals worldwide and it is 20 times more present than Parkinson's disease. The prevalence of ET in the western countries is around 0.3%–4.0%. Males and females of 40 years old are affected approximately equally (incidence of 23.7 per 100,000 people per year). Several studies suggest prevalence from 3.9% to 14.0% in these patients, and 50% to 70% of the cases of ET seems to have a genetic origin [5]. Essential tremor is a rhythmic tremor (4–12 Hz) that only occurs when the affected muscle is exerting. The amplitude of the tremor increases and varies with age, but there is no gender distinction. Additionally, physical or mental stress could worsen the tremor. Parkinson's disease (PD) and Parkinsonism can occur simultaneously with essential tremor. In fact, the prevalence of PD in people with ET is greater than in the general population. Concerning symptoms, hand tremor is the predominant one (as it is in PD) and occurs in nearly all cases, followed by other sort of tremors like voice, head, face, neck, tongue, leg and trunk. Due to its (partially) genetic origin, PD and ET often occur in individuals of the same family [5].

In order to manage and palliate the symptoms, it is of great importance to clinically detect the earliest manifestations of the disorder. In the past few years, approaches seeking early diagnosis of ET have made significant advances towards the development of consistent clinical biomarkers. Despite the valuableness of those biomarkers, cost and technology requirements make unviable to apply such tests to all patients with motor disorders. In this framework, early detection performed through non-invasive intelligent techniques could be a good alternative. Non-technical staff could use those methodologies without altering the patients' abilities, because handwriting analysis, drawing analysis or speech analysis are not perceived as stressful tests by the patients. Furthermore, these are low-cost techniques and do not require medical equipment nor extensive infrastructures [6–9].

Doctors traditionally use handwritten task in order to diagnose ET. Archimedes' spiral is a well-known and established test [10], therefore we will focus our work on the analysis of such drawings. Traditionally, the analysis of handwriting was performed offline because only the writing itself (strokes on a paper) was available. Currently modern digitizing tablets and pens (with or without ink) can gather data keeping all temporal information and the dynamics of the whole process. In this case, the analysis is referred to as online. Modern digitizing tablets capture the x , y and z coordinates of the movement of the writing process, the applied pressure on the surface and the azimuth and altitude angles (the angles of the pen in the horizontal plane and with respect to the vertical axis, respectively) [11]. This makes it possible to analyze both static (offline) and dynamic (online) features [12].

The results presented here are part of a wider cross study on the diagnosis of ET carried out by the Biodonostia Health Institute with the objective of characterizing this impairment, and it is based on families with identified genetics loci. Among several drawings and handwriting exercises, Archimedes' spiral has been chosen and will be explored to determine the best non-linear biomarkers for early diagnosis and following of ET [13,14]. The following sections analyze classic static and dynamic linear

features and also non-linear ones based on several entropy algorithms. Automatic methodologies will be used in the selection of biomarkers. Finally, the quality of the selected features is measured by ANOVA, multiple comparisons test, and Machine Learning paradigms.

2. Materials

2.1. Acquisition System

The acquisition is carried out by means of an Intuos Wacom 4 digitizing tablet. The pen tablet USB [11] captures the following information (Figure 1) at a sample frequency of 100 Hz [12,13]: the spatial coordinates (x , y), the pressure, and azimuth and altitude angles. Using this set of dynamic data, further information can be inferred such as acceleration, speed, instantaneous trajectory angle, instantaneous movement, tangential acceleration, curvature radius, centripetal acceleration *etc.* [12,15].

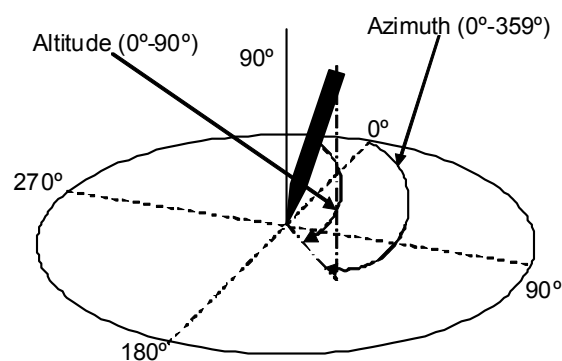


Figure 1. Information extracted from the digitizing tablet.

2.2. Database of Individuals

The database BIODARW consists of 21 healthy subjects (control group, CR) and 29 patients diagnosed with ET (ET group) with identified genetics loci. For all patients, register of electrophysiological test (EPT) and functional magnetic resonance imaging (fMRI) are also available. Both hands are considered in order to perform the handwritten test. Therefore, the total number of samples is $2 \times (21 + 29) = 100$. The handwriting test consists of drawing a line, drawing the Archimedes' spiral and handwriting with dominant hand and non-dominant hand. In this work only Archimedes' spiral is used. The database has variability with regard to: tremor frequency, amplitude and pattern, rating scale values, and demographic data (age and gender). Subjects were recruited from patients of a previous descriptive study that considers familiar and sporadic ET cases and controls from the Movement Disorders Unit at the Donostia University Hospital (San Sebastian, Spain). Written informed consent, fully approved by the local ethics committee at the Donostia University Hospital, was obtained from all participants in this new study. Table 1 summarizes the features of the group with ET with regard to EPT, diagnosis and demography [13,14]. In addition to the standard clinical exploration (Neuropsychological and Electrophysiological Studies), evaluation of ET was carried out by recording the drawing of an Archimedes' spiral. The Montreal Cognitive Assessment [16] with the Fahn–Tolosa–Marin (FTM) tremor rating scale, which assesses different cognitive domains, was used to determine possible cognitive dysfunction; and criteria from the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) of the American Psychiatric Association (APA) were used for the diagnosis of depression and anxiety disorders. Surface electromyography (EMG) was recorded from wrist extensor and flexor muscles using surface electrodes placed over the muscle bellies 3 cm apart. The filters were set with a band-pass of 10–1000 Hz. A triaxial accelerometer was placed over the first dorsal interossei muscle of the hand.

Table 1. Some samples of the electrophysiological test (EPT). Fahn–Tolosa–Marin (FTM) scale values for the selected individuals with ET (ET_X).

ET_X	EPT Features			Diagnosis	Demography	
	Frequency (Hz)	Amplitude (v)	Pattern	FTM Scale	Age	Gender
ET_01	8.5	20	synchronous	1	48	Female
ET_02	6.5	variable	alternating	8	72	Male
ET_03	10.5	200	synchronous	1	46	Male
ET_04	4.5	503.6	synchronous	3	80	Female
ET_05	6.6	298	synchronous	22	68	Female
ET_06	9.5	46	synchronous	2	46	Female
ET_07	5	173	synchronous	50	75	Male
ET_08	6.5	159	synchronous	40	75	Male
ET_09	8	128	asynchronous	9	75	Female

2.3. Individuals Selected for the Study

From the original database of 100 samples, a subset of the samples of Archimedes' spiral is selected. The team of medical doctors carried out this selection. For the control group, the best sample (usually the dominant hand) is kept, but in some small number of cases the non-dominant hand is used as well. For the ET group, only the sample with the best quality is chosen (one hand), but five subjects are discarded due to the poor quality of the samples. Thus, this sub-database BIODARWO consists of 51 samples: 27 samples for the control group and 24 samples for the ET group.

3. Methods

3.1. Online Drawing Applied to Health Analysis

In the medical field, the study of handwriting has proved to be an aid to diagnosing and tracking some diseases of the nervous system. For instance, handwriting skill degradation and Alzheimer's disease (AD) appear to be significantly correlated [17,18], and some aspects of handwriting can be good indicators for its diagnosis [6,17,18] or help to differentiate between mild Alzheimer's disease and mild cognitive impairment [17,18]. Additionally, the analysis of handwriting has proved useful to assess the effects of substances such as alcohol [19], marijuana [20] or caffeine [21]. Thanks to modern acquisition devices, the field of psychology has also benefited from the analysis of handwriting. Rosenblum *et al.* [22] link the proficiency of the writers to the length of the in-air trajectories of their handwritings. The visual inspection of the pen-down image suggests a progressive degree of impairment when drawing becomes more disorganized, and only ET patients in mild states can achieve the three dimensional effect. The visual information provided by the pen-up drawing of pathological individuals also indicates a progressive impairment and disorganization when the individuals try to plan the drawing. It is also important to note that the comparison of pen-up drawings between people with disease and the control group shows noticeable differences as well. In addition to the increased time in-air, there is a higher number of hand movements before putting the pen on the surface to draw. We consider that these graph-motor measures applied to the analysis of drawing and writing functions may be a useful alternative to study the precise nature and progression of the drawing and writing disorders associated with several neurodegenerative diseases.

3.2. Pressure Derived Measures and in-Air Analysis

One of the major advantages of online drawing is the possibility of evaluating the handwriting pressure quantitatively. For instance, people diagnosed with AD produce softer and simpler strokes, and people with ET are less stable. ET and Parkinson's disease are other examples [13,14,23–25]. For example, in [18], they have reported the proficiency of writers with regard to the length of the in-air trajectories and the pressure of their handwritings [26]. A very interesting aspect of modern

online analysis of handwriting is that it can take into account information gathered when the writing device was not exerting pressure on the writing surface. As an example Figure 2 shows the acquisition of the word INEXPUGNABLE: the pen-up (in-air) movement information is represented in red, while the pen-down (on-surface) movement is represented in blue. Our previous experiments on biometric recognition of people revealed that these two kinds of information are complementary [12], and in fact have a similar discriminating capability, even when using a database of 370 users [15,25,26]. Figure 3 shows an example of the Archimedes' spiral performed by a control subject (Figure 3a) and a subject with ET (Figure 3b). The blue line corresponds to pen-down (on surface drawing), and the red line to pen-up (in-air drawing). The drawing is very regular in the case of the control subject (Figure 3a), and with clear irregularities in the case of the ET subject (Figure 3b) not only with regard to Cartesian components but also to the control of the pressure level.

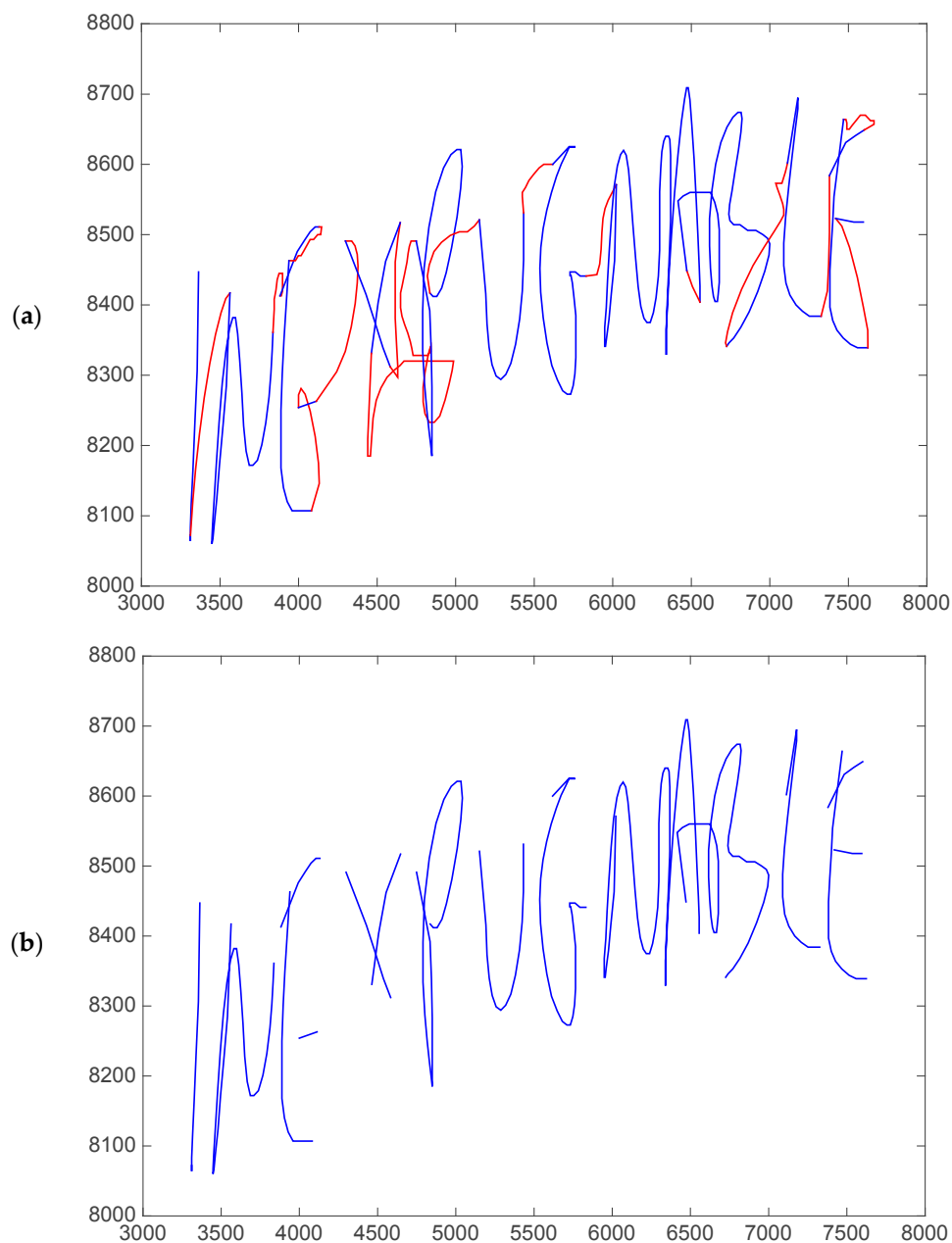


Figure 2. Cont.

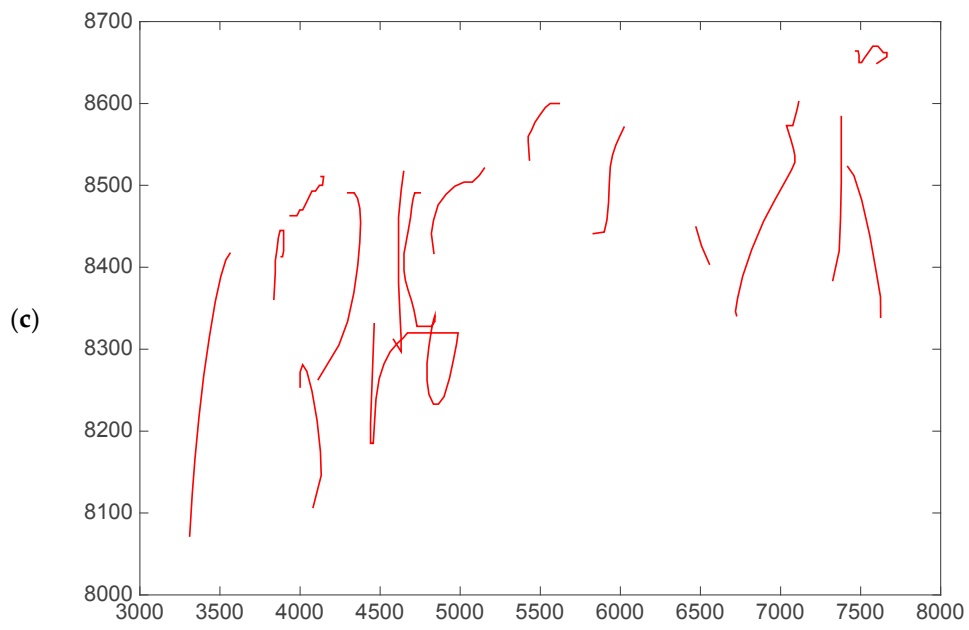


Figure 2. Execution of the word INEXPUGNABLE as captured by the acquisition device: (a) both pen-up and pen-down strokes; (b) pen-down strokes; and (c) pen-up strokes.

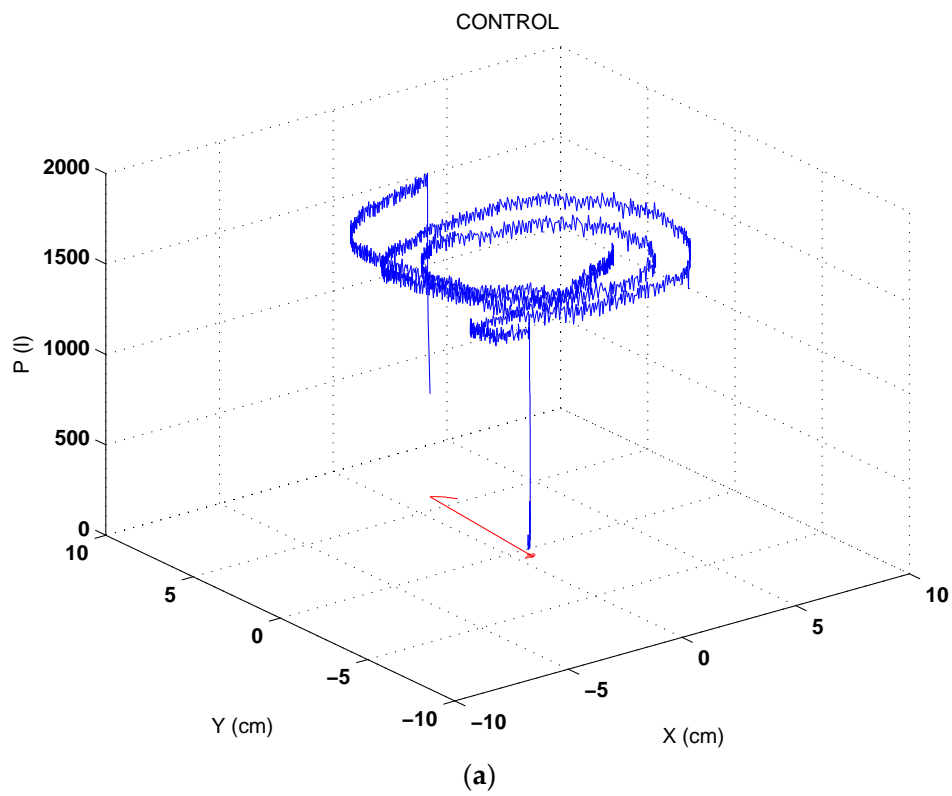


Figure 3. Cont.

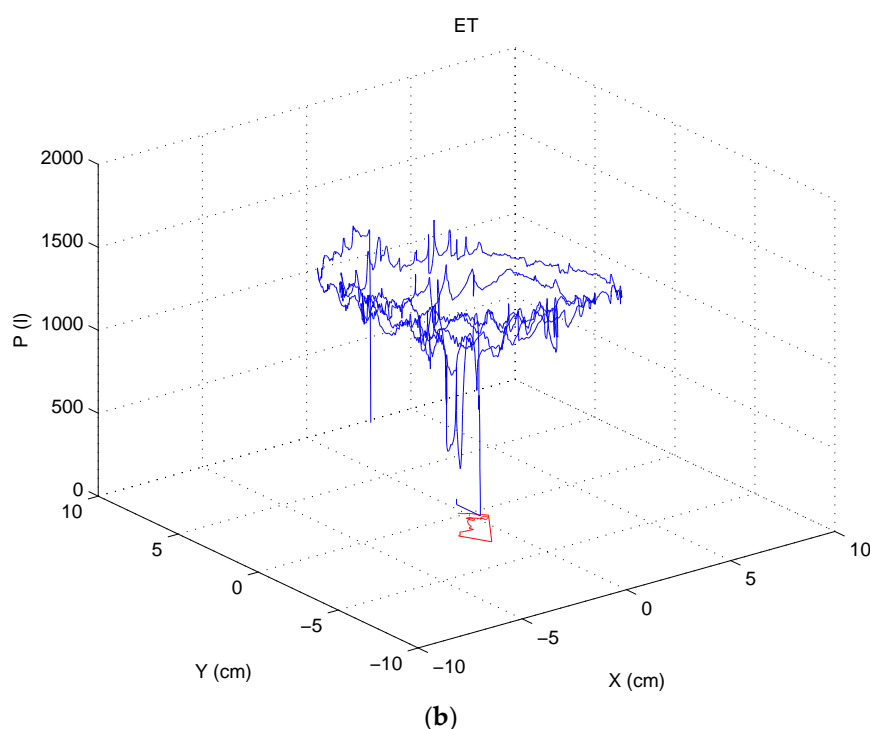


Figure 3. Archimedes' spiral 3D drawings performed by a control subject (a); and a subject with ET (b). Cartesian coordinates X (cm), Y (cm) and Pressure level P (l). The blue line corresponds to pen-down (on surface drawing), and the red line to pen-up (in-air drawing). The drawing is very regular in the case of the control subject (a); and with clear irregularities in the case of the ET subject (b) not only with regard to Cartesian components but also to the control of the pressure level.

3.3. Features Extraction

The research presented here has a preliminary nature; its aim is to define thresholds for a number of biomarkers related to handwriting. It is part of a wider study focused on early ET detection. Feature search in this work aims at preclinical evaluation in order to define useful tests for ET diagnosis [5,13,14].

3.3.1. Linear Features

In this study, the aim is to automatically distinguish handwriting of ET patients from healthy subjects by means of the analysis of different linear features (LF) and their variants (max, min, mean and median) in handwriting, *i.e.*, the following.

1. Time related measures: Time in-air, time on-surface and total time (in-air plus on-surface). Time has been measured as number of samples.
2. Spatial components and their variants: X and Y Cartesian coordinates, altitude (O) and azimuth (A) angles, and angle and modulus polar components (Z and R, respectively) and their projections over a horizontal axis for both pen-down and pen-up signal (see Figure 4 for an example of the distortion of the polar components in the sample of the ET patient).
3. Pressure and its variants.
4. Dynamic features and their variants: Speed and acceleration for both pen-down and pen-up signals.
5. Zero crossing rate: The rate which evaluates the sign-changes along a signal.
6. Frequency domain: Spectral components for both pen-down and pen-up signals (Figure 5).

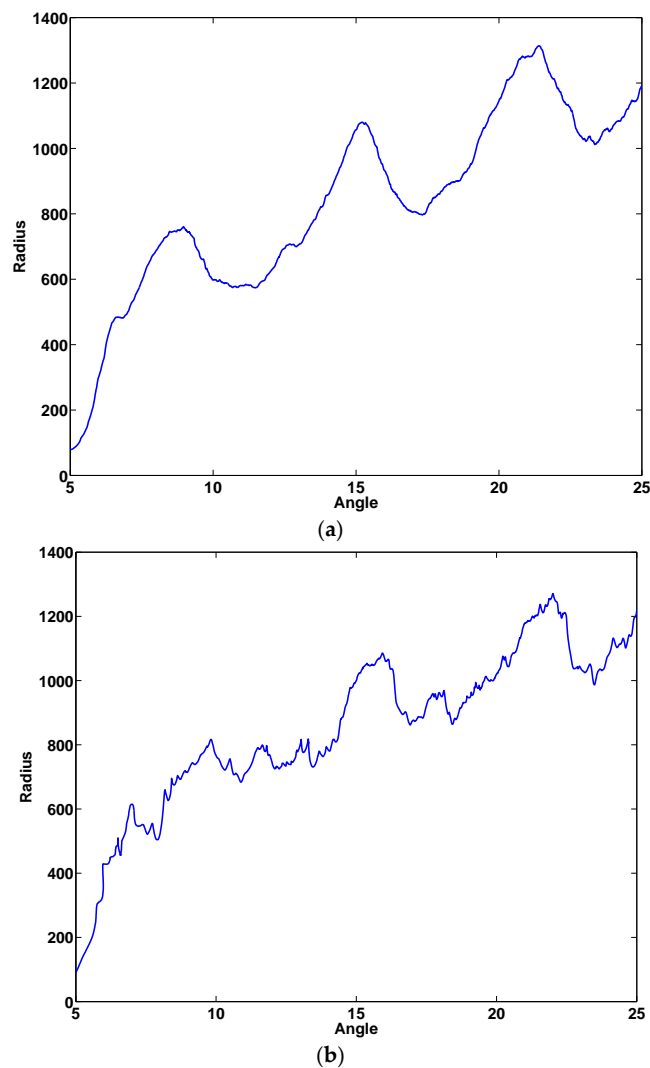


Figure 4. Polar components radius (R) and angle (θ) of the Archimedes' spiral drawing performed by a control subject (a); and a subject with essential tremor (b). The ET patient produces a clear distortion in these components.

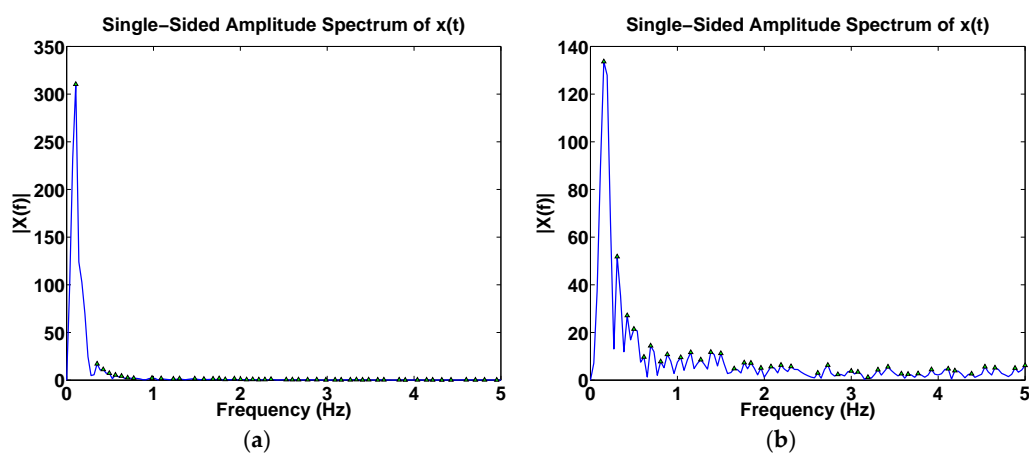


Figure 5. Spectrum of the X component of the *pen-down* (on-surface) signal for a control subject (a); and a subject with essential tremor (b). The spectrum has more components for the subject with ET: it shows many peaks along the whole spectrum, not only in the low frequencies.

3.3.2. Non-Linear Features: Entropy

Entropy is a measure of disorder in physical systems and also a basic quantity with multiple field-specific interpretations. It has been associated with disorder, state-space volume, or lack of information [1,2,27,28]. Shannon entropy is often considered as the classic and most natural way to measure the expected value (average) of the information in a signal [3,4,29,30]. Richman *et al.* analyze that entropy, and in relation to dynamic systems is the rate of information production [31]. On the one hand, some authors point out that calculation of entropy usually requires very long data sets that in the case of biomedical signals can be difficult or impossible to obtain. On the other hand, methods for entropy estimation of a system represented by a time series are not suitable for analyzing the short and noisy data sets of biomedical studies [1,31]. In the following subsections, we present several proposals for calculating entropy used in this work.

Shannon Entropy

The entropy $H(X)$ of a single discrete random variable X is a measure of its average uncertainty. Shannon entropy [29,30] is calculated by the equation:

$$H(X) = - \sum_{x_i \in \Theta} p(x_i) \log p(x_i) = -E[\log p(x_i)] \quad (1)$$

where X represents a random variable with a set of values Θ and probability mass function $p(x_i) = P_r\{X = x_i\}$, $x_i \in \Theta$, and E represents the expectation operator. Note that $p \log p = 0$ if $p = 0$.

For a time series that represents the output of a stochastic process, that is, an indexed sequence of n random variables, $\{X_i\} = \{X_1, \dots, X_n\}$, with a set of values $\theta_1, \dots, \theta_n$, respectively, and $X_i \in \theta_i$, the joint entropy is defined by:

$$H_n = H(X_1, \dots, X_n) = - \sum_{x_i \in \theta} \dots \sum_{x_n \in \theta_n} p(x_1, \dots, x_n) \log p(x_1, \dots, x_n) \quad (2)$$

where $p(x_1, \dots, x_n) = P\{X_1 = x_1, \dots, X_n = x_n\}$ is the joint probability for the n variables X_1, \dots, X_n .

By applying the chain rule to Equation (2), the joint entropy can be written as the addition of conditional entropies, where each of them is a non-negative quantity:

$$H_n = \sum_{i=1}^n H(X_i | X_{i-1}, \dots, X_1) \quad (3)$$

Therefore, as a conclusion, the joint entropy is an increasing function of n . The rate at which the joint entropy grows with n , *i.e.*, the entropy rate h , is defined as:

$$h = \lim_{n \rightarrow \infty} \frac{H_n}{n} \quad (4)$$

Approximate Entropy *versus* Sample Entropy

The approximate entropy is a statistical measure that smoothens transient interference and can suppress the influence of noise by properly setting the parameters of the algorithm. It can be used in the analysis of both stochastic and deterministic signals [32,33]. This is crucial in the case of biological signals, which are outputs of complex biological networks and may be deterministic, stochastic, or both. Approximate entropy provides a model-independent measure of the irregularity of the signals. The algorithm summarizes a time series into a non-negative number, where higher values represent more irregular systems [32,33].

The method examines time series for similar epochs [34]: more frequent and more similar epochs lead to lower values of approximate entropy. ApEn (m, r, n) measures the conditional probability that

two sequences of length n which are similar for m points remain similar at the next sample point, within a tolerance r . Thus, a low value of approximate entropy reflects a high degree of regularity. Approximate entropy algorithm counts each sequence as matching itself to reduce bias; sample entropy $\text{SmEn}(m, r, n)$ was developed in order not to count self-matches.

The sample entropy is defined for a time series of n points. We first define the $n - m + 1$ vectors $\mathbf{x}_m(i) = \{u(i+k) : 0 \leq k \leq m - 1\}$ as the vectors of m data points from $u(i)$ to $u(i+m-1)$. The distance between two such vectors is defined as $d[\mathbf{x}_m(i), \mathbf{x}_m(j)] = \max_k \{|u(i+k) - u(j+k)| : 0 \leq k \leq m - 1\}$, i.e., the maximum difference of their scalar components. The sample entropy $\text{SmEn}(m, r, n)$ is defined as:

$$\text{SmEn}(m, r, n) = \lim_{n \rightarrow \infty} \{-\ln(A^m(r)/B^m(r))\} = -\ln(A/B) \tag{5}$$

where

$$A = [(n - m - 1)(n - m)/2] A^m(r) \tag{6}$$

and

$$B = [(n - m - 1)(n - m)/2] B^m(r) \tag{7}$$

$B^m(r)$ is the probability that two sequences match for m points:

$$B^m(r) = (n - m)^{-1} \sum_{i=1}^{n-m} B_i^m(r) \tag{8}$$

where $B_i^m(r)$ is $(n - m - 1)^{-1}$ times the number of vectors $\mathbf{x}_m(j)$ within a tolerance r of $\mathbf{x}_m(i)$. Similarly, $A^m(r)$ is the probability that two sequences match for $m + 1$ points:

$$A^m(r) = (n - m)^{-1} \sum_{i=1}^{n-m} A_i^m(r) \tag{9}$$

where $A_i^m(r)$ is $(n - m - 1)^{-1}$ times the number of vectors $\mathbf{x}_{m+1}(j)$ within a tolerance r of $\mathbf{x}_{m+1}(i)$. The scalar r is the tolerance for accepting matches. In the present investigation we used the parameters recommended in [35]: $m = 2$ and $r = 0.2$, and standard deviation of the sources is normalized to 1. Sample entropy is a robust quantifier of complexity for instance for electroencephalography (EEG) signals [36], and can be used as a marker for the presence of artifacts in EEG recordings [37]. The quantity A/B is the conditional probability that two sequences within a tolerance r points remain within r of each other at the next point. In contrast to approximate entropy, which calculates probabilities in a template-wise fashion, sample entropy calculates the negative logarithm of a probability associated with the time series as a whole.

Multivariate Multiscale Permutation Entropy

Permutation entropy (PE) directly analyzes the temporal information contained in the time series; furthermore, it has the quality of simplicity, robustness and very low computational cost [1,3,4]. Bandt and Pompe [38] introduce a simple and robust method based on the Shannon entropy measurement that takes into account time causality by comparing neighboring values in a time series. The appropriate symbol sequence arises naturally from the time series with no prior knowledge assumed [1].

Permutation entropy is calculated for a given time series $\{x_1, x_2, \dots, x_n\}$ as a function of the scale factor ϵ . In order to be able to compute the permutation of a new time vector X_j , $S_t = [X_t, X_{t+1}, \dots, X_{t+m-1}]$ is generated with the embedding dimension m and then arranged in increasing order: $[X_{t+j_1-1} \leq X_{t+j_2-1} \leq \dots \leq X_{t+j_n-1}]$. Given m different values, there will be

$m!$ possible patterns π , also known as permutations. If $f(\pi)$ represents its frequency in the time series, its relative frequency is $p(\pi) = f(\pi) / (L/s - m + 1)$. The permutation entropy is then defined as:

$$PE = - \sum_{i=1}^{m!} p(\pi_i) \ln p(\pi_i), \quad (10)$$

Summing up, permutation entropy refers to the local order structure of the time series, which can give a quantitative measure of complexity for dynamic time series. This calculation depends on the selection of the m parameter, which is strictly related to the length n of the analyzed signal. For example, Bandt and Pompe [38] suggested the use of $m = 3, \dots, 7$ following always this rule:

$$m! < n. \quad (11)$$

If m is too small (smaller than 3), the algorithm will work incorrectly because it will only have few different states for recording, but it depends on the data. When using long signals a larger value of m is preferable but it would require larger computational times.

Multiscale Entropy (MSE) was proposed by Costa *et al.* [4] and has been shown to be a robust method for analyzing structural effects at multiple time scales present in complex real data. As Morabito *et al.* explain in their application to EEG signal processing [2], usual coarse-graining procedure can be implemented as follows:

- (i) From the original time series, multiple successive coarse-grained versions are extracted by $y_j(\varepsilon)$, where ε is the scale factor. Each element of the coarse-grained time series is calculated as:

$$y_j(\varepsilon) = \frac{1}{\varepsilon} \sum_{i=(j-1)\varepsilon+1}^{j\varepsilon} x_i \quad (12)$$

- (ii) For each scaled series, the PE is calculated.

In Morabito *et al.* [2], multivariate multiscale permutation entropy (MMSPE) is used as a methodology that can integrate information loss related to relevant cross-channel variability and to channel correlation. These authors refer to Keller and Laufer [39] that didn't take into account variation over multiple scales. In previous works, Ahmed and Mandic [40] analyzed a similar proposal, multivariate multiscale entropy (MMSE), for sample entropy.

Thus the cross-channel complexity with $f_s = \frac{1}{T}$, representing the multivariate PE (MPE), can be calculated for all time $s \in [f_s T - m]$ as the permutation entropy (PE) of p_j :

$$H_{MPE}(s) = \sum_{j=1}^{d!} p_j \log_2 p_j. \quad (13)$$

And the MMSPE algorithm is implemented according to the two following steps:

- (i) Different time scales of increasing length are defined by coarse-graining the original multivariate time series, *i.e.*, $\{x_i, t\}$, for $i = 1, \dots, c$ (where c is the number of channels) and for $t = 1, \dots, n$ (where n is number of samples in each time series). For a scale factor ε , the elements of the multivariate coarse-grained time series can be derived as:

$$y_{i,j}(\varepsilon) = \frac{1}{\varepsilon} \sum_{t=(j-1)\varepsilon+1}^{j\varepsilon} x_{i,t} \text{ for } i = 1, \dots, c, \text{ and, } 1 \leq j \leq \frac{n}{\varepsilon}. \quad (14)$$

- (ii) Calculate the multivariate permutation entropy, MPE, for each coarse-grained multivariate $y_{i,j}(\varepsilon)$ and all variants of average.

3.3.3. Feature Sets

In the experimentation, the following feature sets will be used (see Abbreviations Section for a complete list of acronyms and their meanings):

1. Linear features set (LF), the set described in Section 3.3.1
2. Non-linear features sets (NLF) that consist of LF and the features described in Section 3.3.2: linear features + Shannon entropy (SE), linear features + Approximate Entropy (ApEn), linear features + Sample Entropy (SmEn) and linear features + permutation entropy (PE).
3. Set after selection of features by ANOVA: selection of linear features (SLF), Selection of linear features + Shannon Entropy (SSE), Selection of linear features + Approximate Entropy (SApEn), Selection of linear features + Sample Entropy (SSmEn) and Selection of linear features + permutation entropy (SPE).

3.4. Automatic Selection of Features by ANOVA

In a next stage the feature set will be optimized and the best feature with regard to common significance level will be selected automatically. Thus, automatic feature selection is performed by an ANOVA one-way test [41]. This test analyzes the p -value under the null hypothesis that all samples in a matrix X are drawn from populations with the same mean. If p is near zero, it casts doubt on the null hypothesis, and suggests that at least one mean is significantly different from the others. Common significance level will be less than 0.05 to select a feature as discriminative. In the box-plot graphic the columns of X suggests the size of the F -statistic and the p -value. Large differences in the centre lines of the boxes correspond to large values of F and correspondingly small values of p and therefore useful feature for discrimination tasks. Then, in order to confirm the selection, a *multiple comparison test* of the means of the group is performed by the MATLAB (R2014) [41] function *multcompare*.

3.5. Modeling and Automatic Classification

The main goal of the present work comprises feature search in handwriting aiming at preclinical evaluation in order to define tests for ET diagnosis. These features will define the control group (CR group) and the essential tremor group (ET group). A secondary goal is the optimization of computational cost (measured as execution time of the algorithms) with the aim of making these techniques useful for real-time applications in real environments. Thus, automatic classification will be modeled taking this into consideration. We used three different classifiers based on WEKA (3.7) software suite [42]:

1. A Support Vector Machine (SVM) with polynomial kernel;
2. A Multi Layer Perceptron (MLP) with number of units in the hidden layer given (NNHL) by $= \max(\text{Attribute/Number} + \text{Classes/Number})$ and training step (TS) = NNHL · 10; and
3. A k Nearest Neighbor (k -NN) k -NN algorithm.

The results were evaluated using Accuracy (Acc, in %) and Classification Error Rate (CER, in %) [7,42–44]. For training and validation steps, we used k -fold cross-validation with $k = 10$, where accuracy will be the average of the k iterations. Cross-validation is a robust validation method for variable selection [43]. Repeated cross-validation (as calculated by the WEKA environment) allows robust statistical tests. We also use the measurement automatically provided by WEKA “Coverage of cases” (0.95 level); that is, the confidence interval at 95% level. Despite the small sample size that could point to the use of Leave One Out Cross-Validation (LOOC), we were oriented to the selected k -fold cross-validation by the model number, the sample richness and previous works [14,22].

3.6. General Procedure of the Experimentation

In previous works, both linear and non-linear features (fractal dimension and Shannon entropy) have been used with good results [22]. However, the system was not able to detect subtle changes in the case of level tremor. Therefore, the main goal of these experiments was to examine the potential of other entropy algorithms, and to select the optimum entropy based features for automatic measurement of the degradation of the drawing of Archimedes' spiral with ET. All experiments will be oriented to subtle changes and prodromal stages.

In this work, the general procedure of the experimentation is divided into two phases: a feature selection phase and an optimization phase (Figure 6).

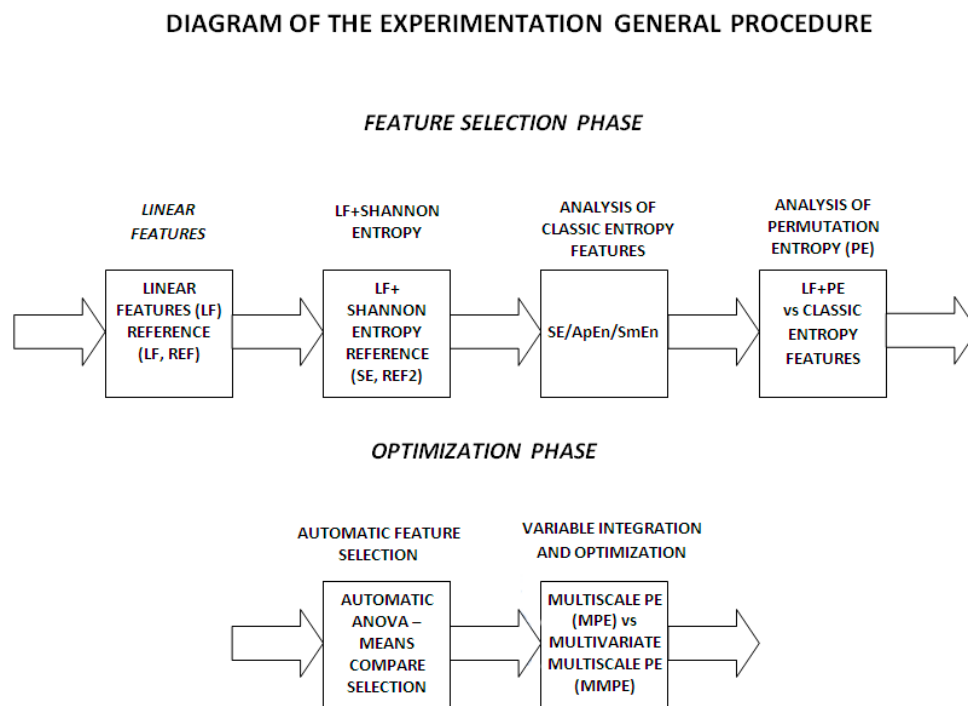


Figure 6. Diagram of the general procedure of the experimentation. The experimentation is divided into two phases. Firstly, an *entropy feature selection phase* is carried out across four steps: (i) analysis of classic linear features (this will be the reference); (ii) analysis of linear features and Shannon entropy; (iii) analysis of classic entropy features ApEn and SmEn and adjustment and selection of optimum parameters, m and r ; and (iv) analysis of previously adjusted classic entropy features against permutation entropy. Secondly, an *optimization phase* is carried out following the next two steps: (i) an automatic selection of the best linear and non-linear features by statistical-medical criteria based on ANOVA and multiple comparison test of group means; and (ii) an optimization analysis of the PE based features based on signal scale decomposition and the integration of signals by the MPE and the MMSPE algorithms.

Firstly, an *entropy feature selection* is carried out following the next four steps:

- (1) *Analysis of classic linear features.* An automatic classification is carried out in order to obtain the reference rates (LF) for linear features.
- (2) *Analysis of linear features and Shannon entropy.* A second reference (SE) is calculated integrating linear and the classic entropy feature, Shannon entropy.
- (3) *Analysis of classic entropy features.* ApEn and SmEn are analyzed and compared in order to adjust and select optimum parameters for the algorithms (m and r).

- (4) *Entropy based classic features vs. permutation entropy.* An analysis of the previous adjusted classic entropy features against permutation entropy is carried out in order to obtain the optimum parameters for permutation entropy.

Then, an *optimization* phase is carried out across two steps:

- (1) *Automatic feature selection.* An automatic feature selection is carried out based on statistical-medical criteria by ANOVA and a multiple comparisons test of the group means. Thus, only the linear and non-linear features with a p -value under a fixed threshold are selected and an optimum feature set is obtained.
- (2) *Optimization.* Finally, an optimization analysis of the PE based features is carried out. Two new algorithms are used based on: (1) scale analysis by multiscale permutation entropy (MPE); and (2) the integration of signals and scale analysis by the novel multivariate multiscale permutation entropy (MMSPE) algorithm. This last step is oriented to integrate signal correlations and to reduce even more the number of features for real-time system purposes.

4. Results and Discussion

The experimentation has been carried out with the balanced subset BIODARWO. In this section, we have used the automatic classification and modeling, the general procedure of experimentation and the feature sets described in Section 3.3.

4.1. Phase of Entropy Feature Selection

In a first stage the reference rates are calculated for both *linear features* (LF, 186 features) and a non-linear proposal that consist of *linear features and Shannon entropy* (SE, 198 features). Thus, integration of Shannon entropy outperforms LF reference for MLP and k -NN with similar results for SVM paradigm but with an increase in the number of features of around 7% (Table 2).

Table 2. Feature sets and feature number (FN). The number following the small letter is the value of that parameter.

	LF	SE	ApEn-m3	SmEn-m3	PE-m7t7	SLF	SSE	SAPEn-m3	SSmEn-m3	SPE-m7t7
FN	186	198	198	198	198	70	76	73	77	78

Then, an analysis of classic entropy features has been carried out. ApEn and SmEn are used and compared in order to adjust and select the optimum parameters (m and r). For both algorithms, $m = 2, 3$ and tolerance $r = 0.2$ have been evaluated. Figure 7 shows the obtained CER (%) values for the three paradigms with regard to the references LF and SE. ApEn cannot improve SE, but promising results are obtained for SmEn and $m = 3$, outperforming the rates for MLP and SVM. However, there is not a clear improvement in performance in the case of k -NN, the option with less computational cost with regard to the model generation and to the classification process.

Secondly, the integration of classic entropy based features is compared with permutation entropy features (PE) for different orders (m) and time delays (t). In our particular case and due to the length of the signals that have a mean of 3497 samples, m parameter was fixed up to $m = 7$. In this step, the previous best option, SmEn with $m = 3$, is evaluated with the references and different options of PE. Figure 8 shows the obtained CER (%) values for the three paradigms. PE based features improve the CER (%) rates in most of the cases. The best results are obtained for $m = 7$ and $t = 7$; this configuration improves the previous references with 198 features for MLP (15.69%) and SVM (17.65%) and it maintains a similar rate for k -NN. Good results are achieved also with k -NN with less computational cost (21.57%).

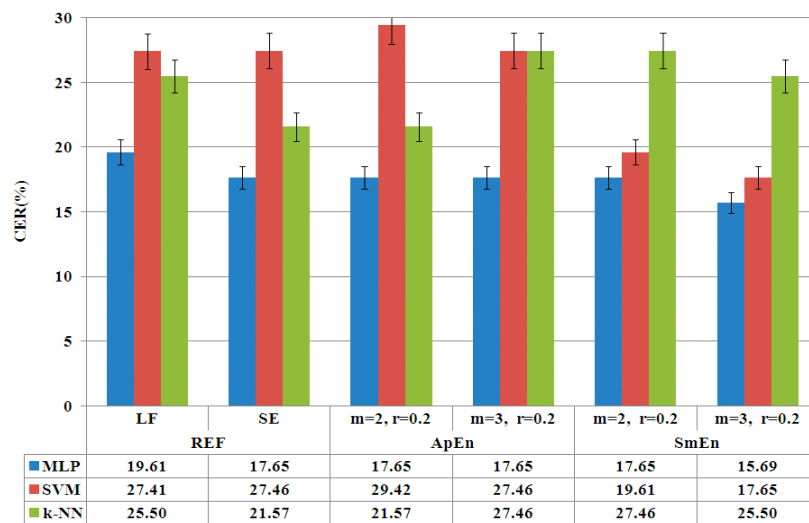


Figure 7. Analysis of classic entropy features. ApEn and SmEn are analyzed and compared in order to adjust and select the optimum parameters (m and r). The graphic presents CER (%) value for the three paradigms compared with the references for linear features and Shannon entropy (SE). The best results are obtained for $m = 3$ and SmEn, but there is not an improvement in the case of k -NN with a worse result with regard to SE [14].

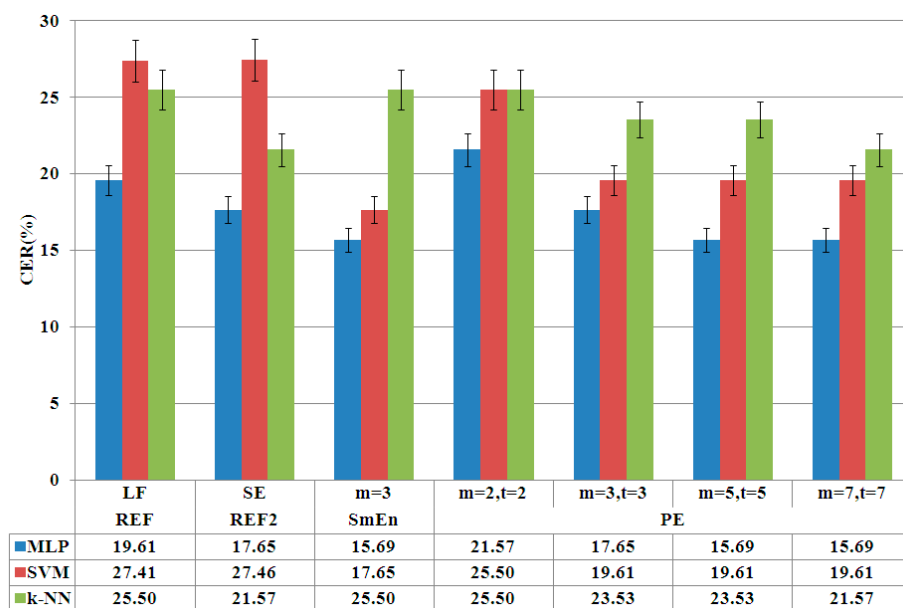


Figure 8. LF and classic entropy based features *vs.* SE and permutation entropy (PE). The best option, LF with SmEn for $m = 3$, is compared with the references and different options of LF and PE. The graphic presents CER (%) value for the three paradigms. The PE based features improve the CER (%) rates in most of the cases. The best results are obtained for $m = 7$ and $t = 7$, which improve the previous references for MLP and SVM and are equal to the k -NN system [14].

4.2. Optimization Phase

In the next stage, during the optimization phase an automatic feature selection is carried out based on statistical-medical criteria by ANOVA and multiple comparison test of the group means. In this automatic process medical and statistical criteria are combined and only the linear and non-linear features with p -value < 0.05 are selected. Thus, the number of features is reduced by around 60%

and an optimum feature set is obtained. Table 2 summarizes the results for the references and the previous best options. In order to show the values considered for the parameters of the algorithms, a number is added after the letter that identifies each parameter. Figure 9 shows an example of the statistical comparison for *pen-up* (in air) time. This feature is relevant with $p < 0.05$ as it can be seen in the analysis of the means of the groups CR and ET, which are significantly different.

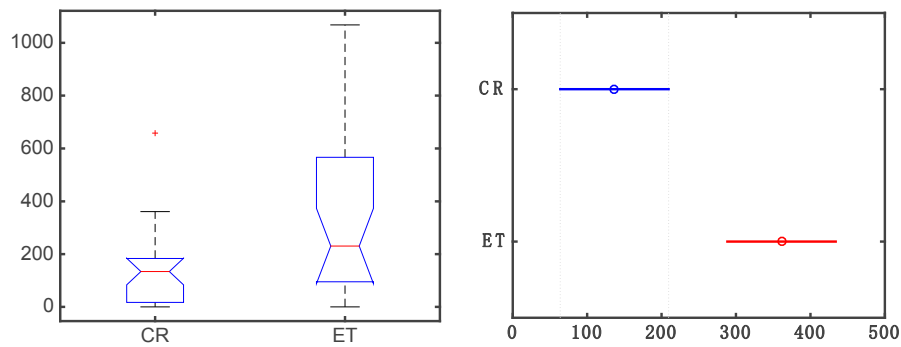


Figure 9. In the optimization phase an automatic feature selection is carried out based on statistical-medical criteria by ANOVA and the analysis of the means of the groups (p -value). In this process, only the linear and non-linear features with a $p < 0.05$ are selected, and an optimum feature set is obtained. The example shows the box-plot for the ANOVA test (left); and analysis of the group means for *pen-up* (in air) time (right). The means of CR and ET groups are significantly different.

After the optimization, an automatic classification is carried out with the selected feature sets in order to analyze the robustness of the generated sets. Figure 10 shows CER (%) for the references and the three paradigms (MLP, SVM and k -NN). There is a clear improved performance of the models after feature optimization for all the options. The best option is clearly MLP with the SPE-mt7t set (3.93%), and also SVM (7.85%) and k -NN obtain promising results for future real time developments (Table 3). Moreover, permutation entropy seems to be a powerful non-linear feature for modeling the non-linear dynamic of the system.

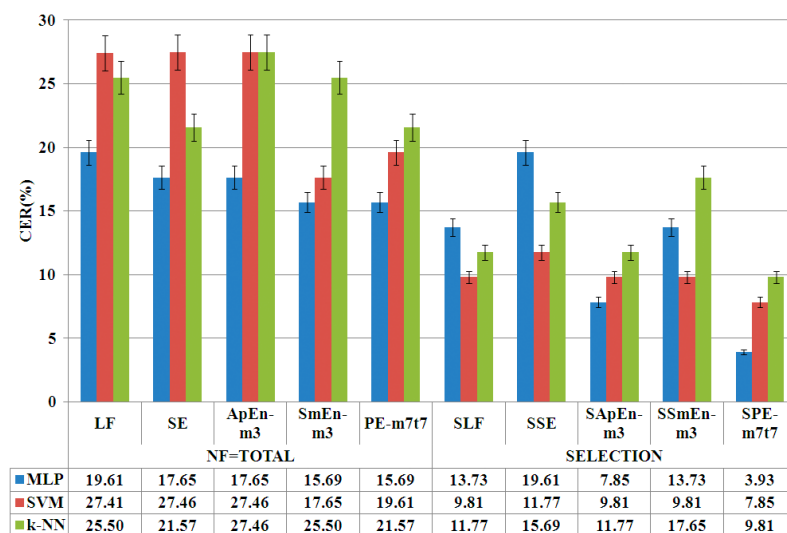


Figure 10. Automatic classification and obtained CER (%) for the references and the three paradigms with feature sets, before and after the automatic selection. The results show a clear improved performance of the models after feature optimization for all the options. The best option after the optimization is MLP with the SPE-m7t7 set (Selection of linear features + permutation entropy for $m = 7$ and $t = 7$).

Table 3. Results summary. CER (%) for the three paradigms with LF, SLF and the optimum feature set.

	MLP	SVM	k-NN
LF	19.61	27.41	25.50
SLF	13.73	9.81	11.77
SPE-m7t7	3.93	7.85	9.81

On the other hand, an analysis for classes has also been carried out. Figure 11 shows in detail the Accumulative Classification Error Rate (% ACCERR) for CR and ET as a measure of optimization impact by class. It can be seen that the optimization and the integration of the novel entropy algorithms improve the ACCERR in all the cases. The best option is SPE-m7t7 for the three paradigms, even for *k*-NN, which yields appropriate rates for both classes with less computational cost in modeling and classification phases. The computational cost is also reduced in both phases with regard to the original feature sets. SPE-m7t7 improves not only the global rates, but also the class rate even for the less powerful model: *k*-NN. Figure 12 shows the analysis for classes with regard to Acc (%) with similar results. SPE-m7t7 is confirmed as the best option.

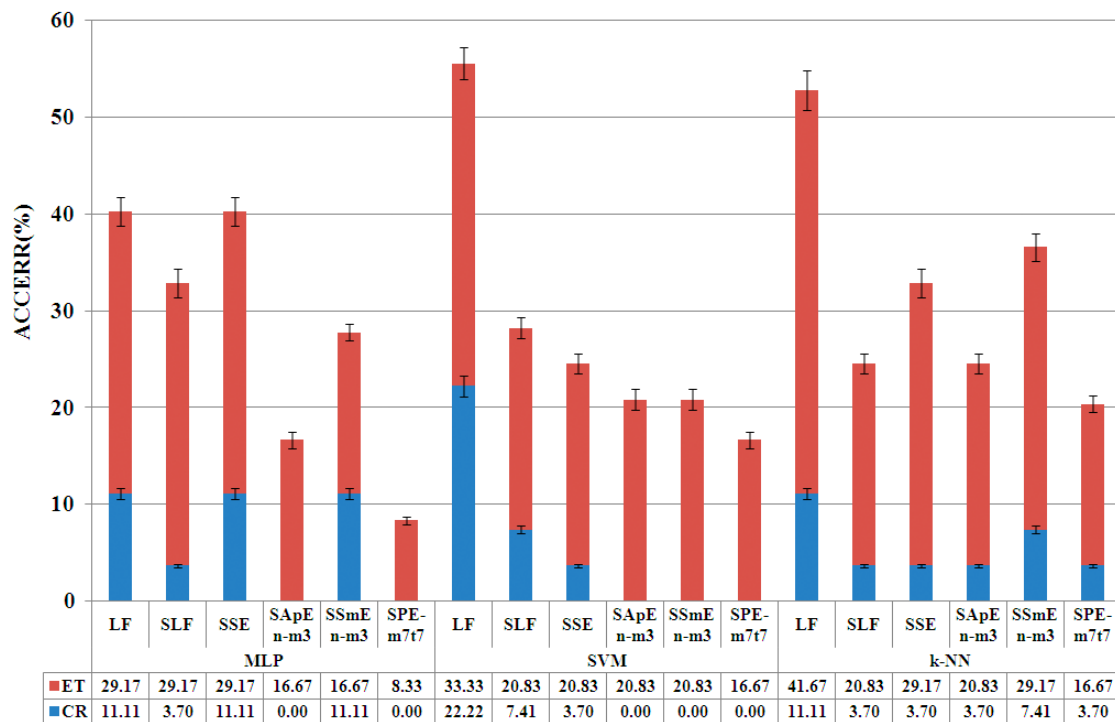


Figure 11. Analysis of the entropy based feature sets and impact on classes by Accumulative Classification Error Rate (% ACCERR). The optimization and the integration of the novel entropy algorithms improve the ACCERR in all the cases. The best option is SPE-m7t7 for the three paradigms even for *k*-NN, which presents appropriate rates for both classes.

In the final stage, an optimization analysis of the PE is carried out. This analysis is based on the signal scale decomposition (MPE) and the integration of signals (*X*, *Y*, pressure) by the novel MMSPE algorithms. This last step is oriented to integrating signal correlations and to further reducing the number of features for real-time system purposes. Figure 13 shows the effect of the introduction of different scales in the PE calculation. These differences in the scales could contribute to the detection of subtle changes in the signal that could be very useful for early ET detection.

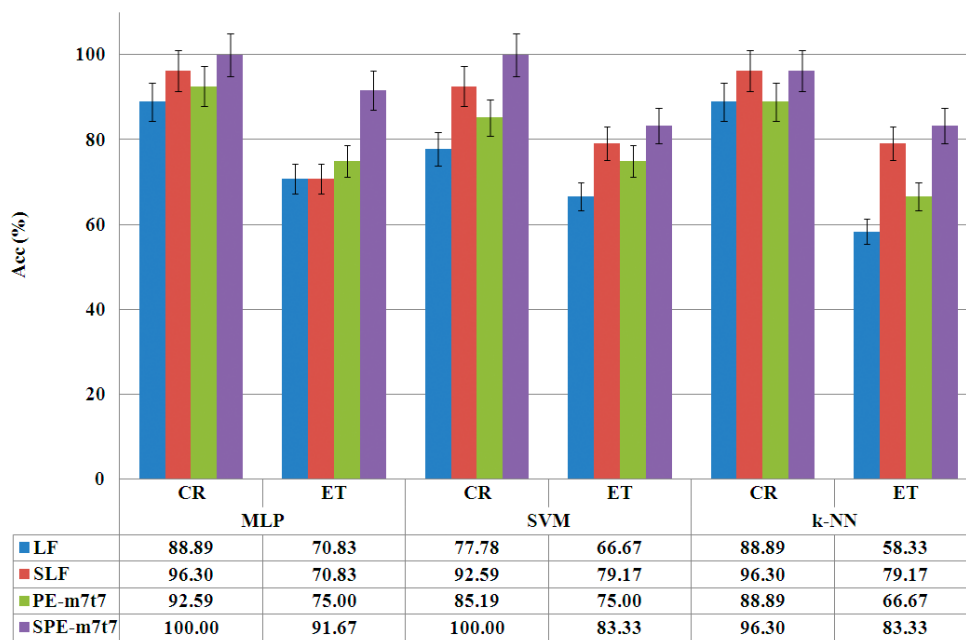


Figure 12. Accuracy (%) of classes for the paradigms for the references and the best options. SPE-m7t7 improves not only the global rates, but also the class rate even for the less powerful model: *k*-NN.

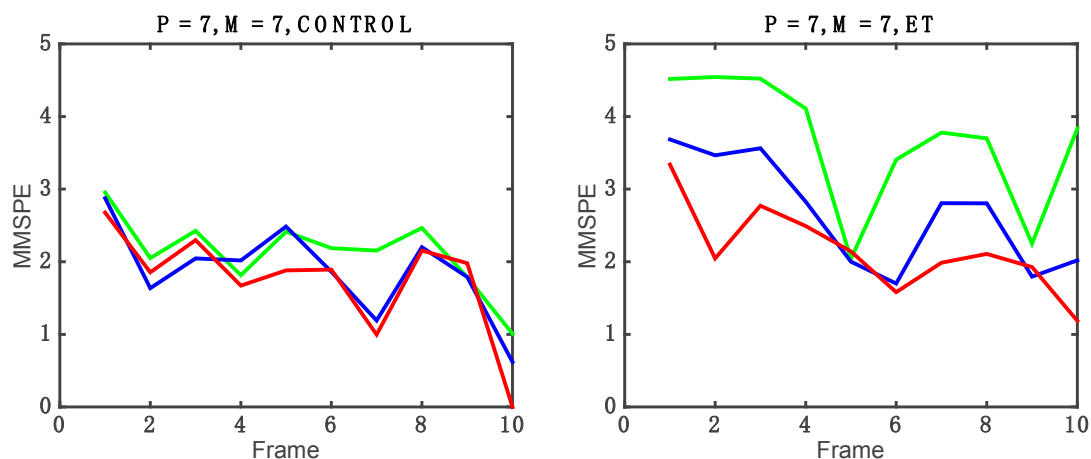


Figure 13. Detail of MMSPE (value: *y*-axis) across the time (frame: *x*-axis) of the drawing of the Archimedes’ spiral performed by a control subject (left) and a subject with essential tremor (right) for three scales ($\epsilon = 3$: green line, $\epsilon = 1$: blue line, $\epsilon = 2$: red line), $p = 7$ and $t = 7$.

These novel approaches provide very promising results. Figure 14 shows the obtained results. On the one hand, the SMPE approach with $\epsilon = 3$, $p = 7$ and $t = 7$, which includes three scales, outperforms SPE-m7t7 with MLP, and obtains similar rates to the ones of the other two paradigms but with an increase in the number of features (84). On the other hand, the option that integrates three signals with $\epsilon = 3$, $p = 7$ and $t = 7$, SMMSPE (76 features) set, decreases the number of features around 10% with regard to SMPE and 6% with regard to SPE-m7p7, but with similar CER for most of the paradigms. Only for MLP the results are worse due to a reduction in the relevant information of independent signals. Moreover, the computational cost decreases with regard to SPE not only in the training phase, but also during the classification process. The SMMSPE methodology seems to be very promising to analyze both interaction between different signals and subtle changes in ET subjects.

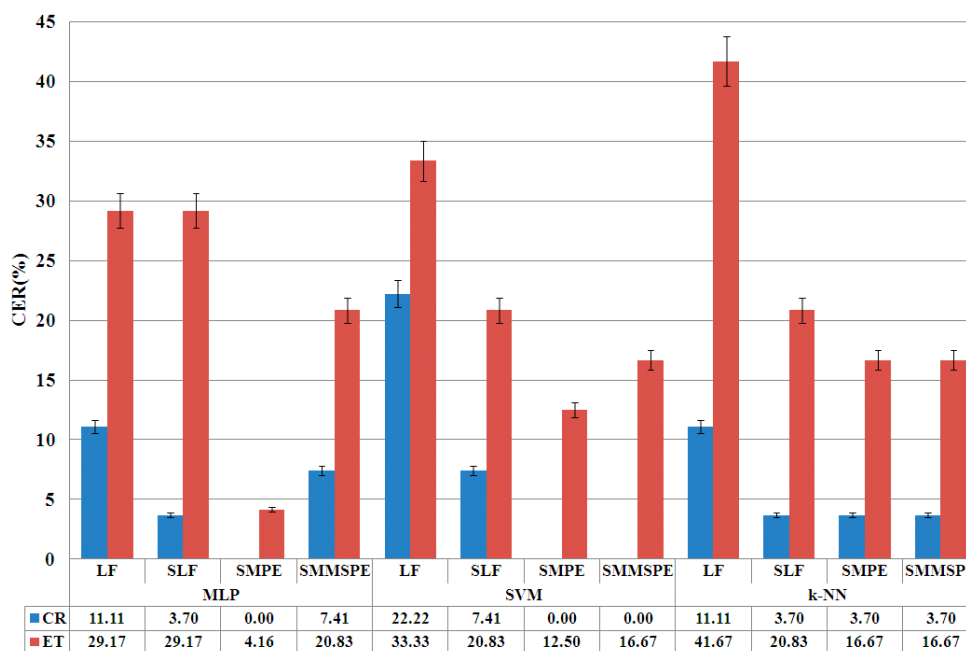


Figure 14. CER (%) for the paradigms with the references LF and SLF and the selected features set SMPE and SMMSPE.

Finally, the health specialists have noticed the relevance and usefulness of the system for selection of handwriting biomarkers for early diagnosis of ET.

5. Conclusions

This work focused on the selection of non-linear biomarkers from drawings and handwriting is part of a wider cross study on the diagnosis of essential tremor, which is developed in the Biodonostia Health Institute. Specifically, the main goal of the present work is the analysis of features in Archimedes' spiral drawing, one of the most used standard tests for clinical diagnosis of ET. In this sense, entropy based features have been added to a set of classic linear features (static and dynamic). Several entropy algorithms have been evaluated: Shannon entropy, approximate entropy, sample entropy and permutation entropy. The automatic analysis system consists of several Machine Learning paradigms, automatic features selection by ANOVA, and multiple comparison test (*multcompare*). The results are optimal even with a reduction of around 60% in the number of features after the selection process. The best option is MLP with permutation entropy. However, for real-time applications, MMSPE appears to be a promising methodology to analyze with less computational cost, both interactions between different signals and subtle changes in ET subjects. In future works, new non-linear features, entropy algorithms and automatic selection methodologies will be evaluated.

Acknowledgments: This work has been partially supported by the University of the Basque Country under project ref. UPV/EHU—58/14, SAIOTEK program and others from the Basque Government, the Spanish Ministerio de Ciencia e Innovación TEC2012-38630-C04-03, the University of Vic—Central University of Catalonia under the research grant R0904, INNPACTO program from the Spanish Government, and UPV/EHU Summer Courses Foundation.

Author Contributions: Karmele López-de-Ipiña, Jordi Solé-Casals, Marcos Faundez-Zanuy, Pilar M. Calvo, Unai Martínez De Lizarduy, Enric Sesa and Blanca Beitia have pre-processed the handwritten signals, and designed, developed and evaluated the system and tools. Alberto Bergareche, Patricia De La Riva and Jose F. Martí-Masso have designed and acquired the signals and collaborated in the system design and evaluation. All authors developed the system and participated writing of the manuscript, which has been read and approved by all of them

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CR	control group
ET	essential tremor
EPT	electrophysiological test
fMRI	functional magnetic resonance imaging
MoCA	Montreal Cognitive Assessment
EMG	electromyography
AD	Alzheimer's disease
LF	linear features (reference)
SLF	selection of linear features
NLF	non-linear features
SE	linear features + Shannon entropy (reference)
ApEn	linear features + Approximate Entropy
SmEn	linear features + Sample Entropy
EEG	electroencephalography
PE	permutation entropy
MPE	multi scale permutation entropy
MSE	multiscale entropy
MMSE	multivariate multiscale entropy
MMSPE	multivariate multiscale permutation entropy
MLP	Multi Layer Perceptron
NNHL	number of hidden layers units
SVM	Support Vector Machine
CER	classification error rate
ACCERR	Accumulative Classification Error Rate
ACC	Accuracy
SE	linear features + Shannon entropy
SSE	Selection of linear features + Shannon Entropy
SAPEn	Selection of linear features + Approximate Entropy
SSmEn	Selection of linear features + Sample Entropy
SPE	Selection of linear features + Permutation Entropy
SMPE	Selection of linear features + Multi scale Permutation Entropy
SMMSPE	Selection of linear features + Multivariate Multiscale Permutation Entropy

References

- Zanin, M.; Zunino, L.; Rosso, O.A.; Papo, D. Permutation entropy and its main biomedical and econophysics applications: A review. *Entropy* **2012**, *14*, 1553–1577. [[CrossRef](#)]
- Morabito, F.C.; Labate, D.; La Foresta, F.; Bramanti, A.; Morabito, G.; Palamara, I. Multivariate Multi-Scale Permutation Entropy for Complexiy Analysis of Alzheimer's Disease EEG. *Entropy* **2012**, *14*, 1186–1202. [[CrossRef](#)]
- Eguiraun, H.; López-de-Ipiña, K.; Martinez, I. Application of, Entropy and Fractal Dimension Analyses to the Pattern Recognition of Contaminated Fish Responses in Aquaculture. *Entropy* **2014**, *16*, 6133–6151. [[CrossRef](#)]
- Costa, M.; Goldberger, A.; Peng, C.K. Multiscale entropy analysis of biological signals. *Phys. Rev.* **2005**, *71*. [[CrossRef](#)] [[PubMed](#)]
- Louis, E.D.; Vonsattel, J.P. The emerging neuropathology of essential tremor. *Mov. Disord.* **2007**, *23*, 174–182. [[CrossRef](#)] [[PubMed](#)]
- Faundez-Zanuy, M.; Hussain, A.; Mekyska, J.; Sesa-Nogueras, E.; Monte-Moreno, E.; Esposito, A.; Chetouani, M.; Garre-Olmo, J.; Abel, A.; Smekal, Z.; *et al.* Biometric Applications Related to Human Beings: There Is Life beyond Security. *Cogn. Comput.* **2013**, *5*, 136–151. [[CrossRef](#)]
- López-de-Ipiña, K.; Alonso, J.B.; Solé-Casals, J.; Barroso, N.; Faundez-Zanuy, M.; Travieso, C.; Ecay-Torres, M.; Martinez-Lage, P.; Martinez-de-Lizardui, U. On the selection of non-invasive methods based on speech analysis oriented to Automatic Alzheimer Disease Diagnosis. *Sensors* **2013**, *13*, 6730–6745. [[CrossRef](#)] [[PubMed](#)]
- Laske, C.; Sohrobi, H.R.; Frost, S.M.; López-de-Ipiña, K.; Garrard, P.; Buscem, M.; Dauwels, J.; Soekadar, S.R.; Mueller, S.; Linnemann, C.; *et al.* Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer Dement.* **2015**, *11*, 561–578. [[CrossRef](#)] [[PubMed](#)]

9. Sesa-Nogueras, E.; Faundez-Zanuy, M. Biometric recognition using online uppercase handwritten text. *Pattern Recognit.* **2012**, *45*, 128–144. [[CrossRef](#)]
10. Pullman, S.L. Spiral Analysis: A New Technique for Measuring Tremor with a Digitizing Tablet. *Mov. Disord.* **1998**, *13*, 85–89. [[CrossRef](#)] [[PubMed](#)]
11. WACOM. Available online: <http://www.wacom.com> (accessed on 10 May 2016).
12. Faundez-Zanuy, M. Online signature recognition based on VQ-DTW. *J. Pattern Recognit.* **2007**, *40*, 981–982. [[CrossRef](#)]
13. López-de-Ipiña, K.; Bergareche, A.; De La Riva, P.; Faundez-Zanuy, M.; Calvo, P.M.; Roure, J.; Sesa-Nogueras, E. Automatic non-linear analysis of non-invasive writing signals, applied to essential tremor. *J. Appl. Log.* **2015**. [[CrossRef](#)]
14. López-de-Ipiña, K.; Iturrate, M.; Calvo, P.M.; Beitia, B.; Garcia-Melero, J.; Bergareche, A.; De La Riva, P.; Marti-Masso, J.F.; Faundez-Zanuy, M.; Sesa-Nogueras, E.; *et al.* Selection of Entropy Based Features for the Analysis of the Archimedes' Spiral Applied to Essential Tremor. In Proceedings of the 4th IEEE International Work Conference on Bioinspired, Intelligence, Donostia, Spain, 9–12 June 2015; pp. 157–162.
15. Ortega-Garcia, J.; Gonzalez-Rodriguez, J.; Simon-Zorita, D.; Cruz-Llanas, S. From Biometrics Technology to Applications Regarding Face, Voice, Signature and Fingerprint Recognition Systems. In *Biometric Solutions for Authentication in an E-World*; Kluwer Academic Publishers: Berlin, Germany, 2002; pp. 289–337.
16. The Montreal Cognitive Assessment (MoCA). Available online: <http://www.mocatest.org/> (accessed on 10 May 2016).
17. Faundez-Zanuy, M.; Sesa-Nogueras, E.; Roure-Alcobé, J.; Garré-Olmo, J.; López-de-Ipiña, K.; Solé-Casals, K. Online Drawings for Dementia Diagnose: In-Air and Pressure Information Analysis. In *XIII Mediterranean Conference on Medical and Biological Engineering and Computing*; Springer International Publishing: New York, NY, USA, 2014; pp. 567–570.
18. Neils-Strunjas, J.; Groves-Wright, K.; Mashima, P.; Harnish, S. Dysgraphia in Alzheimer's disease: A review for clinical and research purposes. *J. Speech Lang. Hear. Res.* **2006**, *49*, 1313–1330. [[CrossRef](#)]
19. Phillips, J.G.; Ogeil, R.P.; Müller, F. Alcohol consumption and handwriting: A kinematic analysis. *Hum. Mov. Sci.* **2009**, *28*, 619–632. [[CrossRef](#)] [[PubMed](#)]
20. Foley, R.G.; Miller, A.L. The effects of marijuana and alcohol usage on handwriting. *Forensic Sci. Int.* **1979**, *14*, 159–164. [[CrossRef](#)]
21. Tucha, O.; Walitza, S.; Mecklinger, L.; Stasik, D.; Sontag, T.; Lange, K.W. The effect of caffeine on handwriting movements in skilled writers. *Hum. Mov. Sci.* **2006**, *25*, 523–535. [[CrossRef](#)] [[PubMed](#)]
22. Rosenblum, S.; Parush, S.; Weiss, P.L. The in Air Phenomenon: Temporal and Spatial Correlates of the Handwriting Process. *Percept. Mot. Skills* **2003**, *96*, 933–954. [[CrossRef](#)] [[PubMed](#)]
23. Sadikov, A.; Groznic, V.; Žabkar, J.; Mozina, M.; Georgiev, D.; Pirtosek, Z.; Bratko, I. Parkinson Check smart phone app. In *Frontiers in Artificial Intelligence and Applications*; IOS Press: Amsterdam, The Netherlands, 2014; Volume 263, pp. 1213–1214.
24. Georgiev, D.; Groznic, V.; Sadikov, A.; Mozina, M.; Guid, M.; Kragelj, V.; Bratko, I.; Ribaric, S.; Pirtosek, Z. Digitalised spirography and clinical examination based decision support system of differentiating between tremors. *Eur. J. Neurol.* **2012**, *19*, 298.
25. Bolle, R.; Pankanti, S. *Biometrics, Personal Identification in Networked Society*; Jain, A., Ed.; Kluwer Academic Publishers: Norwell, MA, USA, 1998.
26. Faundez-Zanuy, M. Privacy issues on biometric systems. In *IEEE Aerospace and Electronic Systems Magazine*; IEEE Xplore: New York, NY, USA, 2005; Volume 20, pp. 13–15.
27. Gray, R.M. *Entropy and Information Theory*; Springer: Berlin/Heidelberg, Germany, 1990.
28. Brissaud, J.B. The meaning of entropy. *Entropy* **2005**, *7*, 68–96. [[CrossRef](#)]
29. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
30. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Champaign, IL, USA, 1949.
31. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, 2039–2049.
32. Pincus, S.M.; Huang, W.M. Approximate entropy, statistical properties and applications. *Commun. Stat. Theory Methods* **1992**, *21*, 3061–3077. [[CrossRef](#)]

33. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [[CrossRef](#)] [[PubMed](#)]
34. Dragomir, A.; Akay, Y.; Curran, A.K. Investigating the complexity of respiratory patterns during the laryngeal chemoreflex. *J. NeuroEng. Rehabil.* **2008**, *5*. [[CrossRef](#)] [[PubMed](#)]
35. Yentes, J.M.; Hunt, N.; Schmid, K.; Kaipust, J.P.; McGrath, D.; Stergiou, N. The use of approximate entropy and sample entropy with short data sets. *Ann. Biomed. Eng.* **2013**, *41*, 349–365. [[CrossRef](#)] [[PubMed](#)]
36. Ramanand, P.; Nampoori, V.P.; Sreenivasan, R. Complexity quantification of dense array EEG using sample entropy analysis. *J. Integr. Neurosci.* **2004**, *3*, 343–358. [[CrossRef](#)] [[PubMed](#)]
37. Solé-Casals, J.; Vialatte, F.B. Towards Semi-Automatic Artifact Rejection for the Improvement of Alzheimer's Disease Screening from EEG Signals. *Sensors* **2015**, *15*, 17963–17976. [[CrossRef](#)] [[PubMed](#)]
38. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*. [[CrossRef](#)] [[PubMed](#)]
39. Keller, K.; Lauffer, H. Symbolic analysis of high-dimensional time series. *Int. J. Bifurc. Chaos* **2003**, *13*, 2657–2668. [[CrossRef](#)]
40. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy. In *IEEE Signal Processing Letters*; IEEE Xplore: New York, NY, USA, 2012; Volume 19, pp. 91–95.
41. Mathworks. Available online: <http://www.mathworks.com> (accessed on 10 May 2016).
42. Weka. Available online: <http://www.cs.waikato.ac.nz/ml/weka> (accessed on 10 May 2016).
43. Picard, R.; Cook, D. Cross-Validation of Regression Models. *J. Am. Stat. Assoc.* **1984**, *79*, 575–583. [[CrossRef](#)]
44. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed.; Morgan Kaufmann: San Francisco, CA, USA, 2005.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).