XXXVIII Jornadas de Automática
https://doi.org/10.17979/spudc.9788497497749.0347

ASSESSMENT OF TREMOR SEVERITY IN PATIENTS WITH ESSENTIAL TREMOR USING SMARTWATCHES

M.A. Velasco[1], R. López-Blanco[2], J.P. Romero[3], M.D. del Castillo[1], J.I. Serrano[1], J. Benito-León[2], E. Rocon[1]

[1]Neural and Cognitive Engineering group, Centro de Automática y Robótica (CAR) CSIC-UPM, Ctra. Campo Real, km 0.2, 28500, Arganda del Rey, e.rocon@csic.es; [2]Instituto de Investigación (i+12), Hospital Universitario 12 de Octubre, Av. de Córdoba s/n, 28041, Madrid, jbenitol67@gmail.com; [3]Unidad de Daño Cerebral, Hospital Beata María Ana, Madrid, C. del Dr. Esquerdo, 83, 28007, Madrid, p.romero.prof@ufv.es

Abstract

This paper presents a classification model for the automatic quantification of tremor severity in patients with essential tremor (ET). The system is based on the signals measured by two commercial smartwatches that the patients wear on their wrist and ankle. The smartwatches register acceleration and angular velocity in these body segments. A set of nine tremor features were used to train the classification algorithm. The proposed algorithm is based on a C4.5 decision tree classifier. It is able to assess rest and kinetic (postural or action) tremor. The method was evaluated using data collected from thirty-four patients with ET. The algorithm classifies the severity of tremor in five levels 0-4 corresponding to those in the Fahn-Tolosa-Marin tremor rating scale with a 94% accuracy. The method can be implemented in a networked platform for the remote monitoring and assessment of movement disorders such as ET or Parkinson's disease.

Keywords: tremor assessment; time series classification; Essential Tremor.

1 INTRODUCTION

Essential tremor (ET) is considered the most prevalent type of tremor in adults, affecting ~ 5% of people over age 65 [1]. Bilateral postural with or without kinetic (in posture or action) tremor are the hallmarks of this entity, but mild rest tremor can be observed in some long-standing severe kinetic tremor [2]. Remarkably, 75% of the patients with ET report significant disability [3], consisting in relevant interference with employment, activities of daily living (ADL), and social function [4]. Improving tremor management in these patients could therefore drastically reduce direct and indirect costs related to the disease. It could also improve the quality of life and independence of both patients and caregivers.

Many objective transducer-based measures, such as electromyography (EMG), vocal acoustic analysis, accelerometers, or gyroscopes, have been used for the quantification and characterization of tremor [5]. They proved to be more sensitive than clinical rating scores to changes in tremor amplitude and frequency in specific scenarios [6]. The miniaturization of inertial measurement units (IMU) makes wearable technology closer to be ready for the clinical practice and long-term ambulatory tremor monitoring [7], [8]. However, new machine learning algorithms are needed to translate the high-dimensional data provided by wearables into clinically meaningful information [9].

Kubota et al. reviewed [9] unsupervised (clustering) and supervised machine learning algorithms such as linear regression, neural networks classifiers, support vector machines, k-nearest neighbors, naïve Bayes, or decision trees for measuring tremor symptoms in Parkinson's disease [9]. Whereas these methods can achieve very high values of accuracy, an incorrect training of the classifier can lead to errors in the prediction of new tremor episodes due to the overfitting of the trained model. Furthermore, these algorithms are trained with the assumption that the distribution of the training data is static and unchanging. Hence, the algorithms must be re-trained periodically in order to prevent invalid predictions. On the other hand, many of the reviewed algorithms focus on the detection of tremor and rely on further analysis for the quantification of severity [10].

In this paper, we propose an automated method for the classification of tremor severity in patients with ET. The methodology is based on the analysis of the signals registered by accelerometers and gyroscopes during standard clinical tasks to assess rest, postural and action tremor. A simple set of 9 features and a C4.5 decision tree classifier can be used to build a model that can be used later for the online classification of tremor in ambulatory monitoring applications. The automatic detection and characterization of tremor are two of the main goals of the NetMD Project[1].

2 METHODOLOGY

2.1 PARTICIPANTS

Thirty-four patients (ages 18-81) with ET were recruited by the Neurology Department of the University Hospital “12 de Octubre” in Madrid. Eighteen of them were taking medication for their disease during the tests. Table 1 depicts the main clinical features of the participants in the study.

Table 1: Demographic and clinical data of patients with essential tremor (N = 34).

Age	
Mean ± SD	64 ± 14.4
Gender	
Female Male	14 (41.1%) 20 (59.9%)
Disease duration (years)	
Mean ± SD	12.5 ± 10.4
Global FTM-TRS score at recruitment	27.9 ±12.6
FTM-TRS-A	9.1 ± 4.9
FTM-TRS-B	13.5 ± 6.4
FTM-TRS-C	5.3 ± 2.6

2.2 APPARATUS

An expert neurologist in movement disorders examined the patients. He used the Fahn-Tolosa-Marin (FTM) tremor rating scale (TRS) to assign a score to several items measured during a clinical examination: rest tremor, postural tremor, and action tremor. All the patients followed a specific protocol, which includes the following tasks:

A. Measurement of rest tremor. Sitting on a chair with their hands resting in their lap. Count from 100 to 0.

B. Measurement of kinetic (postural) tremor. Holding the arms outstretched with the hands in pronation.

C. Measurement of kinetic (action) tremor. Finger to nose movements starting and ending with the arms outstretched to the sides.

D. Measurement of kinetic (action) tremor. Pouring water between two glasses starting and ending with the arms resting.

These items are enumerated in Table 2 and correspond to the FTM-TRS part A and B [11]. The whole session was videotaped. The neurologist assigned a score to each task and patient after a thorough examination of the video.

Table 2: Fahn-Tolosa-Marin (FTM) tremor rating scale (TRS) score pattern for the tasks A-D.

FTM-TRS-A: Items 5 or 6	
Tasks A,B,C: rest and kinetic tremor	0: None 1: Slight. May be intermittent. 2: Moderate. Intermittent (< 2 cm). 3: Marked amplitude (2-4 cm). 4: Severe amplitude (> 4 cm).
FTM-TRS-B: Item 14	
Task D: Pouring water from a glass	0: Normal 1: Slow, but no water is spilled. 2: Spills 10% of water. 3: Spills ~50% of water. 4: Unable to complete the task.

During the experiments, the patients wore two Sony Smartwatch3 located on the wrist ($SW3_w$) and ankle ($SW3_a$) of the most affected hemibody. They also carried an Android Smartphone ASUS inside of a belt-pouch on the waist. An ad-hoc Android application acquired raw data obtained from the gyroscopes and the accelerometers at a sampling frequency of 50Hz. The Smartphone stored a timestamp and angular velocity and linear acceleration in three axis in a text file (txt). See Figure 1.

Figure 1: System of reference of the SW3.

The experiments were approved by the ethical standards committees on human experimentation at the University Hospital ‘‘12 de Octubre’’ (Madrid). The participants read and signed informed consent prior to the tests.

[1] http://g-nec.com/project_NetMD.html

2.3 CLASSIFICATION ALGORITHM

2.3.1 Signal preprocessing

The txt files were processed on a 2.83 GHz Inter Core 2 Quad Q9500 machine running Windows 7 Professional 32-bit. The preprocessing of the signal was done off-line using Matlab software (version 7.11.0 (R2010b); MathWorks, Natick, MA). A total of 136 segments (4 tasks x 34 patients) were annotated by the neurologists in terms of clinical task (A-D) and severity (0-4). The durations (mean ± SD) of the segments corresponding to the tasks A-D are:

- Task A (rest tremor): 20.1 ± 11.4 s
- Task B (kinetic tremor): 23.3 ± 20 s
- Task C (kinetic tremor): 12.8 ± 3.5 s
- Task D (kinetic tremor): 54.8 ± 18 s

Figures 2-4 illustrate three examples of the gyroscope signal measured during the tasks which assessed kinetic tremor.

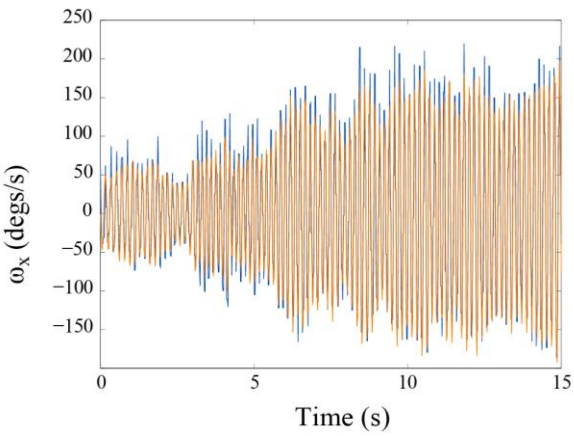

Figure 2: Gyroscope signal measured in the x-axis of the $SW3_w$. The blue curve represents the velocity of the wrist prono-supination in task B; the orange, the band-pass filtered signal.

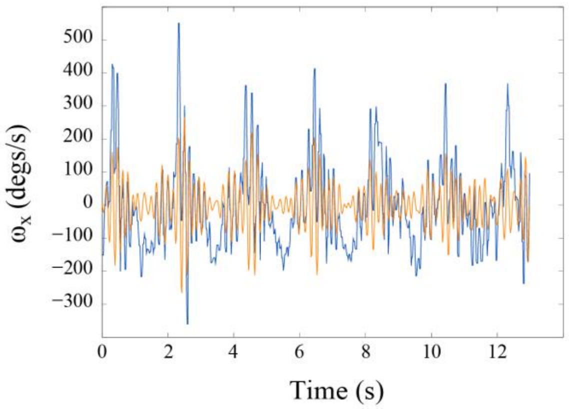

Figure 3: Seven repetitions of the finger-to-nose task (Task C). Raw and filtered wrist prono-supination.

More precisely, Figure 2 shows postural tremor, Figure 3 depicts action tremor in the finger-to-nose task, and Figure 4 exemplifies action tremor in the pouring task.

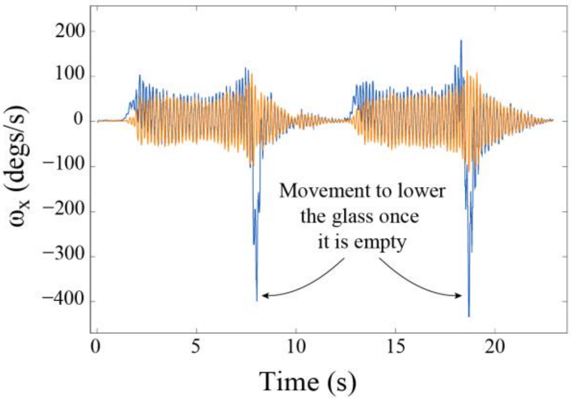

Figure 4: Two repetitions of the pouring task (Task D). Raw (blue) and filtered (orange) wrist prono-supination

The segmented signals were band-pass filtered using a 10th order Butterworth high-pass filter ($freq_1$>3 Hz) which removed the voluntary component, followed by a 10th order Butterworth low-pass filter ($freq_2$<12 Hz) which eliminated tremors of higher frequency.

2.3.2 Feature extraction

Nine tremor features were estimated from the recordings of the smartwatches as determined in [10]. They are enumerated in Table 3. The data registered by the $SW3_w$ in the wrist was used to estimate the features f1-f6, f8, and f9. Note that feature f4 was estimated from the 3 Hz low-passed gyroscope signal. On the other hand, the feature f7, "other body segment energy" was calculated from the data registered by $SW3_a$, located in the ankle. The extraction of the features is based on a 3 s moving window with 0.1 s overlapping.

Table 3: Features for tremor recognition

Feature	Id	Sensor
Dominant Frequency	f1	Gyro
Energy on Dominant Frequency	f2	Gyro
High Frequency Energy	f3	Gyro
Low Frequency Energy	f4	Gyro
Spectrum Entropy	f5	Gyro
Mechanical Energy	f6	Acce
Other Body Segment Energy	f7	Gyro
Ratio Hi/Lo Frequency	f8	Gyro
f1*f2	f9	Gyro

Gyro = gyroscopes; Acce = accelerometers

2.3.2 Tremor classification

We used the Weka [12] collection of machine learning algorithms for the training of a C4.5 decision tree classifier and the evaluation of our set of tremor features.

The method proposed for the selection of the features was the wrapper approach. This method takes into account the classifier chosen. It also uses the best-first search algorithm. The wrapper approach used two classifiers: the C4.5 decision tree classifier and the naïve Bayes classifier consecutively. Additionally, a 10-fold cross-validation procedure was employed to achieve a more robust evaluation.

We also computed a 10-fold cross-validation during the training of the classifiers in order to avoid overfitting [13].

3 RESULTS

A total of 32745 sets of features were analyzed. The patients recruited showed mostly episodes of mild to moderate tremor. The FTM-TRS scores in the instances analyzed is depicted in Figure 5.

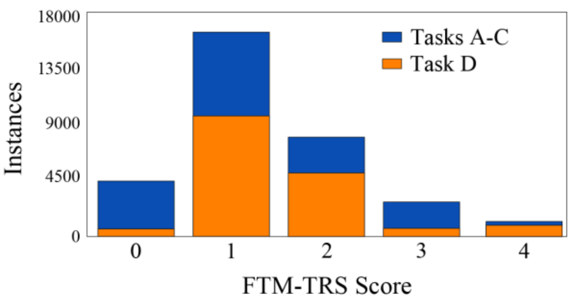

Figure 5: Distribution of instances corresponding to each FTM-TRS level assigned by the neurologist.

3.1 SELECTION OF FEATURES

With the C4.5 decision tree, the most selected features were the dominant frequency (f1), the energy at the dominant frequency (f2), the energy at low frequencies (f4), the spectrum entropy (f5), the mechanical energy (f6), the energy measured in $SW3_a$ (f7), and the ratio of high/low energy (f8). The least selected features were the energy at high frequencies (f3) and f9. On the other hand, the naïve Bayes classifier selected f1, f5, and f8.

3.2 CLASSIFICATION MODELS

Two tremor severity classifiers were built: TC1 and TC2. TC1 was trained with the 9 features. The features f3 and f9, which were discarded by the wrapper approach, are not included in the training and evaluation of the tremor classifier TC2.

The decision tree TC1 was built in 4.9 s. It had 1050 leaves and showed a classification error of 5.85 %. The range of error with a confident interval (CI) of 95 % was 5.55-6.09 %. The areas under ROC for the classes 0 to 4 were 0.969, 0.965, 0.968, 0.976, and 0.974, respectively.

TC2 had 1040 leaves and was able to only classify incorrectly 1804 instances (5.51%), CI at 95 %, 5.25-5.77 %. The time taken to build it with Weka was 3.33 s. In this case, the areas under ROC for the classes 0 to 4 were 0.971, 0.966, 0.969, 0.977, and 0.973. Table 4 depicts the confusion matrix estimated for TC2.

Table 4: Confusion matrix of the C4.5 decision tree classifier TC2 for tremor in ET.

TC2	**Classified as**				
Class	**0**	**1**	**2**	**3**	**4**
0	**4153**	303	17	2	1
1	232	**15701**	348	69	23
2	22	406	**7414**	79	51
3	1	53	92	**2623**	9
4	0	17	58	21	**1050**

We could not achieve a significant improvement of the classification errors after not including f3 and f9 in TC2. However, we improved the time needed to build the model by 32 %.

4 DISCUSSION

In this paper we proposed an algorithm for the automatic assessment of tremor severity in patients with ET. The algorithm is based on the analysis of the signals recorded by the accelerometers and gyroscopes which are integrated in two smartwatches that the patients wear in their wrist and ankle. These signals were preprocessed and characterized by a set of nine features. These extracted features were then used to train two C4.5 decision tree classifiers, TC1 and TC2. The system was validated with thirty-four patients with ET recruited at the University Hospital “12 de Octubre”.

The results show that the classifiers are able to identify the tremor severity among 5 levels of FMT-TRS score. Even though the accuracy of TC1 and TC2 is very similar, we were able to reduce the time needed to build the model in Weka by 32%.

The simplicity of the model makes it very easy to implement in different platforms. Additionally, the smartwatches used for the measurement of tremor showed good wearability and an affordable price. These characteristics make the proposed platform a very interesting solution for the continuous and objective ambulatory monitoring of tremor and other movement disorders.

4.1 EXPERIMENTAL CONCERNS

There are some limitations in the experiment that could affect the results and the reproducibility of the study. The neurologist was blinded to the signals registered by the smartwatches. However, he had treated the patients during the recruitment phase. Hence, his evaluation of the videotapes could be biased. In future studies, a second neurologist with expertise in pathological tremor and movement disorders will be included in the analysis. The agreement between raters will be assessed with a correlation Cohen's kappa coefficient. Additionally, the location of the smartwatch in the distal forearm could affect the registry of tremor. The amplitude of the tremorous signal increases distally and it is maximum in the hand and fingers. Consequently, our smartwatch could be unable to register significant components of distal tremor on these body segments.

Our results reveal a slightly higher number of classification errors between the classes 1 and 2. This could be due to the imbalanced data from these classes, but also to the nature of the FTM tremor rating scale. In the case of the levels 1 "slight, intermittent" and 2 "moderate, intermittent", the rating provided by the neurologists can be extremely subjective. Other well-known clinical scales such as The Essential Tremor Rating Assessment Scale (TETRAS) rate tremor 0-4 in half-point intervals [14]. We will be able to reduce the variability of the rating and improve the accuracy of our tremor classifiers if we use this clinical scale in the future.

4.2 FUTURE WORKS

In future studies, we will improve our tremor classifier by including several neurologists to achieve a more precise rating of the tremor segments. We will also evaluate the performance of other classifiers such as Hidden Markov Models, Support Vector Machines, or k-Nearest Neighbors and extract new tremor features. The model will be implemented in the NetMD online platform. The platform will register the movements of the patient during a 24-hour monitoring and will provide the neurologists, patients and caregivers with an objective full tremor report. More information on the NetMD Project can be found at the link provided in Section 1.

In addition to this, we are interested in introducing new context awareness to our assessment of tremor. More specifically, we will train new classifiers to identify several activities of the daily living. We will detect periods of rest and physical activity, including gait. Tasks of special interest are those related to dressing (putting a shirt on and buttoning it), grooming (combing hair or brushing teeth) or feeding (fine movements and gross movements with a spoon or a fork and a knife). A similar method was presented in [15] using signals registered by four IMUs located in the hand, forearm and arm of patients with ET.

With all this new information, the neurologists will be able to monitor and characterize not only episodes of tremor but also to identify the specific activity that the patient was performing when an onset of tremor was detected. This will be very valuable information to assess the evolution of the disease. It can also help them to detect the possible side effects that the medication can have on the daily activities of the patients and the tremor in their upper and lower limbs.

5 CONCLUSION

In this paper, we used nine features to train an automatic tremor severity classifier. The proposed method showed good accuracy in the identification of tremor severity in signals of gyroscopes and accelerometers. The method was able to assess the severity of tremor in 34 patients with ET. The system classified tremor instances as five levels of severity (0, 1, 2, 3, and 4), corresponding to the levels described by Fahn-Tolosa-Marin tremor rating scale.

The method can be implemented in a networked platform for the remote monitoring and assessment of movement disorders.

Acknowledgements

We would like to thank A. Clemotte and J.A. Gallego for their comments on the methodology. This work was possible thank to the projects NetMD (RTC-2015-3967-1), NeuroMOD (DPI2015-68664-C4-1-R), and InterAAC (RTC-2015-4327-1). They are all financed by the Spanish Ministry of Economy, Industry and Competitiveness.

References

[1] J. Benito-Leon and E. D. Louis, "Essential tremor: emerging views of a common disorder.," *Nat. Clin. Pract. Neurol.*, vol. 2, no. 12, p. 666–78; quiz 2p following 691, Dec. 2006.

[2] G. Deuschl, P. Bain, M. Brin, Y. Agid, L. Benabid, R. Benecke, A. Berardelli, D. J. Brooks, R. Elble, S. Fahn, L. J. Findley, M.

Hallett, J. Jankovic, W. C. Koller, P. Krack, A. E. Lang, A. Lees, C. H. Lucking, C. D. Marsden, J. A. Obeso, W. H. Oertel, W. Poewe, P. Pollak, N. Quinn, J. C. Rothwell, H. Shibasaki, P. Thompson, and E. Tolosa, "Consensus Statement of the Movement Disorder Society on Tremor," *Mov. Disord.*, vol. 13, no. S3, pp. 2–23, 1998.

[3] K. L. Busenbark, J. Nash, J. P. Hubble, and W. C. Koller, "Is essential tremor benign?," *Neurology*, vol. 41, pp. 1982–3, 1991.

[4] E. D. Louis, L. Barnes, S. M. Albert, L. Cote, F. R. Schneider, S. L. Pullman, and Q. Yu, "Correlates of Functional Disability in Essential Tremor," *Mov. Disord.*, vol. 16, no. 5, pp. 914–920, 2001.

[5] R. J. Elble and J. McNames, "Using Portable Transducers to Measure Tremor Severity.," *Tremor Other Hyperkinet. Mov. (N. Y).*, vol. 6, p. 375, 2016.

[6] D. Haubenberger, G. Abbruzzese, P. G. Bain, N. Bajaj, J. Benito-León, K. P. Bhatia, G. Deuschl, M. J. Forjaz, M. Hallett, E. D. Louis, K. E. Lyons, T. A. Mestre, J. Raethjen, M. Stamelou, E. K. Tan, C. M. Testa, and R. J. Elble, "Transducer-based evaluation of tremor," *Mov. Disord.*, vol. 31, no. 9, pp. 1327–1336, 2016.

[7] E. R. de Lima, A. O. Andrade, J. L. Pons, P. Kyberd, and S. J. Nasuto, "Empirical mode decomposition: a novel technique for the study of tremor time series.," *Med. Biol. Eng. Comput.*, vol. 44, no. 7, pp. 569–82, Jul. 2006.

[8] G. Mostile, J. P. Giuffrida, O. R. Adam, A. Davidson, and J. Jankovic, "Correlation between Kinesia system assessments and clinical tremor scores in patients with essential tremor," *Mov. Disord.*, vol. 25, no. 12, pp. 1938–1943, Sep. 2010.

[9] K. J. Kubota, J. A. Chen, and M. A. Little, "Machine learning for large-scale wearable sensor data in Parkinson's disease: Concepts, promises, pitfalls, and futures," *Mov. Disord.*, vol. 31, no. 9, pp. 1314–1326, 2016.

[10] G. Rigas, A. T. Tzallas, M. G. Tsipouras, P. Bougia, E. E. Tripoliti, D. Baga, D. I. Fotiadis, S. Member, S. G. Tsouli, and S. Konitsiotis, "Assessment of Tremor Activity in the Parkinson ' s Disease Using a Set of Wearable Sensors," vol. 16, no. 3, pp. 478–487, 2012.

[11] C. M. Fahn S, Tolosa E, *Clinical rating scale for tremor*, 2nd ed. Baltimore: Williams & Wilkins, 1993.

[12] Hall. M, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA Data Mining Software: An Update," *SIGKDD Explor.*, vol. 11, no. 1, 2009.

[13] S. Lambrecht, J. A. Gallego, E. Rocon, and J. L. Pons, "Automatic real-time monitoring and assessment of tremor parameters in the upper limb from orientation data," *Front. Neurosci.*, vol. 8, no. July, p. 221, Jul. 2014.

[14] R. J. Elble, "The Essential Tremor Rating Assessment Scale," *J Neurol Neuromed*, vol. 1, no. 4, pp. 34–38, 2016.

[15] J. I. Serrano, S. Lambrecht, M. Dolores del Castillo, J. P. Romero, J. Benito-León, and E. Rocon, "Identification of activities of daily living in tremorous patients using inertial sensors," *Expert Syst. Appl.*, Apr. 2017.