



HAL
open science

New insight in spiral drawing analysis methods – Application to action tremor quantification

André Pierre Legrand, Isabelle Rivals, Aliénor Richard, Emmanuelle Apartis,
Emmanuel Roze, Marie Vidailhet, Sabine Meunier, Elodie Hainque

► **To cite this version:**

André Pierre Legrand, Isabelle Rivals, Aliénor Richard, Emmanuelle Apartis, Emmanuel Roze, et al..
New insight in spiral drawing analysis methods – Application to action tremor quantification. *Clinical
Neurophysiology*, 2017, 128 (10), pp.1823-1834. 10.1016/j.clinph.2017.07.002 . hal-04018358

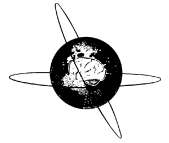
HAL Id: hal-04018358

<https://hal.science/hal-04018358>

Submitted on 7 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



New insight in spiral drawing analysis methods – Application to action tremor quantification



André Pierre Legrand^{a,*}, Isabelle Rivals^b, Aliénor Richard^c, Emmanuelle Apartis^{c,d}, Emmanuel Roze^{c,e}, Marie Vidailhet^{c,e}, Sabine Meunier^c, Elodie Hainque^{c,e}

^a ESPCI Paris, PSL Research University, Paris, France

^b Equipe de Statistique Appliquée, ESPCI Paris, PSL Research University, UMR 1158, Paris, France

^c Inserm U 1127, CNRS UMR 7225, Université de la Sorbonne, UPMC Univ Paris 06 UMR S 1127, Institut du Cerveau et de la Moelle Epinière, F-75013, Paris, France

^d Département de Physiologie, Hôpital Saint-Antoine, AP-HP, Paris, France

^e Département de Neurologie, Hôpital Pitié-Salpêtrière, AP-HP, Paris, France

ARTICLE INFO

Article history:

Accepted 1 July 2017

Available online 17 July 2017

Keywords:

Spiral drawing analysis

Essential tremor

Unravelling

Empirical mode decomposition

HIGHLIGHTS

- New methods for computerized spiral analysis are proposed, which outperform the usual velocity method.
- They provide new insight into tremor amplitude and frequency variations.
- They were implemented on a computer via a digitized tablet.

ABSTRACT

Objective: Spiral drawing is one of the standard tests used to assess tremor severity for the clinical evaluation of medical treatments. Tremor severity is estimated through visual rating of the drawings by movement disorders experts. Different approaches based on the mathematical signal analysis of the recorded spiral drawings were proposed to replace this rater dependent estimate. The objective of the present study is to propose new numerical methods and to evaluate them in terms of agreement with visual rating and reproducibility.

Methods: Series of spiral drawings of patients with essential tremor were visually rated by a board of experts. In addition to the usual velocity analysis, three new numerical methods were tested and compared, namely static and dynamic unraveling, and empirical mode decomposition. The reproducibility of both visual and numerical ratings was estimated, and their agreement was evaluated.

Results: The statistical analysis demonstrated excellent agreement between visual and numerical ratings, and more reproducible results with numerical methods than with visual ratings.

Conclusions: The velocity method and the new numerical methods are in good agreement. Among the latter, static and dynamic unravelling both display a smaller dispersion and are easier for automatic analysis.

Significance: The reliable scores obtained through the proposed numerical methods allow considering that their implementation on a digitized tablet, be it connected with a computer or independent, provides an efficient automatic tool for tremor severity assessment.

© 2017 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Hand drawing and/or writing analyses have been used for a long time to document tremor severity. The Fahn-Tolosa-Marin

(FTM) tremor rating scale (1993), dedicated to essential tremor (ET), integrates visual ratings of hand drawn spirals on a (0–4) point scale (score B). To improve precision and reliability, Bain et al. (1993) provided a broader gradation with an (0–10) point clinical tremor rating scale. These widely used visual scales are subject to intra and interrater variability (Stacy et al. 2007; Haubenberger et al. 2011). Assessment of tremor amplitude by visual inspection may hence be delicate, while signal analysis can

* Corresponding author at: ESPCI, 10 rue Vauquelin 75231 Paris CEDEX 05, France.

E-mail address: andre-pierre.legrand@espci.fr (A.P. Legrand).

provide a reliable measure of it (Miralles et al. 2006; Kraus and Hoffmann, 2010). The Bain rating scale was validated by Kraus and Hoffmann, (2010), except for a few examples of spirals.

Digitising tablets have been increasingly used to automate the quantification of hand drawn spirals and writing capabilities (Lacquaniti et al. 1987; Elble et al. 1990; Marquardt and Mai 1994; Elble et al. 1996; Pullman 1998; Mergl et al. 1999; Erasmus et al., 2001; Liu et al. 2005; Ondo et al. 2005; Wang et al. 2005; Feys et al. 2007; Zeuner et al. 2007; Hsu et al. 2009; Viviani et al., 2009; Schneider et al. 2010; Bajaj et al., 2011; Haubenberger et al. 2011; Louis et al. 2012; Michalec et al., 2014; Hess et al. 2014; Sisti et al. 2017). In these studies, several parameters were used to characterize the tremor. The most frequently used parameter is the Fast Fourier Transform (FFT) of the pen-tip velocity. As the objective is to extract information about the intensity on a frequency domain of interest, the signal needs to be filtered. This may be done through derivation or filtering in a frequency range from one or three to ten Hz. Assuming the stability of the tremor frequency during the record, the FFT provides an estimate of the intensity or amplitude of the tremor by integration on a short frequency domain around the spectral peak (depending on the studies). Such a parameter was shown to correlate with the FTM or Bain scores attributed by raters, and to be less dispersed than these ratings.

In this paper, we propose three new computerized rating methods which compete with the velocity analysis methods. Such methods are compared for reliability in view of their implementation in automatic software analysis. We then correlate the quantitative measures of tremor derived from these methods with the scores from the visual ratings of experts on a series of hand drawn spirals.

2. Methods

2.1. Patients

Thirteen patients diagnosed with essential tremor (mean age: 54.5 ± 19 years, range 23–74) and little or no improvement with pharmacological treatment were included in the study (Table 1). Diagnosis of ET was established on the basis of the *Movement disorders consensus criteria* (Deuschl et al., 1998). Patients had bilateral postural/action tremor for more than 1 year. Two patients were treated for their tremor with beta blockers, and kept the doses unchanged throughout the study. The other patients did not receive any anti tremoring treatment. The patients did not receive benzodiazepines, anti-epileptic or antidepressant medication and were requested to avoid alcohol during the whole course of the study. Only one patient was taking medications that could affect tremor; this patient (#10) received levothyroxine for hypothyroidism. No restriction of caffeine was planned during the procedure.

The experimental procedures were approved by the local ethics committee “Ile de France VI – Groupe Hospitalier Pitié-Salpêtrière” and conformed to the Declaration of Helsinki. All the subjects gave their written informed consent before any procedure.

2.2. Spiral drawing and collection

A digitising tablet (Wacom, Bamboo Fun Medium Pen&Touch; sampling rate 133 Hz, precision ± 0.25 mm) was used to record the drawings. The template is a 6-turn clockwise Archimedean spiral, with an inter-loop of 9 mm printed on a paper centred on the digitizing tablet. Patients had 2 visits with a 3 month gap between them. At each visit, they drew 4 spirals with their dominant hand and the drawing arm unsupported, using a Wacom pen. Before the

patients start to draw, the centring of the spiral using the Wacom pen is done by the operator. For two of the spirals, they were instructed to keep the pen “between” the template’s printed lines, and for the remaining two, to draw “on” the printed lines, in both cases from the inside to the outside, with no limitation of time for drawing. Owing to the tablet used here, patients were not able to visualize their actual drawing during the record. Due to some technical issues and/or bad patients’ compliance to the experiment, 20 spirals had to be withdrawn from the data set, leaving 92 spirals (45 “between”, 47 “on”) for analysis.

2.3. Visual rating

The 92 digitized spiral drawings were printed and anonymized. Spiral drawings were visually rated off line by 7 independent blinded raters¹ twice at a few weeks interval, using Bain rating scale, the scores of which range from 0 (no tremor) to 10 (tremor too severe for a spiral to be recognisable as such).

2.4. Computerized methods of analysis

The digitised spiral drawings included air points, i.e. X, Y coordinates of the pen-tip at less than 16 mm from the tablet surface. When higher distance values appeared (a rare event) during a short time an interpolation between the last and next contacts was done, and some erratic initial or final contacts were excluded. After these pre-processing steps, the 92 spiral drawings were analysed using 4 different methods. In order to enable comparisons between them, the same drawing (a spiral drawing “on” the template) is used all along the next figures.

2.4.1. Velocity based method

Several authors (Elble et al., 1990, 1996; Mergl et al., 1999) used the spectra of the X and Y velocities to estimate the intensity around the tremor frequency in amplitude or power. Others (Wang et al. 2005; Feys et al. 2007) used either the intensity of the main peak or the root mean square of the velocity after a window filtering. (Haubenberger et al., 2011) preferred the area under the curve of a 2 Hz-window around the spectral peak. Here, the spectra of the X and Y velocities were fitted with a Gaussian curve centred on the action tremor frequency, without filtering. To take the anisotropy of the drawing into account, the X and Y areas under the curves were averaged to obtain the characteristic amplitude of the tremor. Spectral tremor peak frequency (Hz), amplitude (mm/s) and Gaussian fits are shown on Fig. 1. The corresponding descriptor will be referred as “V” in the following. The variability of the method depends on the operator’s Gaussian fits. To evaluate the reproducibility of the V values, two couples (X and Y) of Gaussian fits were performed.

In addition to this baseline method, we propose three new methods.

2.4.2. Static and dynamic unravelling based methods

Unravelling consists in the conversion of a spiral drawing from Cartesian to polar coordinates and was introduced by Pullman (1998) and Saunders-Pullmann et al. (2008). It was shown (Wang et al. 2008) that an inaccurate centring of the spiral results in an oscillation around the polar line, and that was also observed in our records, see Fig. 2. It can be interpreted as an involuntary positioning error made by the operator and/or patient when handling the Wacom pen. This was observed for “on” as well as for “between” template spirals. In order to analyse the fluctuations of

¹ Neurologists and movement disorder specialists experienced in the collection of Archimedean spirals.

Table 1
Clinical profile of the patients.

Patient	Sex	Age	Disease duration (years)	FMT Score Total [/156]	FMT Score Part A [/88]	FMT Score Part B [/36]	FMT Score Part C [/32]	Familial History	Treatment with beta blockers
1	M	71.6	4.2	34	11	16	7	na	no
2	M	44.3	35.2	28	8	14	6	yes	no
3	M	64.8	52.3	43	13	19	11	no	no
4	M	23.4	8.3	30	7	17	6	yes	no
5	M	70.3	16.4	59	15	22	22	yes	yes
6	F	63.1	28.9	26	9	11	6	yes	no
7	M	22.7	8.0	25	9	12	4	yes	no
8	M	74.4	23.0	34	5	18	11	no	no
9	F	72.4	20.1	51	16	20	15	no	no
10	M	38.1	24.1	38	6	23	9	yes	no
11	M	52.3	34.2	30	11	13	6	yes	no
12	F	71.5	8.2	42	8	23	11	no	no
13	F	39.1	30.0	41	9	18	14	yes	yes
Mean		54.5	21.9	37.0	9.8	17.4	9.8		
(SD)		(19.0)	(14.2)	(10.1)	(3.3)	(4.1)	(5.0)		

Abbreviations: FMT = Fahn-Tolosa-Marin rating scale; SD = standard deviation; na = not available; M = male; F = female.

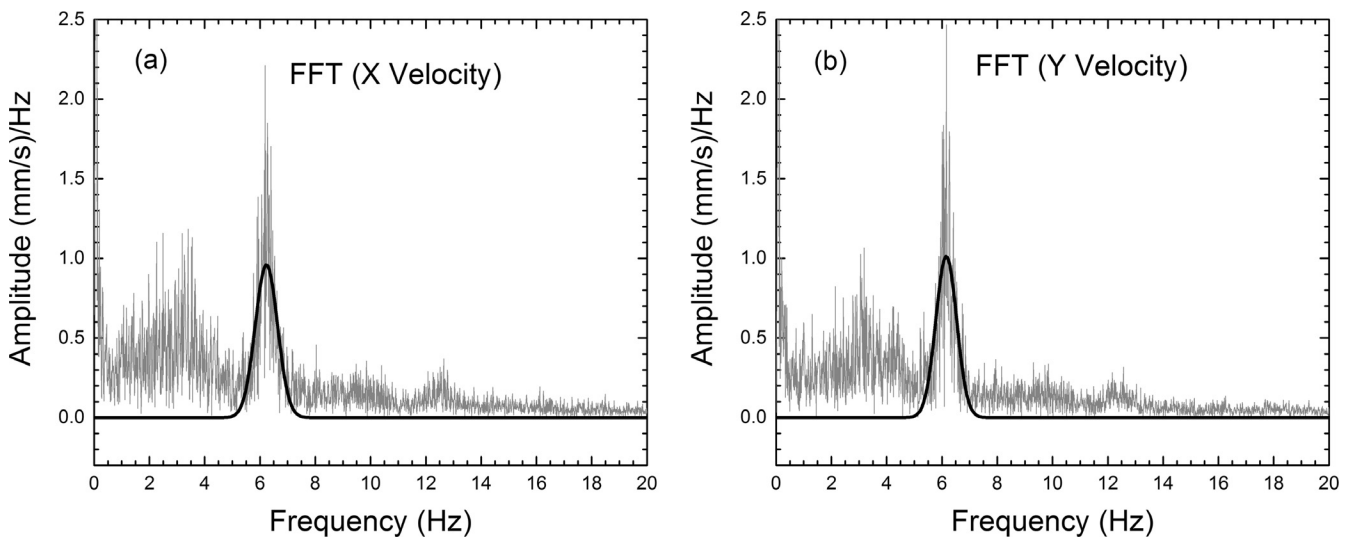


Fig. 1. X (a) and Y (b) velocities and their spectra obtained by Gaussian fit around tremor frequency. The drawing was performed “on” the template spiral. The areas under the two curves are averaged to obtain the V descriptor value reflecting the tremor intensity.

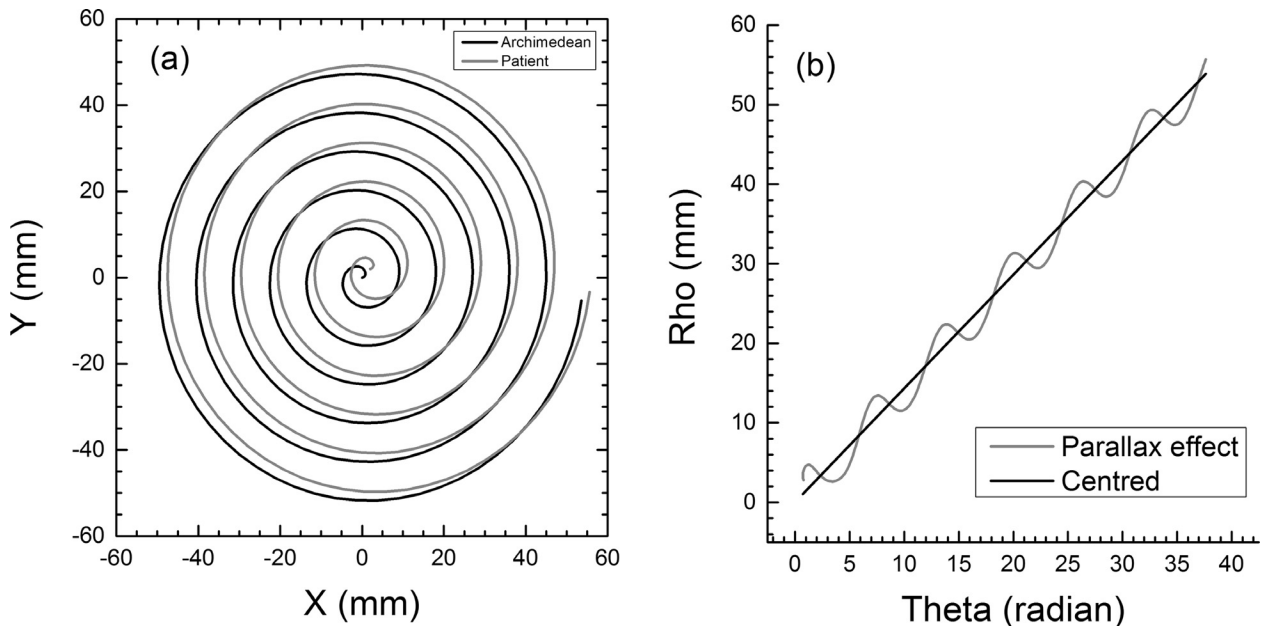


Fig. 2. (a) Ideal Archimedean spiral centred at (0, 0) (black) and offset (2, 2 mm) (grey); (b) Polar transform for centred (black) and offset (grey).

the drawing around the template spiral, this positioning or parallax error has to be taken into account.

As shown on Fig. 3, the focused point ($\rho = a\theta$) is different from the recorded point (ρ', θ') with a parallax (α, β) assumed not to change along the drawing. Consequently:

$$\rho'^2 = a^2 \cdot \theta^2 + (\alpha^2 + \beta^2) + 2 \cdot \alpha \cdot a \cdot \theta \cdot \cos(\theta) + 2 \cdot \beta \cdot a \cdot \theta \cdot \sin(\theta)$$

which can be approximated as:

$$\rho' \approx a \cdot \theta + \alpha \cdot \cos(\theta) + \beta \cdot \sin(\theta) \tag{1}$$

assuming that:

$$a^2 \cdot \theta^2 \gg (\alpha^2 + \beta^2)$$

Eq. (1) can be used to fit the unravelled polar plot (Fig. 4), leaving the residual fluctuation for analysis. Two new descriptors were

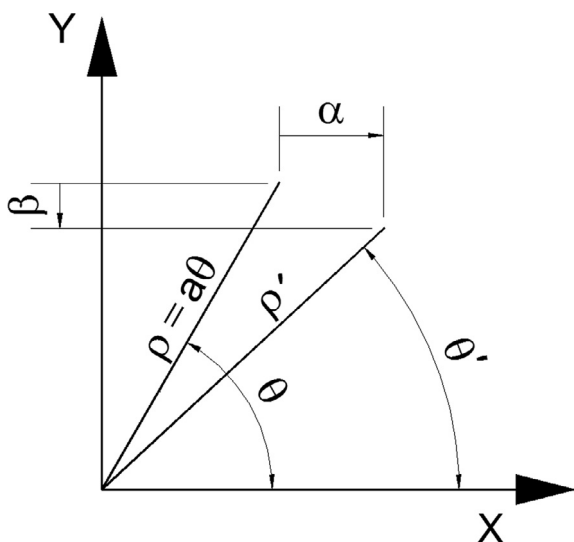


Fig. 3. Parallax effect on the recorded point.

derived from this residual in order to characterize the amplitude of the tremor:

- a. *First descriptor: The mean square deviation of the residual:* This descriptor is referred to as “URS” (URS for Unravelling Static) in the following. Note that a similar characterisation was done using the velocity instead of the above residual in Feys et al. (2007). Operator-dependent choice of the start- and end-points of the spirals is the main source of variability. To assess the reproducibility of the URS value, three different choices of 0.25 radian start- and end-angle cuts were tested.
- b. *Second descriptor: The area under a Gaussian curve fitted to the FFT of the residual around the typical tremor frequency,* which is the same as the one observed on accelerometric or electromyographic records, see Fig. 5. This descriptor is referred to as “URD” (URD for Unravelling Dynamic) in the following. Choice of the start- and end-points of the spirals, and of the Gaussian fits participate both in the variability of the results. In order to quantify the variability of both sources, two different fits were performed for the three choices of the start- and end-points described for the URS values.

2.4.3. Empirical Mode Decomposition (EMD) based method

This method was first used to analyse gyroscope displacement during physiological movements, in particular for spiral drawing follow up (de Lima et al. 2006). It was also applied to study effects of fatigue on electromyographic signals (Xie and Wang 2006).

Whereas the velocity method as well as the two previously introduced unravelling methods assumed a stationary state all along the time window, EMD is suited for the nonlinear analysis of non-stationary signals. EMD can be viewed as a sifting process, decomposing the signal into its different components i.e. the Intrinsic Mode Functions (IMFs), from the faster to the slower. These components can be analysed as function of frequency, amplitude and time. The goal is here is to split signals linked to fast motion from those linked to slow motion, i.e. the tremor from the voluntary drawing (de Lima et al. 2006). We performed EMD using the software developed by Flandrin et al. (<http://perso.ens-lyon.fr>).

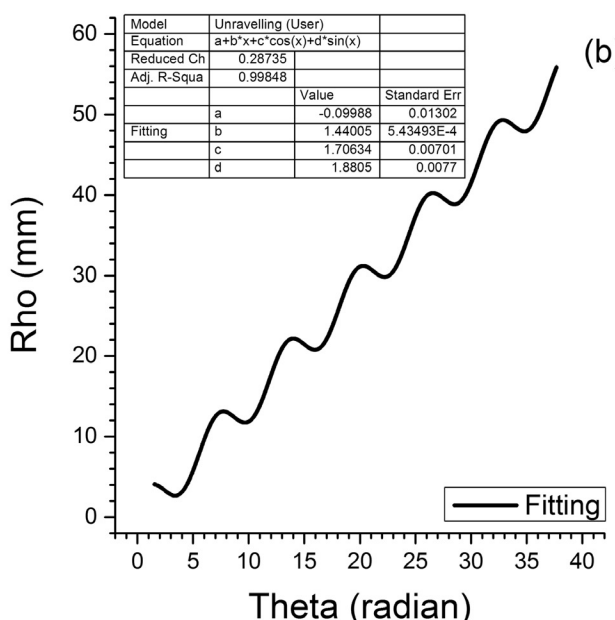
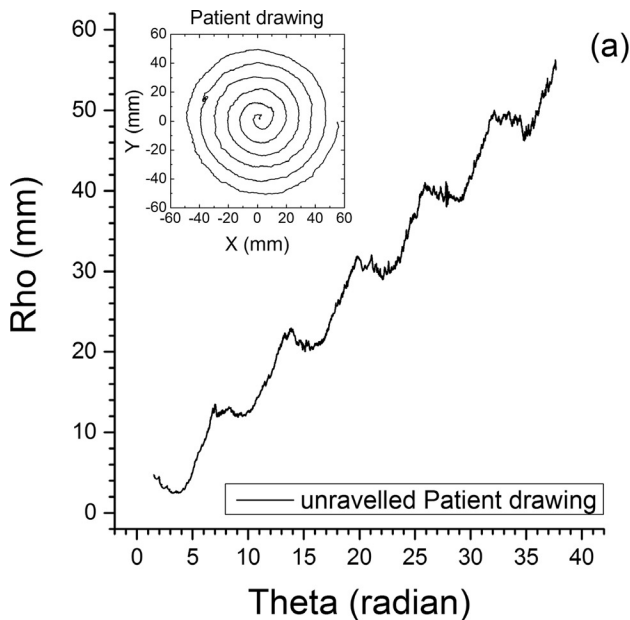


Fig. 4. A typical patient drawing with a noticeable parallax effect. (a) Unravalled Patient drawing; (b) Fit obtained using Eq. (1).

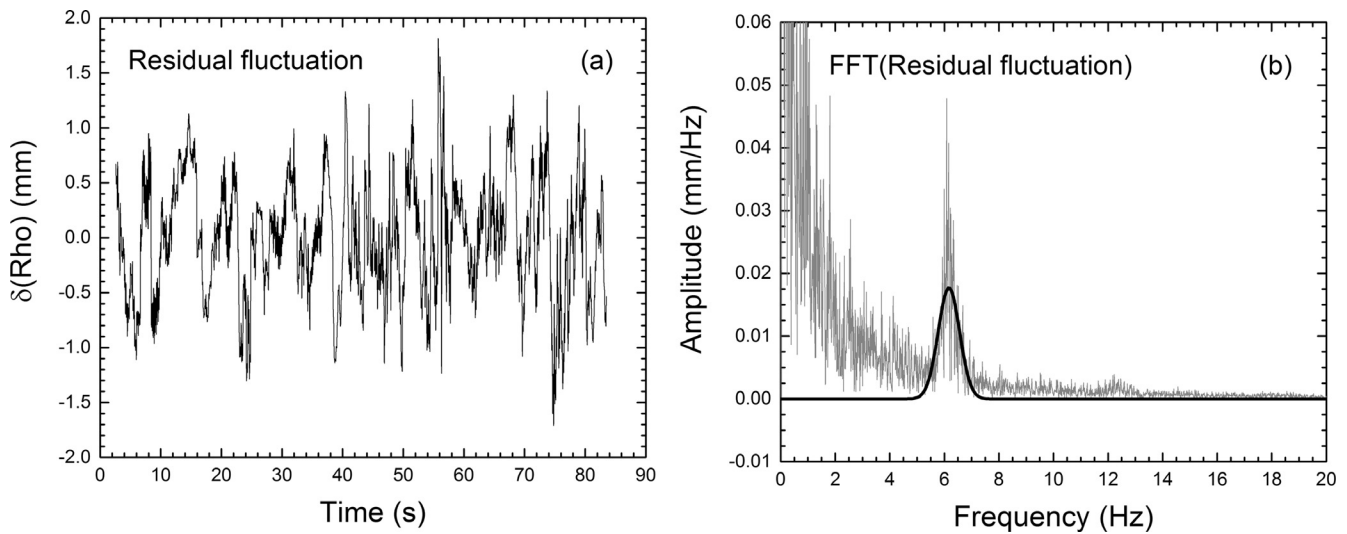


Fig. 5. (a) Residual fluctuation (URS = mean square deviation of $\delta(\text{Rho})$); (b) Spectrum and Gaussian fit (URD = Gaussian amplitude).

fr/patrick.flandrin/emd.html). Fig. 6 shows the different IMFs of the X component of our spiral example.

The possibility to dissociate tremor from voluntary movement is illustrated on Fig. 7, where the X and Y components of the signal were decomposed into IMF1 (tremor) on one hand, and the remaining IMFs + the residue (voluntary movement) on the other hand. Third descriptor: The average areas under Gaussian curves fitted of first FFT IMFs

Fig. 8 shows that the first IMF contains the main part of the tremor. Thus, we chose to characterize the amplitude of the tremor using the first IMF. Like for the velocity analyses the areas were averaged over X and Y. The resulting descriptor is referred to as “EMD” in the following. As for the V descriptor, the variability of the EMD based method is mainly due to the operator-dependent Gaussian fits. To evaluate the reproducibility of the EMD values, two different couples (X and Y) of Gaussian fits were performed for each spiral.

2.4.4. Evolution of the tremor amplitude and frequency over time

Visual or numerical ratings aim to determine the average tremor frequency as well as its intensity. Due to the implicit assumption of the stationarity of the signal representing the spiral (see above), these ratings do not take the temporal aspects of the drawing into account. Time-frequency analysis could help to characterize the non-stationarity of the signal. Indeed, wavelet based analysis of the evolution of the tremor frequency and amplitude peak of the drawing was used to distinguish ET from Parkinson (Kragelj et al. 2014).

Such an analysis could also be performed coupled to EMD. The time course of the first IMF, characterizing the tremor itself, can be conveniently studied using the Hilbert-Huang Transform (HHT) or Short Time Fourier Transform as shown in Fig. 9, fluctuations of tremor frequency and intensity are evidenced. The amplitude of the main frequency peak clearly oscillates around its main value of 6 Hz.

Frequency peaks of variable amplitude can be observed. In the 1–6 Hz domain, different peaks are associated to periodical fluctuations of the main ET amplitude. This phenomenon was observed on many patients of this series. Although these variations through time are rarely tackled in the literature, they might be of interest for further studies.

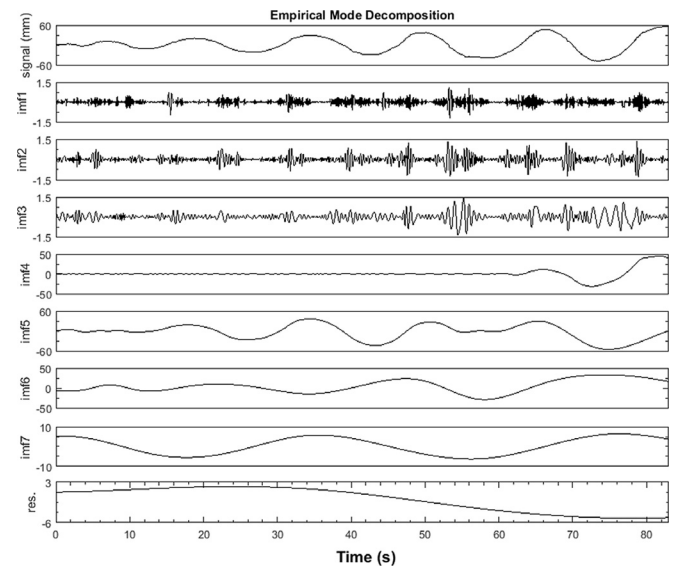


Fig. 6. EMD of the X component. For this drawing, 7 IMFs were obtained from the X-position signal, with a residue. Vertical scales are in mm. The three first IMFs have amplitudes lower than the four next ones.

2.4.5. Statistical analysis

The interrater agreement, i.e. the agreement between the two scores of the raters at a few weeks interval was evaluated using Cohen's weighted kappa with quadratic weights (Fleiss et al., 1969). Cohen's weighted kappa was also used to quantify the interrater agreement, i.e. the agreement between the mean scores of a couple of raters. Note that the intraclass correlation coefficient ICC(2,1) of Shrout and Fleiss (1979) was shown to be asymptotically equivalent to Cohen's weighted kappa with quadratic weights (Fleiss and Cohen 1973), and indeed provides the same values on our data. The intraclass correlation coefficient ICC(2,1) was used to estimate the overall interrater agreement, i.e. the agreement of the mean scores of the seven raters.

As already observed by several authors, the relation between the raters overall mean score and the different computerized scores is essentially logarithmic. Haubenberger (2011) mentions Fechner's law of psychophysics “which predicts a relationship between perceived magnitude (e.g. clinical tremor rating scale;

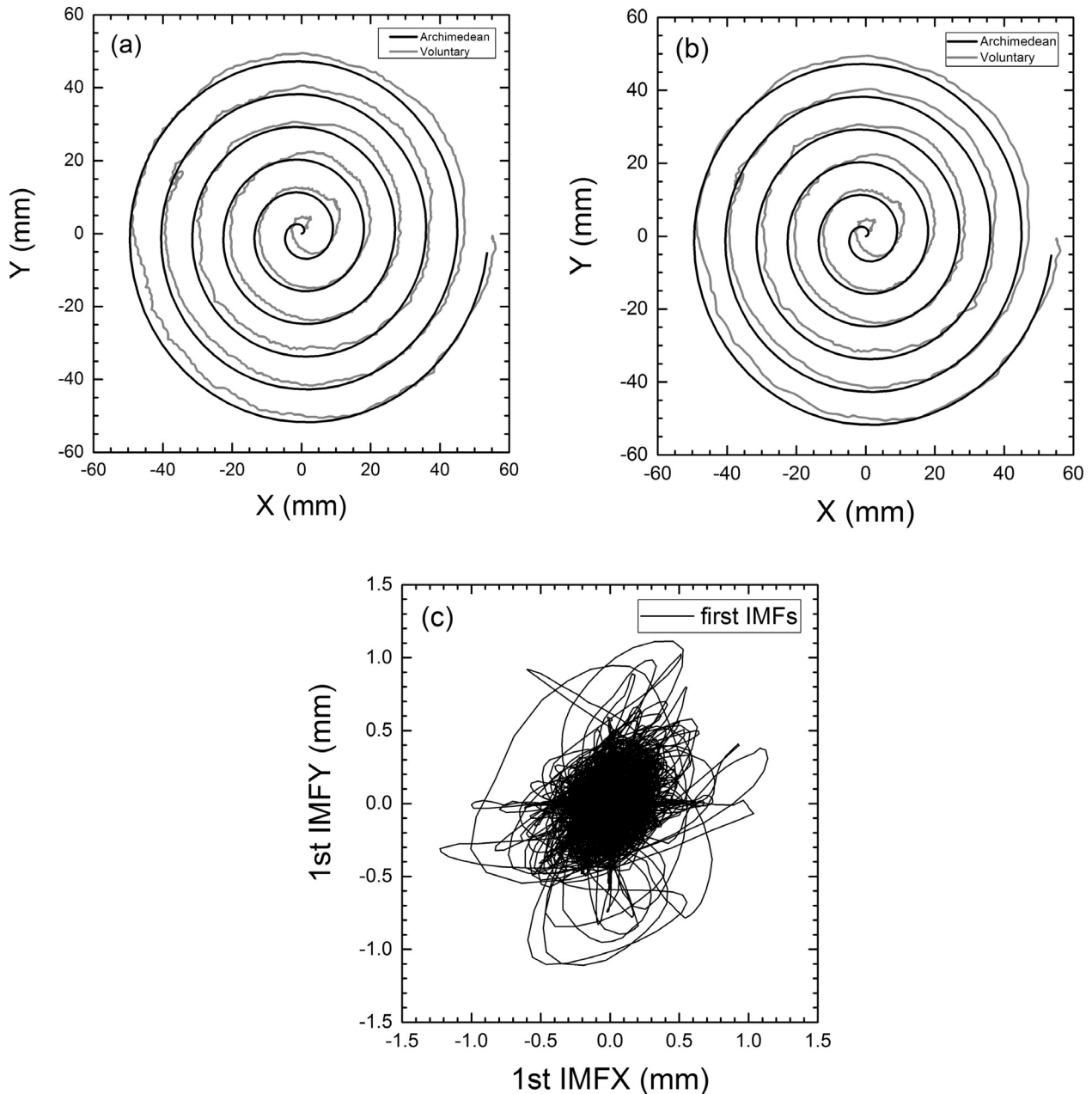


Fig. 7. Decomposition of a spiral drawing (“on” the template spiral) into tremor and voluntary movement. (a) Original patient drawing; (b) Template and summation of IMFs 2 to 7 + residue (res); (c) first X and Y IMFs amplitudes.

TRS) and the log of the measured amplitude”, as also demonstrated in Elble et al. (2006). Thus, the four descriptors were log-transformed, i.e. transformed into magnitudes. Furthermore, in order to provide scores comparable to those of the raters, these magnitudes were rescaled in the interval [0; 10] using an affine transformation calibrated on the first series of measurements, namely V, EMD, URS and URD scores.

Thanks to the repeats (2 for the V score and the EMD score, 3 for the URS score and 6 for the URD score), the variance of the differences sources of variability (Gaussian fits and choice of the start- and end-points of the spirals) could be estimated.

The ICC(2,1) was used to estimate the agreement between the log-transformed and rescaled repetitions of a given computerized method, as well as the overall agreement between the four computerized methods.

The linearity between the mean expert rating and the mean score of a method (in rescaled magnitude) was tested using the

Kolmogorov based non-parametric test developed by Passing and Bablok (1983).

3. Results

3.1. Intra and interrater agreement

The intrarater agreement, i.e. the agreement between the two scores of each of the 7 raters was first evaluated separately for the spirals drawn on and between the printed lines (47 and 45 spirals respectively), using Cohen’s weighted kappa with quadratic weights. The intrarater agreement was excellent for all the raters, both for the spirals “on” (kappa between 0.87 and 0.96) and the spirals “between” (kappa between 0.91 and 0.96). Since the slight average superiority of the agreement spirals “between” was not systematic and not significant considering the confidence intervals, we decided not to make the distinction between the two types of

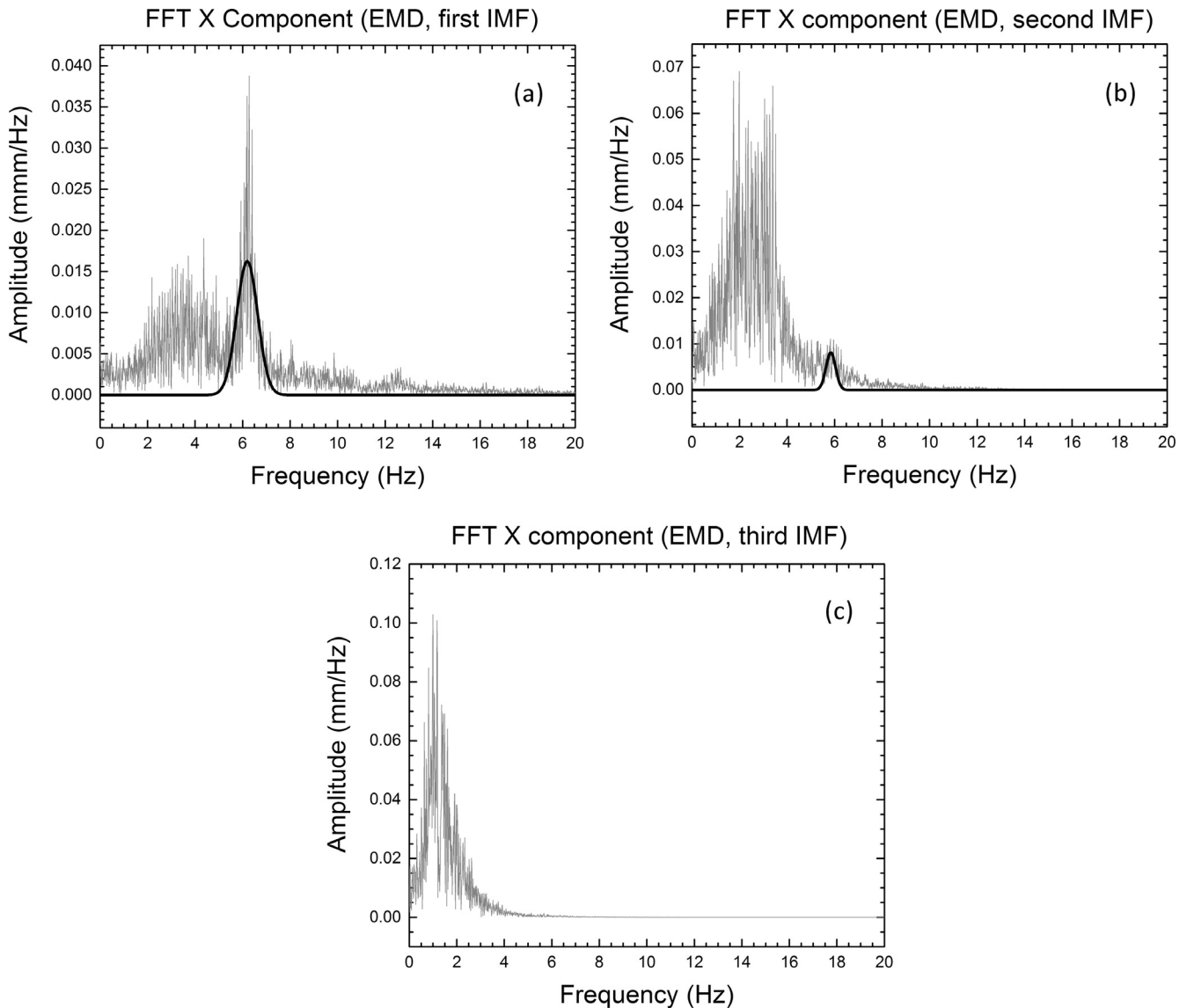


Fig. 8. FFT of the 3 first IMFs of X component. Spectra and Gaussian fits around tremor frequency.

spirals anymore. Evaluated on the 92 spirals, the kappa values for the two scores of the 7 raters were in the interval 0.896 – 0.955.

The interrater agreement was then measured for all couples of raters using the mean of their two scores: the corresponding kappa values were in the interval 0.80 – 0.95. The overall interrater agreement was measured by the intraclass correlation coefficient ICC (2,1) of Shrout and Fleiss, with a point estimate of 0.88 and a 95% confidence interval of [0.80; 0.92]. This interrater agreement can be conveniently visualized on Fig. 10.

3.2. Reproducibility and agreement between the computerized methods

Thanks to the repeats, the variance due to operator choices (Gaussian fits and choice of the start- and end-points of the spirals) could be estimated for each of the computerized methods on the log-transformed and rescaled magnitudes. For both the V and the EMD scores, the variation was due to the Gaussian fits and two repeats were available, leading to standard error estimates with 92 degrees of freedom (d.o.f.). For both methods, the standard error

estimate equals 0.38, with ICC(2,1) values of 0.97 and 0.96 for the V and EMD score respectively. The variability of the URS score, due to choice of the start- and end-points of the spirals, could be estimated using three repeats, leading to a standard error estimate of 0.11 (with 184 d.o.f.) and to an ICC(2,1) of 1.00. The variability of the URD score is due both to Gaussian fits and spiral cuts. Since 2 Gaussian fits were performed for 3 different cuts, we could estimate the standard error due to the Gaussian fits at 0.13 with 276 d.o.f. Using only the first of the two Gaussian fits for each spiral cut, we obtain a global standard error (due to both Gaussian fit and spiral cuts) of 0.16 (with 184 degrees of freedom). Hence a standard error due to the spiral cuts of 0.09. The major part of the variance is hence due to the Gaussian fits. Thus, in the following, we consider the 6 repeats of URD as independent, with a global standard deviation of 0.16, and an ICC(2,1) of 0.99. Thus, the variability of the four scores is lower than the variability of an expert's notation. The reproducibility of the four computerized methods can be visualized on Fig. 11.

Finally, the overall agreement of the four computerized scores can be characterized by their intraclass correlation coefficient ICC

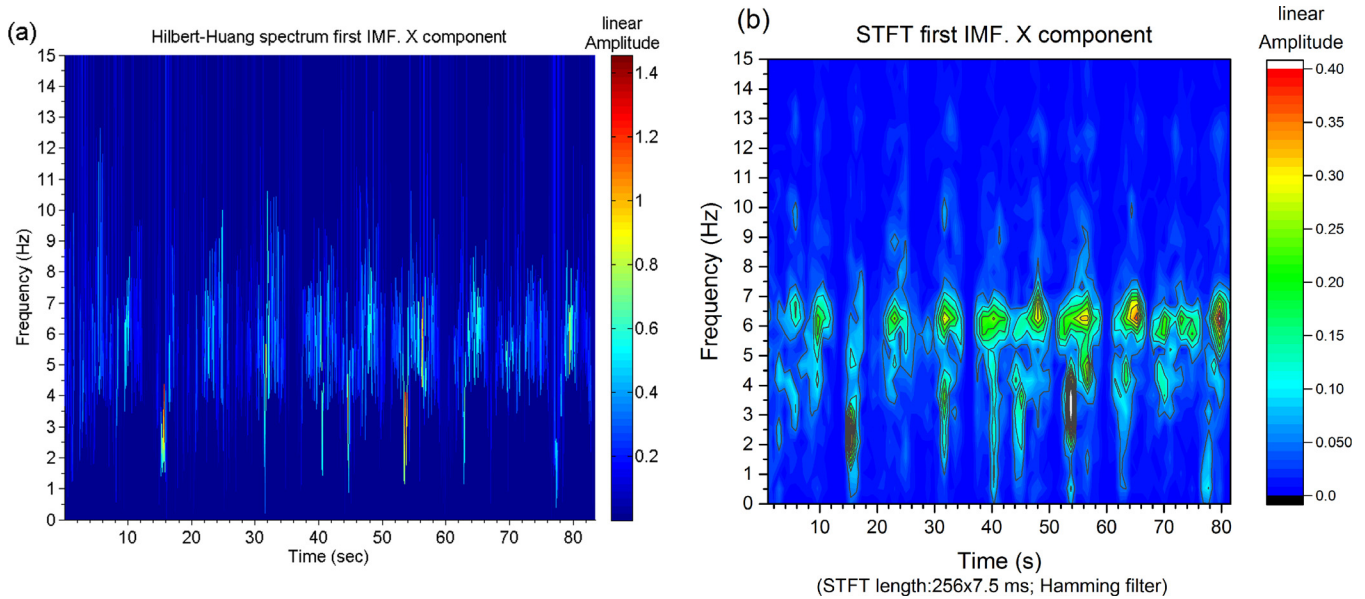


Fig. 9. Transforms of the first IMF X component. (a) Hilbert-Huang spectrum; (b) Short Time Fourier Transform. Colour bars correspond to a linear amplitude scale.

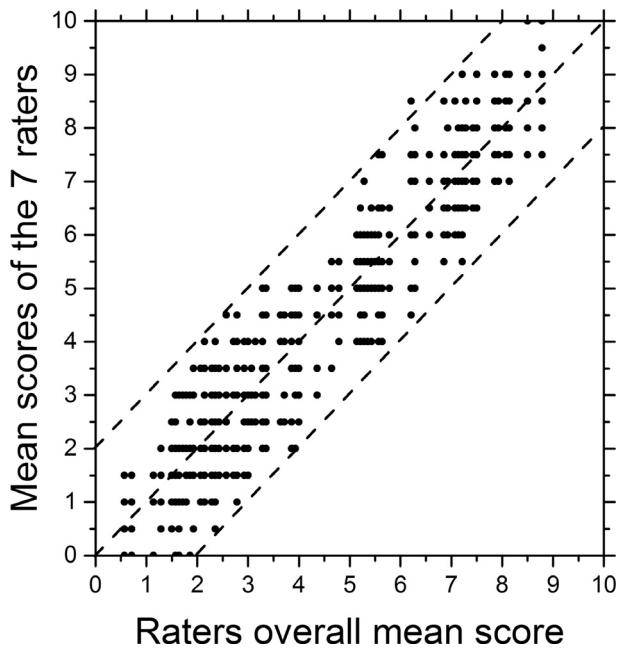


Fig. 10. Interrater agreement. The mean score of a given rater almost never differs by more than 2 points from the raters overall mean score, with an intraclass coefficient value of 0.88 (95% confidence interval [0.80; 0.92]).

(2,1) of their mean values, which equals 0.91 (95% confidence interval [0.87; 0.93]).

3.3. Agreement between the visual rating and the computerized methods

The agreement between the mean rating of the experts and the mean rating of each computerized method can be evaluated using the ICC(2,1), leading to values of 0.74, 0.74, 0.84 and 0.85 for the V, EMD, URS and URD scores respectively. Mean expert ratings and scores are shown in Fig. 12.

For the URS and URD scores, similarly to the expert ratings, the mean score almost never differs by more than 2 points from the

raters overall mean score. Moreover, one has to note the less dispersed values obtained with the computerized methods, especially URS and URD. However, a test of linearity shows that none of the rescaled magnitudes is a strictly linear function of the expert rating.

4. Discussion

4.1. Overall comparison

Intrarater agreement was good for all the raters, for both the spirals “on” and the spirals “between”. Interrater agreement was also high (ICC(2,1) of 0.88): visual rating scores were, for a given spiral, at a distance of 2 or less rating points from the overall mean on Bain (0–10) point scale.

The overall agreement of the four numerical methods of analysis was excellent (ICC(2,1) of 0.91). Like the visual rating scores, the computerized spiral score analysis did not show any significant difference between the spirals “on” and the spirals “between”.

In order to compare the numerical methods of analysis with visual rating, the parameter values obtained via these methods had to be converted in their logarithm values. Once expressed in magnitudes, the numerical ratings were shown to be in good correlation with the visual rating (Pearson correlation coefficients between 0.79 and 0.87), with a slight advantage for our three new methods, especially URD (0.87). Once converted in scores between 0 and 10 with a fixed affine relationship calibrated on the first measurement, the numerical ratings were also in excellent agreement with the visual rating (CC(2,1) between 0.74 and 0.85), the best agreement being again obtained with URD.

Thus, our new numerical methods are able to provide more reproducible ratings than the experts, while being in excellent agreement with their mean rating.

4.2. New insights provided by the proposed numerical methods

In addition to their reproducibility and accuracy, the unravelling methods enable to obtain two types of information. Like the Bain rating scale, unravelling static (URS) provides a quantification of the amplitude of the tremor. Unravelling dynamic (URD) pro-

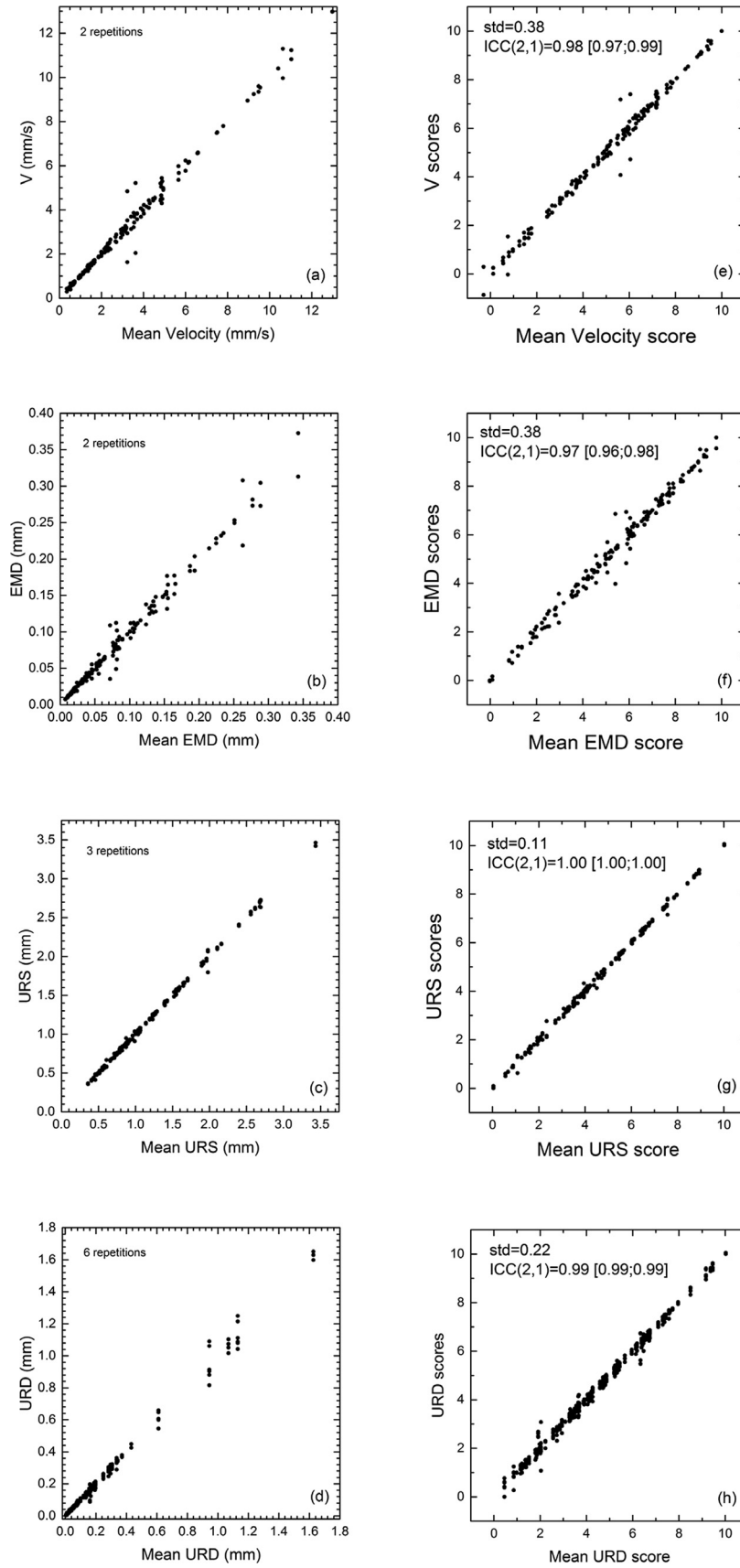


Fig. 11. Reproducibility of the computerized scores. Column 1: linear scale value plot. Column 2: affine transform magnitude plot, in the interval [0; 10] using a calibration obtained on the first series of measurements.

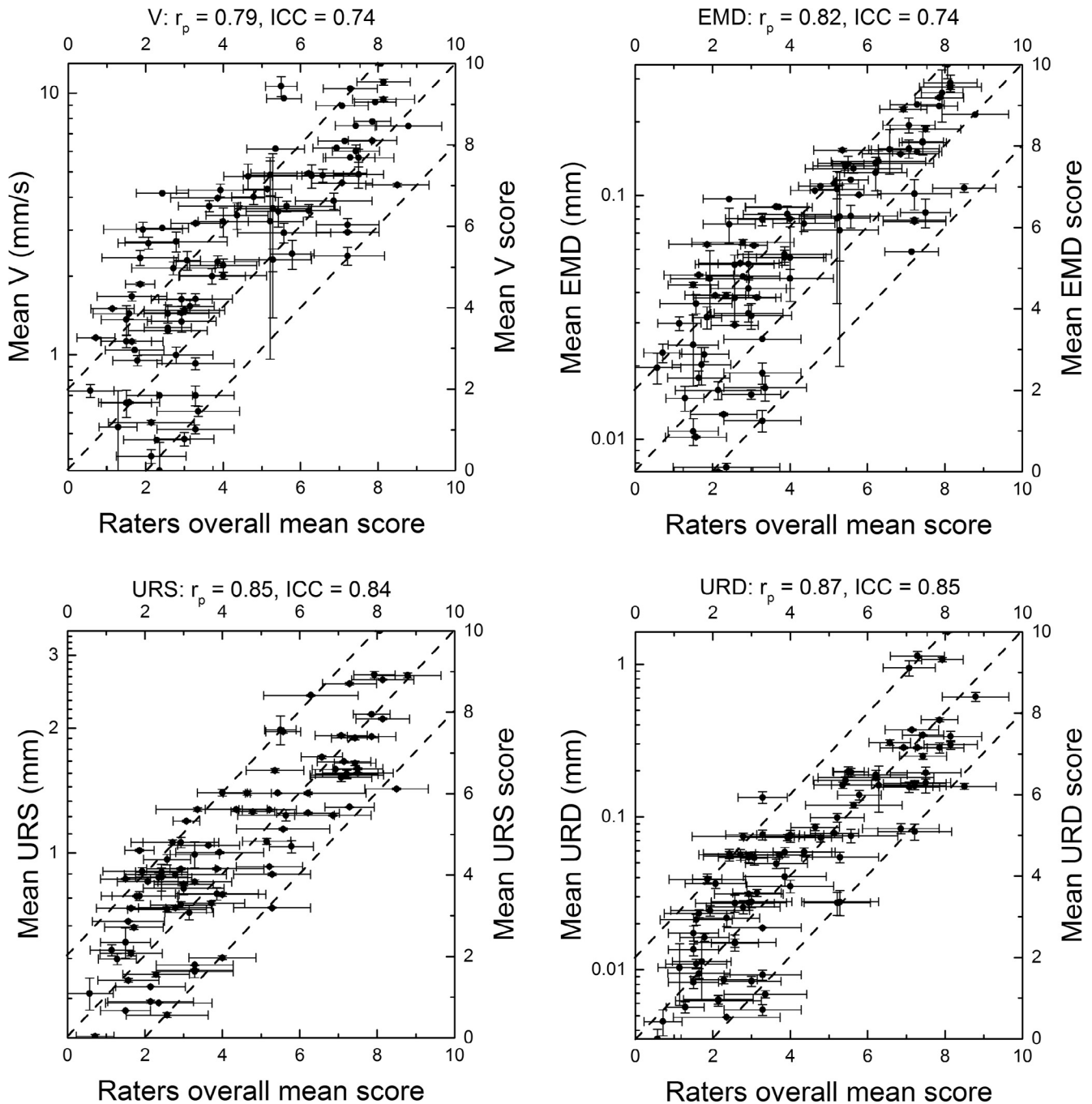


Fig. 12. Means of the values (left vertical scale) and means of the log and affine transformed computerized scores (right vertical scale) as a function of the raters overall mean scores, and corresponding Pearson correlation and ICC(2,1) values. For each spiral, the vertical and horizontal bars are the standard deviations estimated with the repeats.

vides information on the amplitude and also on the frequency of the tremor.

Moreover, as shown on Fig. 12, the mm URD scale value range is of three orders of magnitude, which may enable a more sensitive quantification of the effect of the clinical treatment. Finally, fluctuations of frequency and amplitude of ET during the time of observation are likely to be evidenced with such methods of analysis. The latter information may also be of interest for the study other movement disorders.

Finally, although EMD proves less efficient than URD and URS, it might be used for different types of freehand drawing analysis. As a matter of fact, the method does not depend on assumptions made about the drawing, such as the Archimedean shape, so that the

analysis of the first IMF's gives insight into any type of drawing done by a patient.

4.3. Spiral Analyser software

In order to enable an easy and fast analysis, we have developed a stand-alone Matlab application, available upon request², which determines the URS and URD values directly from the data record. Moreover, the operator is free to cut beginning and/or end parts of the patient drawing by different amounts and to visualize the corresponding evolution of the URS and URD values as well as their means

² Email for requests at: andre-pierre.legrand@espci.fr

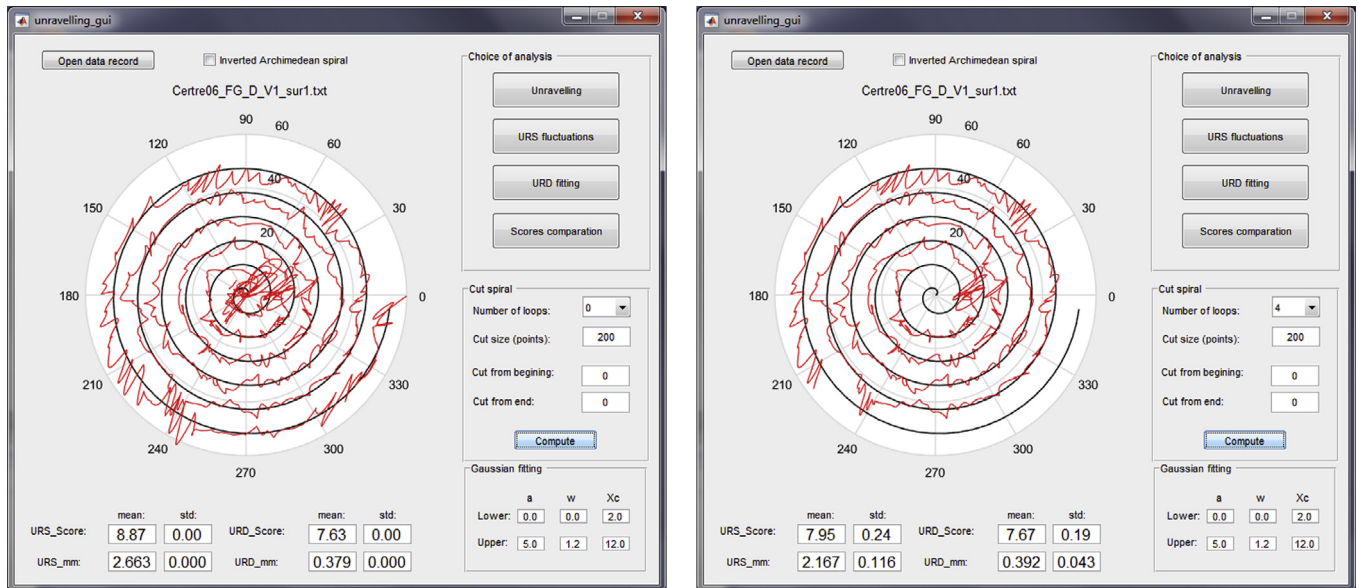


Fig. 13. Tremor determination. Left: data just as imported. Right: analysis done with 4 different sequences, 200 points being removed at the beginning and at the end from one sequence to the next one. Standard deviations of scores and in mm give information on the reliability of such analysis.

and standard deviations. Different facilities are provided to improve the analysis (Fig. 13).

5. Conclusion

Different publications have proposed methods of characterisation to complement and replace visual rating scores on FTM and Bain rating scales, in order to improve reliability and objectivity in ET routine and clinical assessment trials. Spiral drawing analysis, sometimes denominated spirography, is one of the usual tests for documenting tremor. The digital conversion of spiral drawings enables the use of mathematical methods of analysis to quantify tremor amplitude, the most common method of analysis relying on the Fourier spectrum of the velocity. Here, we have proposed three original methods of characterisation which outperform the velocity analysis, are more reproducible than the expert rating, and hence are likely to be implemented in an easy-to-use alternative system. Moreover, they provide valuable complementary information on tremor amplitude and frequency variations over time that could be used to improve the diagnosis.

Acknowledgments

Visual spiral rating were performed among other raters by Chloé Laurencin, MD, Cécile Hubsch-Bonneaud, MD Ph.D and Yara Beaugendre, MD; the software for spiral drawing recording on the digitizing tablet was developed by Dmytro Mykhalyk, Ph.D and Spiral Analyser software by Ivan Mudryk, MS.

Conflicts of interest: Dr. Legrand AP, Dr. Rivals I, Ms. Richard A and Dr. Meunier S report no disclosures.

Dr. Apartis E received research support from INSERM (COSSEC) and APTES (Association des Personnes concernées par le Tremblement ESsentiel).

Dr. Roze E received a "poste d'accueil" AP-HP/CNRS grant; received research support from INSERM (COSSEC), AP-HP (DRC-PHRC), Fondation pour la Recherche sur le Cerveau (FRC), Merz-Pharma, Orkyn, IP santé; served on scientific advisory boards for Orkyn, and Merz-pharma; received speech honorarium from Merz-pharma, Novartis, Ipsen-Pharma and Orkyn, received travel funding from Ipsen-Pharma, Teva, Abbvie, Merz-Pharma, Dystonia

Europe, the Georgian Medical and Public Health Association the International Federation of Clinical Neurophysiology, and the Movement Disorders Society.

Dr. Vidailhet M received research support from INSERM (ANR), APHP(PHRC) and patients association (APTPEs, France Parkinson, AMADYS), and travel grants from the Movement Disorders society (speaker), advisory board for Merz and Medtronic.

Dr. Hainque E received travel funding from JNLF (Journées Neurologiques de Langue Française) and Naturalia and Biologia association.

References

- Bain PG, Findley LJ, Atchison P, Behari M, Vidailhet M, Gresty M, Rothwell JC, Thompson PD, Marsden CD. Assessing tremor severity. *J Neurol Neurosurg Psychiatry*. 1993;56:868–73.
- Bajaj N, Knöbel M, Gontu V, Bain PG. Can spiral analysis predict the FP-CIT SPECT Scan result in tremulous patients? *Mov Disord* 2011;26(4):699–704.
- Deuschl G, Bain PG, Brin M. Consensus statement of the movement disorder society on tremor. *Mov Disord* 1998;13(3):2–23.
- de Lima ER, Andrade AO, Pons JL, Kyberd P, Nasuto SJ. Empirical mode decomposition: a novel technique for the study of tremor time series. *Med Bio Eng Comput* 2006;44(7):569–82.
- Elble RJ, Sinha R, Higgins C. Quantification of tremor with a digitizing tablet. *J Neurosci Methods* 1990;32:193–8.
- Elble RJ, Brilliant M, Leffler K, Higgins C. Quantification of essential tremor in writing and drawing. *Mov Disord*. 1996;11(1):70–8.
- Elble RJ, Pullman SL, Matsumoto JY, Raethjen J, Deuschl G, Tintner R. The tremor research group. tremor amplitude is logarithmically related to 4- and 5-point tremor rating scales. *Brain* 2006;129(10):2660–6.
- Erasmus L-P, Sarno S, Albrecht H, Schwecht M, Pöllmann W, König N. Measurement of ataxic symptoms with a graphic tablet: standard values in controls and validity in multiple sclerosis patients. *J Neurosci Methods* 2001;108:25–37.
- Fahn S, Tolosa E, Conception M. *Clinical Rating Scale for Tremor*. In: Jankovic J, Tolosa E, editors. *Parkinson's disease and movement disorders*. 2nd ed. Baltimore: Williams and Wilkins; 1993. p. 271–80.
- Feys P, Helsen W, Prinsmel A, Ilsbrouckx S, Wang S, Liu X. Digitised spirography as an evaluation tool for intention tremor in multiple sclerosis. *J Neurosci Methods* 2007;160:309–16.
- Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psych Measurement* 1973;33:613–9.
- Fleiss JL, Cohen J, Everitt BS. Large sample standard errors of kappa and weighted kappa. *Psych Bull* 1969;72. 332–327.
- Haubenberger D, Kalowitz D, Nahab FB, Toro C, Ippolito D, Luckenbaugh DA, Wittevrongel L, Hallett M. Validation of digital spiral analysis as outcome parameter for clinical trials in essential tremor. *Move Disord* 2011;26(11):2073–80.

- Hess CW, Hsu AW, Yu Q, Ortega R, Pullman SL. Increased variability in spiral drawing in patients with functional (psychogenic) tremor. *Hum Movement Sci* 2014;38:14–22.
- Hsu AW, Piboolnurak PA, Floyd AG, Yu QP, Wraith JE, Patterson MC, Pullman SL. Spiral analysis in Nieman-Pick disease type C. *Move Disord* 2009;24(13):1984–90.
- Kragelj V, Georgiev D, Pirtošek Z, Ribarič S. Wavelet analysis increases sensitivity and specificity of spirometry for ambulatory tremor discrimination. *Biomed Res Int* 2014;2014:934746. Doi: 10.1155/2014/934746.
- Kraus PH, Hoffmann A. Spiralometry: computerized assessment of tremor amplitude on the basis of spiral drawing. *Move Disord* 2010;25(13):2164–70.
- Lacquaniti F, Ferrigno G, Pedotti A, Soechting JF, Terzuolo C. Changes in spatial scale in drawing and handwriting: kinematic contribution by proximal and distal joints. *J Neurosci* 1987;7(3):819–28.
- Liu X, Carroll CB, Wang S-Y, Zajilek J, Bain PG. Quantifying drug-induced dyskinesias in the arm using digitised spiral-drawing tasks. *J Neurosci Methods* 2005;144:47–52.
- Louis ED, Gillman A, Boschung S, Hess CW, Yu Q, Pullman SL. High width variability during spiral drawing: further evidence of cerebellar dysfunction in essential tremor. *Cerebellum* 2012;11(872):9.
- Marquardt C, Mai N. A computational procedure for movement analysis in handwriting. *J Neurosci Methods* 1994;52:39–45.
- Michalec M, Hernandez N, Clark L.N., Louis E.D. The spiral axis as a clinical tool to distinguish essential tremor from dystonia. *Parkinson Relat Disord* 2014;20:541–544.
- Miralles F, Tarongí S, Espino A. Quantification of the drawing of an Archimedes spiral through the analysis of its digitized picture. *J Neurosci Methods* 2006;152:18–31.
- Mergl R, Tigges P, Schröter A, Möller H-J, Hegerl U. Digitized analysis of handwriting and drawing movements in healthy subjects, results and perspectives. *J Neurosci Methods* 1999;90:157–69.
- Ondo WG, Wang A, Thomas M, Vuong KD. Evaluating factors that can influence spirometry ratings in patients with essential tremor. *Parkinson Relat Disord* 2005;11:45–8.
- Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in Clinical Chemistry, Part I. *J Clin Chem Clin Biochem* 1983;21:709–20.
- Pullman SL. Spiral analysis: a new technique for measuring tremor with digitizing tablet. *Move Disord* 1998;3:85–9.
- Saunders-Pullmann R, Derby CD, Stanley K, Floyd A, Bressman S, Lipton RB, Deligtisch A, Severt L, Yu Q, Kurtis M, Pullman SL. Validity of spiral analysis in early Parkinson's disease. *Move Disord* 2008;23(4):531–7.
- Schneider AS, Baur B, Fürholzer W, Jasper I, Marquardt C, Hermsdörfer J. Writing kinematics and pen forces in Writer's Cramp: Effects of task and clinical subtype. *Clinical Neurophysiol* 2010;121:1898–907.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
- Sisti JA, Christophe B, Seville AR, Garton ALA, Gupta VP, Bandin AJ, Yu QY, Pullman SL. Computerized spiral analysis using the iPad. *J Neurosci Methods* 2017;275:50–4. <http://dx.doi.org/10.1016/j.jneumeth.2016.11.004>.
- Stacy MA, Elble RJ, Ondo WG, Wu S-C, Hulihan J. Assessment of interrater and intrarater reliability of the fahn-tolosa-marin tremor rating scale in essential tremor. *Move Disord* 2007;22(6):833–8.
- Viviani P, Burkhard PR, Chiuvé SC, dell'Acqua CC, Vindras P. Velocity control in Parkinson's disease: a quantitative analysis of isochrony in scribbling movements. *Exp Brain* 2009;194:259–83.
- Wang H, Yu Q, Kurtis MM, Floyd AC, Smith WA, Pullman SL. Spiral analysis—Improved clinical utility with center detection. *J Neurosci Methods* 2008;171:264–70.
- Wang S, Bain PG, Aziz TZ, Liu X. The direction of oscillation in spiral drawings can be used to differentiate distal and proximal arm tremor. *Neurosci Lett* 2005;384:188–92.
- Xie H, Wang Z. Mean frequency derived via Huang-Hilbert transform with application to fatigue EMG signal analysis. *Comput Meth Progr Biomed* 2006;82:114–20.
- Zeuner KE, Peller M, Knutzen A, Holler I, Münchau A, Hallett M, Deuschl G, Siebner H. How to assess motor impairment in writer's cramp. *Move Disord* 2007;22(8):1102–9.