Identification of drowsiness and alertness conditions by means of Spectral F-Test applied to pupillometric signals

V C G Soares¹, J K S Souza¹, G E Ginani², S Pompéia², C J Tierra-Criollo¹,³ and D B Melges¹

¹ Graduate Program in Electrical Engineering - Federal University of Minas Gerais - Av. Antônio Carlos 6627, 31270-901, Belo Horizonte, MG, Brazil
² UNIFESP/ Psychobiology Department, Federal University of São Paulo, São Paulo, Brazil
³ UFRJ/Biomedical Engineering Program, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Abstract. The autonomous regulation of pupil size provides an objective physiological measure of alertness level. Pupil diameter decreases and its fluctuations increase during drowsiness. In this work, a statistical based technique in frequency domain, known as Spectral F-Test (SFT), was applied in order to compare the power of pupillometric signals measured in two conditions, wakefulness and sleepiness, from eleven volunteers. SFT was calculated based on Welch Periodograms of time series of pupil diameter (TSPD) signals. The median power for TSPD in two conditions and each volunteer at frequencies bellow 0.8 Hz were compared by means of the paired Wilcoxon signed-rank test. The percentages of volunteers for whom power during drowsiness were statistically higher than power during alertness achieved 100% for several frequencies bellow 0.2 Hz. For frequencies from 0.2 to 0.8 Hz, the percentages presented high variability, fluctuating from 40 to 90%. The Wilcoxon test indicated that median power of TSPD during drowsiness was higher than TSPD during alertness for frequencies bellow 0.2 Hz presented statistical difference. Hence, the SFT showed to be a suitable tool for drowsiness/alertness differentiation that could be applied to sleep studies.

1. Introduction
Multiple Sleep Latency Test (MSLT) is a current gold standard method to evaluate somnolence, but its main drawback is the expense of the test, which is based on polysomnography [1]. This procedure consists of recording physiological signals during sleep, such as the electroencephalogram (EEG), electro-oculogram (EOG), electromyogram (EMG), electrocardiogram (ECG), capnogram, oximetry and a video recording of the exam. It also requires patients to remain all day in laboratory, the
presence of a specialized professional on sleep analysis during all exam, and an expensive infrastructure.

On the other hand, pupillometry has been used as an alternative method for objective quantification of sleep. This method consists of recording pupil diameter oscillations [1–3], which is a simple and non-invasive test, cheaper than MSLT.

The measurement of pupil parameters is clinically important, since pupil size indicates, apart from adjustments to different lighting conditions and visual accommodation [4], the presence of physical or mental dysfunctions such as: multiple sclerosis [5], autonomic neuropathy [6], depression [7], Alzheimer's and Parkinson disease [8], retinitis pigmentosa [9] and glaucoma [10]. In addition, pupil size increases with cognitive effort [11].

Pupil dilation also varies between conditions of drowsiness and alertness. During alertness, pupil diameter is relatively stable, while drowsiness leads to considerable pupil size fluctuation [12].

Pupil size can be registered by a camera, using techniques of image processing, resulting in a time series of pupil diameter (TSPD). Analytical tools are applied to the TSPD in order to obtain indicators of drowsiness.

Various pupillometry studies [3,13–15] have shown that as subjects become sleepy, the pupil diameter decreases and oscillates at lower frequency and higher amplitude. This parameter, known as hippus, was also investigated by other researchers [14]. Hence, these works suggested that pupil diameter could be used as a drowsiness indicator [1,3,14,15]. Particularly, the power of pupil oscillations in low frequency components has been used for sleep deprivation identification [14–16]. However, there is no consensus about the frequencies that better translate the drowsiness-to-alertness transition, and conversely.

Lüdtke et al. [15], Wilhelm et al. [2] and McLaren et al. [1] proposed indicators based on the signal power at frequencies up to 0.8 Hz in order to objectively assess the alertness level. These three studies used the same methodology by calculating the following parameters: power of pupillary oscillations (PPO) from 0.01 to 0.8 Hz, mean pupil diameter (MPD) and pupillary unrest index (PUI); the latter measures the pupil instability [1,2,14].

Lüdtke et al. [15] concluded that there is difference in PPO and PUI when these parameters are compared for alert and sleepy subjects, suggesting these variables would be useful to quantify drowsiness. After that, Wilhelm et al. [2] used the three parameters (PPO, MPD and PUI) to differentiate successfully healthy patients and hypersomniacs. Moreover, McLaren et al. [1] found correlation between Sleep Latency (MSLT) and the parameters PPO and PUI. However, any variable was able to substitute MSLT [1]. On the other hand, PPO and PUI are still the main parameters studied today [5,12,16–20]. Different from these parameters, an alternative approach using frequency components from the power spectrum density (PSD) of pupillary response has been suggested [20,21]. In these studies, three self-report questionnaires were used to measure sleepiness: the Stanford Sleepiness Score (SSS), Visual Analogue Scales (VAS) and the Epworth Sleepiness Scale (ESS). These questionnaires were associated with the power in four different frequency ranges of pupillary oscillations in order to discriminate between two groups of different sleepiness level.

In the present paper, a statistical based technique in frequency domain [22], known as Spectral F-Test (SFT), was applied in order to compare the power of pupillometric signals in two conditions: alertness and drowsiness. Statistical difference between median power in both conditions was assessed in order to establish the best frequencies to distinguish them.

2. Methods

2.1. Signal acquisition

The pupillogram was recorded from eleven male (aged from 21-33 years) volunteers without a history of neurological or psychiatric diseases, using the pupillometer developed at Biomedical Engineering Laboratory of Federal University of Minas Gerais (UFMG) [23]. The participants were awake and continuously supervised during the experiment, which was performed in Department of Psychobiology.
The pupillometry video was recorded (sampling rate: 120 samples per second) at 8 am after a night of sleep (wake session) and again after 24 h of continuous alertness (sleep deprived condition) for 10 minutes each (Table 1). During pupil size estimation, participants carried out a Psychomotor Vigilance Test (PVT) [24], which involves pressing a button every time they saw a visual stimulus generated by a LED stimulator.

Factors that could interfere with the drowsiness quantification such as caffeine and medication ingestion were used as exclusion criteria [23]. The temperature and light were controlled and the volunteers had access to external stimuli such as day time and day light. This research was approved by local ethics committee (UNIFESP; 0763/10) and all volunteers signed written informed consent [23].

| Section | Time | Day | Signal characteristics |
|---------|------|-----|-------------------------|
| 1st     | 8:00 | 1st | Record after a normal sleep night |
| 2nd     | 8:00 | 2nd | Record after 24 hours* of sleep deprivation |

*Considering the Circadian Rhythm.

2.2. Pre-processing

After video acquisition, pupil size was estimated leading to time series of pupil diameter (TSPD). When blinks occurred (pupillary diameter = 0), this measure was adjusted by interpolation [23].

Considering that for the first two minutes of recording the volunteers were alert even for the 2th session (drowsiness condition), this segment was discarded. The TSPD was filtered by a 2 Hz low-pass, 4th order, zero phase, Butterworth filter.

2.3. Spectral F-Test (SFT)

Spectral F-Test (SFT) assess whether two sample spectra are from populations with identical power spectrum [25]. In this study, the SFT was calculated as the ratio between the power of filtered TSPD in drowsiness condition (2th session) by the power in alertness condition (1st session) [22]:

$$\phi(f) = \frac{\hat{P}_{yy}(f)}{\hat{P}_{xx}(f)} = \frac{1}{M_y} \sum_{i=1}^{M_y} |Y_i(f)|^2 \frac{1}{M_x} \sum_{i=1}^{M_x} |X_i(f)|^2$$

(1)

Where $\hat{P}_{yy}(f)$ and $\hat{P}_{xx}(f)$ are the periodograms of signal $y[k]$ and $x[k]$, respectively, $f$ is frequency, “\^” superscript denotes estimation, $X_i(f)$ and $Y_i(f)$ are respectively, the Fourier Transform of i-th epoch of $x[k]$ and $y[k]$, and $M$ is the number of epochs.

Considering that $\hat{P}_{yy}(f)$ belongs to the same population that $\hat{P}_{xx}(f)$, one can establish the null hypothesis ($H_0$) of power equality. Assuming that TSPD, in both conditions, is a zero mean Gaussian random signal, it can be showed that the probability density function of the ratio in equation (1) tends asymptotically to an F-distribution with $2M_x$ and $2M_y$ degrees of freedom [25]. Hence, critical values for $H_0$ and a given significance level $\alpha$ [22], can be calculated as:

$$\phi_{crit} = F_{crit}(2M_x, 2M_y, \alpha)$$

(2)
Statistical difference between powers of distinct sessions is given when SFT estimation exceeds the critical value for a frequency \( f(\phi(f) > \phi_{\text{crit}}) \), which leads to the null hypothesis rejection of power equality.

2.4. **Processing**

After pre-processing, the Power Spectral Density (PSD) for signals in both conditions, \( P_{xx}(f) \) and \( P_{yy}(f) \), were calculated using Welch Periodogram for 8-minutes recordings, windowed in 2-minutes epochs (spectral resolution of 0.008 Hz) with 5% overlapping, leading to \( M_x = M_y = 61 \) epochs. The significance level \( \alpha = 0.05 \) was used in equation (2) for critical value calculation.

3. **Results**

Figure 1 shows pupillograms (TSPD) for subject #11 in two different conditions, after filtering process. The drowsiness signal clearly presents higher amplitude fluctuations than alertness one. PSD of TSPD for the same subject shows that the pupillogram power at lower frequencies, mainly below 0.2 Hz, increases during somnolence (Figure 2). This behavior described for both time and frequency for volunteer #11 was similar for other individuals.

The percentages of volunteers for whom power during drowsiness were statistically higher than the power during alertness, estimated by means of SFT for frequencies up to 0.8 Hz are presented in figure 3. It is worth noting, below 0.2 Hz, several frequencies presented percentages of 100%, whereas above this frequency, only two components (0.337 and 0.747 Hz) showed this rate. Percentages higher than 90% are also more common below 0.2 Hz.

In order to verify the frequency range that is more suitable to differentiate drowsiness from alertness condition, a paired Wilcoxon signed-rank test was applied to PSD for all volunteers. P-values for the test are showed in figure 4 that evidences significant differences between median powers for two conditions powers, particularly for the frequency range below 0.2 Hz (p-values lower than 0.01, except for 0.08 Hz). For frequencies over 0.2 Hz, the p-values presented high variability, with values often over 0.01.

4. **Discussion**

The phenomenon of pupillary fluctuations, known as “waves of fatigue”, was firstly described by Lowenstein et al. in 1963 [18 apud 1]. Later, other studies [3,13,14,15] involving pupillary activity became important references when they noted the increasing of pupillary oscillations at low frequencies.

Similar to Lüdtke et al. [15] work, in this study, power of pupil oscillations in low frequency range was used to differentiate two conditions: alertness and drowsiness for eleven volunteers. However, the power of time series of pupil diameter (TSPD) was compared by means of a statistical test, the Spectral F-Test (SFT), applied to two different conditions, for each volunteer. The SFT results showed that the 2th session (sleep deprivation) presented statistically significant higher power than the 1st in frequency range below 0.8 Hz for the majority of volunteers. These findings agree with other works that employed different parameters based on pupil oscillations power [1,2,14]. However, the paired Wilcoxon signed-rank test applied to PSD values of both conditions for all volunteers showed that frequencies up to 0.2 Hz presented significant differences in median power. Hence, frequencies below 0.2 Hz, instead of 0.5 Hz [12] or 0.8 Hz [1, 2,14], should be used to differentiate alert from sleepy subjects. In fact, these frequencies presented the highest percentages of volunteers for whom power during drowsiness were statistically higher than during alertness. These percentages were often above 90%. For frequencies from 0.2 to 0.8 Hz, the percentages presented high variability, fluctuating from 40 to 90%.

It should be borne in mind, however, that the participants in the present study were carrying out a cognitive task while the pupil size was measured. Hence, it is possible that cognitive effort under situations of different arousal levels led to changes observed here.
Figure 1. TSPD for subject #11 in two states: alertness and drowsiness.

Figure 2. PSD for subject #11 at two states: alertness and sleepiness.
Figure 3. Percentages of volunteers for whom power during drowsiness were statistically higher than the power during alertness, estimated by means of SFT.

Figure 4. Result of Wilcoxon paired test indicating frequencies below 0.2 Hz presented statistical difference between alertness and drowsiness. Y-scale in logarithmic values.
5. Conclusion
This work proposed a novel approach to differentiate alertness from drowsiness condition based on a statistical test. Spectral F-Test showed to be a suitable technique for this purpose. Moreover, a more limited frequency band was identified as the range where the power changes occur. Thus, the best frequencies for evaluating drowsiness seem to be below 0.2 Hz. As conclusion, the analysis of pupillometric signal using SFT and suitable frequencies can be an important tool for sleep studies, since this technique allowed the discrimination of data from two sessions corresponding to different physiological states. Further, this could be applied to identify variations in pupil size associated with circadian rhythm (metabolic cycle involving organic processes that repeat every day, including sleep-wake cycle), by analyzing other sessions performed at different hours of the experiment.

Finally, future studies should focus on the dynamic time evolution of SFT in order to identify drowsiness-to-alertness transitions, and conversely.

Acknowledgments
To CAPES, FAPESP (process 2011/01286-0), CNPq and FAPEMIG for financial support.

References
[1] McLaren J W, Hauri P J, Lin S-C and Harris C D 2002 Pupillometry in clinically sleepy patients. Sleep Med 3 347–52
[2] Wilhelm H, Lüdtke H and Wilhelm B 1998 Pupillographic sleepiness testing in hypersonniacs and normals. Graef Arch Clin Exp 236 725–9
[3] Morad Y, Lemberg H, Yofe N and Dagan Y 2000 Pupillography as an objective indicator of fatigue. Curr eye res 21 353–42
[4] Kandel E R, Schwartz J H and Jessel T M 2000 Principals of neural science (New York: McGraw-Hill)
[5] Frauscher B, Egg R, Brandauer E, Ulmer H, Berger T, Poewe W and Högl B 2005 Daytime sleepiness is not increased in mild to moderate multiple sclerosis: a pupillographic study. Sleep Med 6 543–7
[6] Ferrari G L, Marques J L B, Gandhi R a, Heller S R, Schneider F K, Tesfaye S and Gamba H R 2010 Using dynamic pupillometry as a simple screening tool to detect autonomic neuropathy in patients with diabetes: a pilot study. Biomed Eng Online 9 26
[7] Siegle G J, Steinhauer S R and Thase M E 2004 Pupillary assessment and computational modeling of the Stroop task in depression. Int J Psychol 72 307–13
[8] Fotiou D F, Stergiou V, Tsipotsios D, Lathari C, Nakou M and Karlovatisotou A 2009 Cholinergic deficiency in Alzheimer’s and Parkinson’s disease: evaluation with pupillometry. Int J Psychol 73 143–9
[9] Kardon R, Anderson S C, Damarjian T G, Grace E M, Stone E and Kawasaki A 2011 Chromatic pupillometry in patients with retinitis pigmentosa. Ophthalmology 118 376–81
[10] Link B, Jünemann A, Rix R, Sembritzki O, Brenning A, Korth M and Horn F K 2006 Pupillographic measurements with pattern stimulation: the pupil’s response in normal subjects and first measurements in glaucoma patients. Invest Ophth Vis Sci 47 4947–55
[11] Stanners R, Coulter M, Sweet A and Murphy P 1979 The pupillary response as an indicator of arousal and cognition Motiv Emotion 3
[12] Wilhelm B J, Widmann A, Durst W, Heine C and Otto G 2009 Objective and quantitative analysis of daytime sleepiness in physicians after night duties Int J Psychol 72 307–13
[13] Lowenstein O and Loewenfeld I E 1958 Electronic Pupillography: A New Instrument and Some Clinical Applications Arch Ophthalomol 59 352–63
[14] McLaren J W, Erie J C and Brubaker R F 1992 Computerized analysis of pupillograms in studies of alertness. Invest Ophth Vis Sci 33 671–6
[15] Lüdtke H, Wilhelm B, Adler M, Schaeffel F and Wilhelm H 1998 Mathematical procedures in data recording and processing of pupillary fatigue waves. Vision Res 38 2889–96
[16] Souza J K S, Ginani G E, Pompéia S, Baron J and Criollo C J T 2012 Identificação de privação de sono utilizando pupilômetro XXIII Congr Bras Eng Biomed 5

[17] Warga M, Lüdtke H, Wilhelm H and Wilhelm B 2009 How do spontaneous pupillary oscillations in light relate to light intensity? Vision Res 49 295–300

[18] Prasad B, Choi Y K, Weaver T E and Carley D W 2011 Pupillometric assessment of sleepiness in narcolepsy. Frontiers in Psychiatry / Frontiers Research Foundation 2 35

[19] Eggert T, Sauter C, Popp R, Zeithofer J and Danker-Hopfe H 2012 The Pupillographic Sleepiness Test in adults: effect of age, gender, and time of day on pupillometric variables. Am J Hum Biol 24 820–8

[20] Nakayama M, Yamamoto K and Kobayashi F 2012 Estimation of sleepiness using pupillary response and its frequency components Int J Bioinformatics Res Appl 8 342–65

[21] Nakayama M and Eye A 2008 Estimation of Sleepiness using Frequency Components of Pupillary Response BioCAS 357–60

[22] Melges D, Infantosi A and Miranda de Sá A 2012 Frequency-Domain Objective Response Detection Techniques Applied to Evoked Potentials: A Review Applied Biological Engineering - Principles and Practice (Rijeka, Croatia: InTech) pp 47–84

[23] Souza J K S, Ginani G, Pompéia S and Tierra-Criollo C 2013 Equipamento portátil para ensaios simultâneos de teste de vigilância psicomotora e pupilometria Rev Bras Eng Biomed 29 97–109

[24] Dinges D F and Powell J W 1986 Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations Behav Res Meth Ins C 17 652–5

[25] Infantosi A F C, Miranda de Sá A M F L and Sft A T S F 2007 A Statistical Test for Evaluating the Event-Related Synchronization / Desynchronization and its Potential use in Brain-Computer-Interfaces IV Latim Am C Biomed Eng 2 1122–6