Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 64 (2015) 879 – 884

Procedia Computer Science

Conference on ENTERprise Information Systems / International Conference on Project MANagement / Conference on Health and Social Care Information Systems and Technologies, CENTERIS / ProjMAN / HCist 2015 October 7-9, 2015

Electroencephalogram Cepstral Distances in Alzheimer's Disease Diagnosis

Pedro Miguel Rodrigues[a,]*, Diamantino Freitas[a] and João Paulo Teixeira[b]

[a] *Faculty of Engineering, University of Porto, Porto, Portugal*
[b] *Politechinic Institute of Bragança, Bragança, Portugal*

Abstract

Alzheimer's disease (AD) represents one of the greatest public health challenges worldwide nowadays, because it affects millions of people all over the world and it is expected that the disease will increase considerably in the near future. This study is the first application attempt of cepstral analysis on Electroencephalogram (EEG) signals to find new parameters in order to achieve a better differentiation between EEGs of AD patients and Control subjects. The results show that the methodology that uses a combined Wavelet (WT) Biorthogonal (Bior) 3.5 and cepstrum analysis was able to describe the EEG dynamics with a higher discriminative power than the other WTs/spectrum methodologies in previous studies. The most important significance figures were found in cepstral distances between cepstrums of theta and alpha bands ($p=0.00006<0.05$).

Peer-review under responsibility of SciKA - Association for Promotion and Dissemination of Scientific Knowledge

Keywords: Alzheimer's Diasease; Cepstrum; Wavelet Transform; Electroencephalogram signals; Cesptral distances;

1. Introduction

Alzheimer's disease (AD) largely strikes older adults and cannot be prevented or cured. AD is an irreversible neurodegenerative brain disorder and the most common cause of dementia in the elderly [1]. This is worrying

* Corresponding author. Tel.: Tel.: +351 96 891 67 85; fax: +351 22 508 14 40.
E-mail address: pedro.luis.rodrigues@fe.up.pt

Peer-review under responsibility of SciKA - Association for Promotion and Dissemination of Scientific Knowledge
doi:10.1016/j.procs.2015.08.600

because the elderly population will rise at a global level in the coming years. An early diagnosis is very important to determine the treatment and try to prevent the disease progression [1].

One of the most common signs of Alzheimer's disease, especially in the early stages, is forgetting recent information. It is therefore very important to pay attention to these warning signs, because dementia affects mainly the elderly.

The cause of AD is not yet known [2]. So far, no one single factor has been identified as being responsible to cause AD. It seems that a combination of factors, such as: age, genetic inheritance, environment, lifestyle, education, obesity, diabetes, hypertension, cholesterol, tobacco, alcohol lead to the appearance of AD [2].

The diagnostic accuracy is relatively low and there is not a biomarker able to detect AD without invasive tests. An autopsy or brain biopsy is the only way to make a definitive AD diagnosis. Effectively diagnosis can provide opportunities for AD patients to get involved in clinical trails and to get the best treatment. Moreover, beginning the treatment early can help to preserve a few cerebral functions [3].

In the last years, many research groups have started investigating the potential of electroencephalograms (EEGs) for AD diagnosis. For several decades the EEG has been used as a diagnostic tool for dementia as it is a very useful tool in the study of brain disorders being a non-invasive technique that records the electromagnetic fields produced by the brain activity.

EEG is a sensitive diagnostic test, non-invasive, economically accessible and of high temporal and spatial resolution. EEG is typically divided in different frequency bands, such as delta (δ, 1-4 Hz), theta (θ, 4-8 Hz), alpha (α, 8-13Hz), beta (β, 13-30 Hz), and gamma (γ, >30 Hz). AD seems to affect the signal power in those different bands. The major effect is known as the EEG "slowing", that means an increase of power in low frequency bands such as delta and theta, and a decrease of power in higher frequency bands such as alpha and beta [3].

In this study the authors intend to assess whether a combination of wavelet-based and Cepstral analysis can be useful to describe the abnormalities in EEG activity associated with AD. Furthermore, they want to explore the diagnostic ability of several Cepstral distances parameters to distinguish between AD patients and control subjects.

2. Materials and Proposed Methodology

Twenty subjects participated in the study (10 as Controls and 10 with AD). EEGs were organized in segments without artefacts.

The proposed methodology is composed by four different steps. The first two steps are related to the EEG signal processing and the last two to the process of features extraction and statistical analysis. The methodology steps organization is illustrated in Fig. 1.

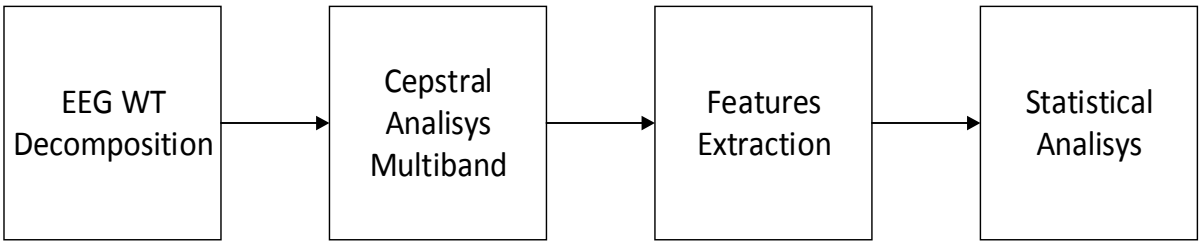

Fig. 1. Methodology description

2.1. EEG WT Decomposition

The Wavelet Transform (*WT*) was chosen to be the main technique to process the EEG signals. This signal processing tool provides a good time resolution for the high frequencies and good frequency resolution for low frequencies, which is useful for processing EEG signals because the target frequencies bands were localized in low frequencies.

The Discrete Wavelet Transform (*DWT*) was used in the implementation. Its multiresolution analysis consists in obtaining increasingly smaller signal resolution versions by successive filtering. The *DWT* uses a two functions set: a scale function ($\varphi[n]$) and a Wavelet function ($\psi[n]$) [4].

$$\varphi[n]=\sum_k h[k]\cdot\phi[2\cdot n-k] \tag{1}$$

$$\psi[n]=\sum_k g[k]\cdot\phi[2\cdot n-k] \tag{2}$$

where k is the discrete translation parameter and $h[k]$ and $g[k]$ respectively are the impulse responses of the lowpass and highpass filters used in the *WT* analysis. The signal decomposition in different frequency bands is achieved by successive low-pass filters and high-pass filters in time domain, followed by subsampling by factor of two until the maximum level of decomposition ($log_2(N)$) as is illustrated in the following equations 3 and 4 [4].

$$DWT_A^{\varphi}[j,k]=\sum_{n=0}^{\frac{N}{2^{(j-1)}}} DWT_A^{\varphi}[j-1,n]\cdot h[2\cdot k-n] \tag{3}$$

$$DWT_D^{\psi}[j,k]=\sum_{n=0}^{\frac{N}{2^{(j-1)}}} DWT_A^{\psi}[j-1,n]\cdot g[2\cdot k-n] \tag{4}$$

where, k=0, ..., N/(2^j); j=0, ..., $log_2(N)$, A represents the *DWT* approximated coefficients, D the *DWT* detail coefficients and N the signal length [4].

2.2. Cepstral Analysis - Multiband

Cepstral analysis is a signal analysis technique based on a homomorphic transformation called cepstrum. The cepstrum enables two or more signals deconvolution in the time-domain. It was proposed in 1963 by Bogert, Healy and Tukey for detecting echoes in seismic signals [5]. So, the cepstrum is useful to separate source and filter components which is one of the main reasons why it is highly applied for instance in speech signal processing.

The real cepstrum is defined as the inverse Discrete Fourier Transform (*DFT*) of the log magnitude of the signal's *DFT*.

$$c_i = c_j = IDFT\left(\log\left(\left|DFT\left(x_i(n)\right)\right|\right)\right) \tag{5}$$

where $i=j=\{\delta,\theta,\alpha,\beta,\gamma\}$, $x_i(n)$ represented the WT reconstructed EEG signal components in each EEG conventional band [5].

2.3. Features Extraction

In order to extract EEG signal information of AD patients and Control subjects, a global average of EEG signals were performed and the resulting signals, 10 AD patients signals and 10 Controls subjects signals, were decomposed down to the level 5 (correct level of EEG signal decomposition to reach the conventional frequency bands of EEG - δ, θ, α, β and γ bands) by different kinds of *WT* families.

The delta band corresponded to the approximated coefficients $A5$ of *DWT* decomposition, the theta band to the detail coefficients $D5$, the alpha band to $D4$, the beta band to $D3$ and the gamma band to $D2$. Each EEG conventional band was reconstructed in the time domain by different king of *DWT* families with the same length of the original segment of EEG signal.

Afterward several cepstral distances (*CD*) were calculated as features from the cepstrum of each EEG signal reconstructed through the WT components that corresponded to the δ, θ, α, β and γ bands:

- The smoothed distance between two cepstra [6]:

$$CD1_{i,j} = l * \sqrt{\left(\left(c_i(1) - c_j(1)\right)^2 + p * \sum_{n=2}^{N} \left(c_i(n) - c_j(n)\right)^2 \right)}; \quad (6)$$

where, l and p are the normalization factors of the smoothed distance calculation process, N the cepstral signal length, c_i and c_j the cepstral signals and $i=j=\{\delta,\theta,\alpha,\beta,\gamma\}$;

- The smoothed distance between two cepstra without the first cepstral coefficient [6]:

$$CD2_{i,j} = l * \sqrt{\left(p * \sum_{n=2}^{N} \left(c_i(n) - c_j(n)\right)^2 \right)}; \quad (7)$$

- The Euclidian distance between two cepstral vectors [6]:

$$CD3_{i,j} = \sqrt{\sum_{n=1}^{N} \left(c_i(n) - c_j(n)\right)^2}; \quad (8)$$

- Weighted Euclidian distance between two cepstral vectors [6]:

$$CD4_{i,j} = \sqrt{\left(\sum_{n=1}^{N} w(n) * \left(c_i(n) - c_j(n)\right)^2 \right)}; where\ w = [1, N] \quad (9)$$

where, $w(n)$ is a vector of weights;

- Non-linear weighted distance of two cepstral vectors [6]:

$$CD5_{i,j} = \sqrt{\left(\sum_{n=1}^{N} \sqrt{w(n)} * \left(c_i(n) - c_j(n)\right)^2 \right)}; where\ w = [1, N] \quad (10)$$

- Exponentially weighted distance of two cepstral vectors [6]:

$$CD6_{i,j} = \sqrt{\left(\sum_{n=1}^{N} w(n) * w(n) * \left(c_i(n) - c_j(n)\right)^2 \right)}; where\ w = [1, N] \quad (11)$$

To calculate all this cepstral distances parameters it should be ensured that both compared cepstra (c_i and c_j) have the same dimensionality.

2.4. Statistical Analysis

The normality of the cepstral distance extracted parameters was assessed with the Kolmogorov-Smirnov test and the homoscedasticity with the Levene's test. Data distributions did not meet the hypotheses of parametric tests. Thus, differences between groups were analyzed using the Mann-Whitney U test ($p<0.05$). As the authors only have found statistical differences between AD and control subjects on $CD_{\theta,\alpha}$ and $CD_{\theta,\beta}$, the statistical analysis of results will be focused on those bands (see Table 1 and Table 2). In Table 1 and Table 2 the best p-value achieved for each feature are represented in bold.

Despite the highly significant differences found in *CD* between the theta and beta bands it can be seen that the lowest *p*-value was reached with *CD* between theta band and alpha band ($CD2_{\theta,\alpha}$ and $CD3_{\theta,\alpha}$). A possible hypothesis to explain those results is that in those bands the phenomenon of "EEG slowing" is more notable. This phenomenon involves a loss of neurotransmitter acetylcholine and appears in intermediate states of AD. It was proved again with this new method, that the methodology that involves Biorthogonal *WT* level 3.5 showed the highest performance to identify and emphasize the EEG activity of AD patients and with the cepstrum combination it was possible to achieve a lower *p*-value than in previous works [7,8,9], where it was used a different methodology based on *WT*. So, when *WT* and cepstrum are combined, the cepstrum emphasizes the differences found by *WT* in EEG signals of AD Patients and Control subjects. Despite that, the comparison should be carefully established because the authors had used a different dataset in those previous works [7,8,9]. It was also found that *CD*5 and *CD4* are not good parameters to identify AD, in spite of the p-value<0.05 with the methodology that involved the *WT* Coiflet6. The *CD2* and *CD*3 proved to be the best parameters to identify AD as they achieved the lowest *p*-values. Anyhow, also *CD1* and *CD6* can be used because also achieved a p-value<0.05 for most of the *WT* families, except *CD1* with Meyer *WT* family.

Table 1. *p*-values of statistical Analysis of cepstral distances between theta and alpha bands

WT family band reconstruction	**Features**					
	$CD1_{\theta,\alpha}$	$CD2_{\theta,\alpha}$	$CD3_{\theta,\alpha}$	$CD4_{\theta,\alpha}$	$CD5_{\theta,\alpha}$	$CD6_{\theta,\alpha}$
Biorthogonal 3.5	**0,00014**	**0,00006**	**0,00006**	N.S	N.S	**0,00007**
Symlet6	0,00459	0,00233	0,00233	N.S	N.S	0,00262
Coiflet6	0,00233	0,00145	0,00145	**0,01905**	N.S	0,00184
Meyer	0,01734	0,00961	0,00961	N.S.	N.S	0,00329
Daubechies6	0,00368	0,00061	0,00061	N.S	N.S	0,00128

N.S. – Not significant

Table 2. *p*-values of statistical Analysis of cepstral distances between theta and beta bands

WT family band reconstruction	**Features**					
	$CD1_{\theta,\beta}$	$CD2_{\theta,\beta}$	$CD3_{\theta,\beta}$	$CD4_{\theta,\beta}$	$CD5_{\theta,\beta}$	$CD6_{\theta,\beta}$
Biorthogonal 3.5	**0,00041**	**0,00014**	**0,00014**	N.S.	N.S	**0,00047**
Symlet6	0,00634	0,00294	0,00294	N.S.	N.S	0,00262
Coiflet6	0,00233	0,00128	0,00128	**0,03896**	N.S	0,00207
Meyer	N.S.	0,03896	0,03896	N.S.	N.S.	0,00368
Daubechies6	0,00634	0,00262	0,00262	N.S.	**0,01175**	0,00459

N.S. – Not significant

3. Conclusion

So far there is no cure for Alzheimer's disease and patients often start the treatment when they already have significant symptoms. Therefore, a timely, effective and correct diagnosis can greatly help to reduce the disease evolution effects.

This study represents a further contribution in the search for new techniques of AD detection based on the EEG in order to increase its value as a diagnosis tool. Most of the literature focuses primarily on the frequency analysis of the EEG signal but this study has showed that the *WT* and cepstrum combination provided a good tool to obtain good features for AD detection through the EEG. In short, the authors of this study processed the global average of each EEG signal of AD patients and control subjects with the *WT* to get each reconstructed conventional band, from that signal, respectively. After that, each conventional band was processed with a cepstrum technique to obtain a cepstral distance between that conventional bands. Significant differences were found in *CD* between theta band and beta band and in *CD* between theta band and alpha band, namely *CD1*, *CD2*, *CD3* and *CD6*. It was concluded that two features $CD2_{\theta,\alpha}$ and $CD3_{\theta,\alpha}$ provided a high discrimination between AD patients and controls subjects (p=0.00006<0.05).

With this study it was revealed that the cepstrum representation can be a good tool for EEG signal processing in order to help in AD identification.

References

[1] Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E. Alzheimer's disease. The Lancet 2011; **377**:1019-1031.
[2] Jonker C, Launer L, Hooijer C, Lindeboom J. Memory complaints and memory impairment in older individuals. J Am Geriatr Soc 1996; **44**:44-49.
[3] Jeong J. Eeg dynamics in patients with alzheimer's disease. Clin. Neurophysiol 2004; **115**:1490-1505.
[4] Rioul O, Vetterli, M. Wavelets and signal processing. IEEE Signal Processing Magazine 1992; **8**:14-38.
[5] Bogert B.P., Healy M.J.R., Tukey J.W. The Quefrency Analysis of Time series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum, and Saphe Cracking, in Time Series Analysis, by M. Rosenblatt (Ed.), Wiley, NY 1963, 209-243.
[6] Junqua J. C., Wakita H. and Hermansky H. Optimizing Perceptually Based ASR Front End, IEEE Transactions on Speech and Audio Processing 1993, **1**:39-49.
[7] Rodrigues P, Teixeira J. Artificial neural networks in the discrimination of alzheimer's disease. Communications in Computer and Information Science 2011; **221**:272-281.
[8] Rodrigues P, Teixeira J, Hornero R, Poza J, Carreres A. Classification of alzheimer's electroencephalograms using artificial neural networks and logistic regression. Japan - Portugal Nano-Biomedical Engineering Symposium 2011; **1**:33-34
[9] Rodrigues P.M., Teixeira J.P. Alzheimer's Disease Recognition with Artificial Neural Networks, In Information Systems and Technologies for Enhancing Health and Social Care, by R. Martinho, R. Rijo, M. Cruz-Cunha and J. Varajão (Eds.), IGI Global 2013,102-118, 2013.