Stationary Epoch-based Entropy Estimation for Early Diagnosis of Alzheimer's Disease

N. Houmani[1], F. B. Vialatte[1], C. Latchoumane[2], J. Jeong[3], G. Dreyfus[1]
[1] ESPCI Paris Tech, Laboratoire SIGMA, 10 rue Vauquelin 75005 Paris, France.
[2] Center for Neural Science, Korea Institute of Science and Technology, 39-1 Hawlgok-Dong, Seongbuk-Gu, Seoul 136-791, South Korea.
[3] Korea Advanced Institute of Science and Technology, Department of Bio and Brain Engineering, Daejeon, 305-701, South Korea.

Abstract- Several studies showed that EEG signal of Alzheimer's disease patients is less complex than that of healthy subjects. In this article, we propose to characterize the complexity of the EEG signal by an entropy measure based on local density estimation by a Hidden Markov Model. We first show that this measure leads to consistent results qualitatively and quantitatively (in terms of classification accuracy). Indeed, it discriminates AD patients, at an early stage of Alzheimer's disease, from healthy subjects: a classification accuracy of 80% is reached on a dataset including EEG data recorded in different conditions. Based on this measure, we also show that parietal and temporal regions are the first regions affected by complexity loss in the early stage of Alzheimer's disease.

Keywords - EEG signal; Complexity measure; Stationary epochs; Entropy; HMM; Alzheimer's disease.

I. INTRODUCTION

Alzheimer's disease (AD) is a neuro-degenerative disease characterized by a progressive and irreversible brain disorder. It is the main cause of dementia in western countries, affecting 5-10% of the population above the age of 65 [1]. Moreover, as the life expectancy increased significantly in western countries in the last decades, the number of patients is expected to reach 115 million in 2050 [2].

The causes of AD are not clear; however it is characterized by a widespread neuronal cell destruction, cortical atrophy, intracellular deposition of neurofibrillary tangles, and extracellular deposition of senile plaques, especially in the hippocampus and cerebral cortex [3]. Currently, there is no treatment that can reverse the symptoms of the disease but many studies have indicated that therapeutic interventions at the early stage of AD can slow down the evolution of the disease [4,5]. Additionally, a reliable early diagnosis of AD enables a person with AD and his/her family to receive help in understanding this type of dementia and to take appropriate steps for the future. Therefore, a reliable early diagnosis of AD becomes an important issue for the scientific community.

In recent years, several researchers exploited the electroencephalogram (EEG) as a potential tool for diagnosing AD. Nevertheless, diagnosing AD in EEG signals remains a challenging issue as most existing techniques do not lead to a reliable diagnosis [6,7]. Several studies have shown that one of the major effects of AD is the reducing in complexity of the EEG signal compared to that of healthy subjects [8]. However, it is not always easy to detect such effects because of the large inter-variability that exists between AD patients.

Many methods were proposed to estimate the complexity of the EEG time series as a potential marker of early AD diagnosis. The correlation dimension [9] and the Lyapunov exponent [10] were frequently used. It was demonstrated that AD patients exhibit lower values of such measures than healthy subjects. The complexity of the signal was also measured using the fractal dimension [11]. The "complexity" or "irregularity" of the signal was also assessed by different measures stemming from information theory [12-20]: sample entropy [12,13], Tsallis entropy [14], approximate entropy [15,16], multiscale entropy [17], mutual information [18] and Lempel-Ziv complexity [19]. Such measures were shown to be useful in the analysis of EEG activity in AD patients. However, they were computed on the whole EEG signal without addressing the problem of its "*non-stationarity*". Some studies inferred that EEG time series are "*quasi-stationary*" [20,21]: the authors in [20] claimed that the EEG signal could be modeled as a sequence of quasi-stationary segments ("*epochs*") separated by rapid transitions. Also, in [21], the author suggested that perception is based on sequences of stationary patterns demarcated by discontinuities.

In this study, we attack the problem of quantifying the complexity of the EEG signal by means of a refined entropy measure computed on "*stationary*" epochs. The concept of entropy was introduced by Claude Shannon [22] in 1948, and is often referred to as "Shannon entropy". It is widely used in physics and information theory. In information theory, entropy measures the "uncertainty" related to a random variable, relying on its distribution. In physics, entropy is a measure of "chaos" or "disorder": higher entropy is often associated with more randomness and less system order. Thus, the more complex the signal, the higher its associated entropy measure.

Different estimators have been introduced to quantify the entropy of time series. In this work, we propose an entropy measure, which is derived from a refined characterization of the local statistical properties of the EEG signal by means of a Hidden Markov Model (HMM) [23]. We show that a better discrimination between early-stage AD patients and healthy

subjects can be performed by characterizing the EEG signal locally (at the "epoch" level).

The remainder of the paper is organized as follows. In Section 2, the proposed complexity measure is presented after a brief recall of Shannon Entropy. In Section 3, we describe the EEG databases used for experiments. In Section 4, we analyze and discuss the obtained results. Finally, conclusions are stated in Section 5.

II. COMPLEXITY MEASURES

In this section, we describe our proposed entropy measure after a brief recall on the computation of Shannon entropy.

A. Shannon entropy

Shannon entropy of a discrete random variable Z with N possible outcomes z_i, $i = 1 \ldots N$, is defined by:

$$H(Z) = -\sum_{i=1}^{N} p_i . \log_2(p_i) \quad (1)$$

where p_i is the probability mass of outcome z_i with $\sum_i p_i = 1$. Entropy is thus a measure of how uniformly distributed the random variable Z is across its possible outcomes: if there are N outcomes, then entropy is maximized at $log_2(N)$ with all $p_i = 1/N$. Base 2 logarithms are used, in order to express entropy in bits. Shannon entropy is the number of bits on average required to describe the random variable.

B. The proposed entropy-based complexity measure

The entropy measure that we propose is based on the assumption that the EEG signal is piecewise stationary, i.e. can be viewed as being stationary at the time scale of an epoch. In this framework, left-to-right Hidden Markov Models [23] (HMMs) are good candidates for performing the task, because they can segment the EEG signal into epochs (states of HMMs) and perform a local estimation of the probability density on each epoch. The states of HMMs will correspond to the stationary parts of the EEG signal, and the transitions of HMMs will correspond to the variations of the signal.

We thus consider the EEG signal of a given subject as a succession of epochs, obtained by segmenting such a signal via the Viterbi algorithm [23] using the corresponding subject's HMM. Viterbi algorithm is a widely used algorithm in HMMs to find the best state sequence [23]. It can be viewed as a modified forward algorithm: instead of summing up the probabilities from all the different paths, we pick only the optimal path, called "Viterbi path".

We thus obtain as many epochs in each EEG signal as there are states in the subject's HMM. Then we consider each local observation (i.e. each point) in a given epoch S_i as the outcome of a random variable Z_i that follows a given probability mass function $P_i(z)$. Thus, a random variable is associated to each stationary epoch of the signal (Figure 1), and the entropy of the considered epoch is that of an ensemble of outcomes of Z_i :

$$H(Z_i) = -\sum_{z \in S_i} p(z).\log_2(p(z)) \quad (2)$$

The sampling period (typically 8ms) is small with respect to the epoch length (typically 250ms). Thus, although Z is a discrete variable, we take advantage of the continuous emission probability law estimated on each epoch by the HMM. In other words, we consider $p(z)$ as the value of the probability density function estimated at the outcome z. The density functions were modeled as mixtures of Gaussians.

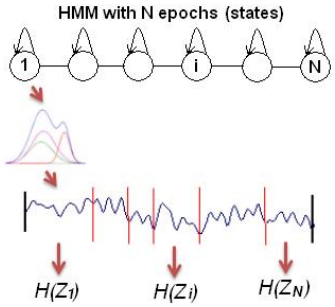

Figure 1: Epoch-based Entropy computation

Finally, by averaging the entropy over all the epochs of the EEG signal of the considered subject, we obtain an entropy-based complexity value, called in the following "Epoch-based Entropy".

III. DESCRIPTION OF DATABASES

Three datasets containing EEG recordings of healthy subjects and AD patients at rest and with closed-eyes conditions were used [24,25,26]. These datasets differ in terms of the number of used electrodes and the sampling frequencies.

A. Database A

This dataset was acquired at the Derriford Hospital, Plymouth, UK. It contains EEG data of 24 healthy subjects (aged 69.4 ± 11.5 years) and 17 diagnosed with mild form of AD patients (aged 77.6 ± 10 years). Patients underwent different neuroimaging and cognitive tests. EEG signals were recorded during 4 min at a sampling frequency of 256 Hz, later down-sampled to 128 Hz using 19 electrodes disposed according to the Maudsley System.

B. Dataset B

This dataset contains EEG data of five age-matched healthy subjects (aged 76.6 ± 5.6 years) and five AD patients (aged 78.8 ± 2.4 years). Patients were diagnosed with early stage, mild form AD according to NINCDS-ADRDA and DSM IV criteria; they underwent general medical, neurological and psychiatric tests. EEG signals were recorded during 1 min at a sampling frequency of 128 Hz using 21 electrodes disposed according to the 10-20 international system at the University of Malta.

C. Dataset C

This dataset is obtained from the Ecological University of Bucharest. It consists of three healthy subjects (aged 73.5 ± 2.2 years) and eight age-matched AD patients (aged 75 ± 3.4

years). Patients were diagnosed with a mild form of AD using psychometric tests, neuroimaging and clinical examinations. EEG signals were recorded during 10 to 20 min at a sampling frequency of 512 Hz using 22 electrodes disposed according to the international federation of clinical neurophysiology standards for digital recording of clinical EEG. In this work, EEG signals of this dataset were down-sampled to 128 Hz as done for dataset A and dataset B.

IV. Results and Discussion

In order to assess the effectiveness of the proposed entropy measure to discriminate AD patients from healthy subjects, we compare it in the following section to the classical Shannon entropy in terms of classification accuracy.

A. Epoch-based Entropy vs. Shannon Entropy

For comparison purposes, we compute both Shannon Entropy and Epoch-based Entropy associated to the EEG signal of each subject from the considered dataset. Then, classification performance is computed directly on the whole dataset by comparing the considered entropy value of each person to a decision threshold.

Table I presents the classification accuracies with Shannon Entropy and Epoch-based Entropy for the three datasets. As Receiver Operating Characteristic Curve (ROC) analysis was used to compare the two entropy measures in terms of classification accuracy (see Figure 2), we also report in Table 1 the area under the curve (AUC) values.

TABLE I. Classification accuracies (in %) and AUC values on each dataset with Shannon Entropy and Epoch-based Entropy

Datasets	Shannon Entropy		Epoch-based Entropy	
	Accuracy	AUC	Accuracy	AUC
Set A	73.2%	0.739	87.8%	0.914
Set B	90%	0.880	90%	0.960
Set C	81.8%	0.333	81.8%	0.583

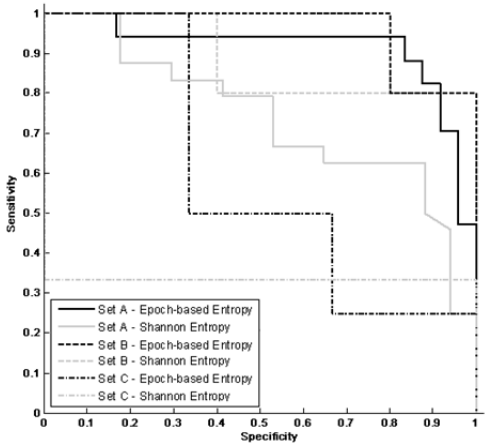

Figure 2: ROC Curves on each dataset with Shannon Entropy and Epoch-based Entropy.

Results show that the proposed entropy outperforms Shannon Entropy in terms of discrimination between AD patients and healthy subjects. This can be explained by the fact that even if our proposed entropy measure is averaged over all the epochs of the EEG signal, it is derived from a refined characterization of the local statistical properties of the EEG signal by means of Hidden Markov Models.

Figure 3 displays examples of signals with different complexities and different values of Epoch-based Entropy. This entropy measure indeed reflects the complexity of the signal: the more complex the signal, the higher its associated entropy value.

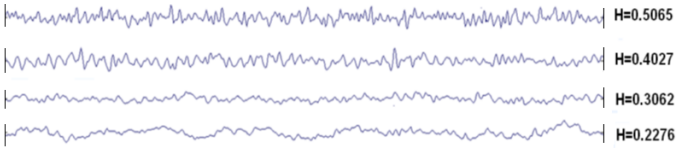

Figure 3: Examples of signals with differents complexities and their corresponding Epoch-based Entropy values.

B. Classification performance with Epoch-based Entropy

In this section, we assess classification performance with Epoch-based Entropy, following a consistent protocol that is more adapted for real clinical applications: we use a development subset containing EEG data of 10 AD patients and 10 healthy subjects belonging to Set A. The entropy values associated to these 20 subjects are computed. By averaging such values per class, we obtain two "Entropy-Prototypes": one associated to the class of AD patients and the other to the class of healthy subjects.

For the test subset, we consider the remaining 17 data of Set A, and in addition the data of Set B and Set C to analyze the ability of the entropy-based classifier to generalize to data that are independent from the development subset. For each subject in the test, we compute the entropy value of his/her EEG signal. Then for classification decision, each subject is associated to the class of the nearest Entropy-Prototype using the Euclidean distance. In order to obtain a reliable classification performance independent from the chosen development subset, five random samplings are carried out on healthy subjects and AD patients of the development subset. We carried out this experiment per brain region (Temporal, Parietal, Frontal and Occipital regions), as these regions are not affected by AD in the same way [27].

For all the subjects, the HMM was trained with 80 states (epochs) and 7 Gaussians per epoch. These values, chosen empirically, are those leading to the best performance classification on the development subset.

Tables II presents the classification accuracies per brain region on the three test subsets (the remaining 17 data of Set A (Set A'), the 10 data of Set B and 11 data of Set C) taken separately or simultaneously. The accuracies are computed considering the five random samplings of the development subset. Results show that best performance is obtained with temporal and parietal regions: a classification accuracy of 80% is reached when considering data of Set A', Set B and Set C simultaneously. This is consistent with the literature [27]: parietal and temporal regions are the first affected regions in the early stages of Alzheimer's disease.

TABLE II. CLASSIFICATION ACCURACY (IN %) WITH EPOCH-BASED ENTROPY PER BRAIN REGION

Test Datasets	*Temporal*	*Parietal*	*Frontal*	*Occipital*
Set A'	83.81%	82.85%	69.52%	63.81%
Set B	90%	80%	84%	80%
Set C	63.63%	78.18%	43.63%	61.81%
Set A' + Set B	85.81%	81.93%	74.19%	69.02%
Set A'+ Set B + Set C	80%	80.95%	66.19%	67.14%

V. CONCLUSION

We proposed an entropy-based complexity measure computed on stationary epochs, using a Hidden Markov Model that performs local density estimation at the epoch level. Such measure discriminates AD patients, at an early stage of Alzheimer's disease, from healthy subjects. The results are promising for early stage AD diagnosis: based on the proposed entropy measure, a classification accuracy of 80% is reached on a dataset including EEG data recorded in different conditions.

However, some limitations of our study deserve consideration. The proposed entropy measure was compared to the classical Shannon Entropy in terms of classification accuracy: we have shown that our entropy outperforms the classical one. However, other experiments, aiming at comparing such entropies, should be considered: studying the effect of noise, sampling frequency and windowing size on both entropies. This will be done in a more extended publication.

Finally, in this work, the entropy was computed locally at the epoch level, and then averaged over all the epochs. It would be interesting to keep the entropy values per epoch; thus each EEG signal would be associated to a sequence of entropy values. This would characterize more finely how the EEG signal fluctuates over time.

ACKNOWLEDGEMENT

N. Houmani thanks Pr. S. Garcia (Institut Mines-Telecom/ Telecom SudParis) for her significant help in the development of entropy-based quality measures during her PhD thesis.

REFERENCES

[1] C. P. Ferri, M. Prince M, C. Brayne, H. Brodaty, L. Fratiglioni, M. Ganguli, et al., "Global prevalence of dementia: a delphi consensus study". Lancet, vol. 366, no. 17, pp. 2012-2017, 2005.

[2] L. Li, D. Ruau, R. Chen, S. Weber, A.J. Butte, "Systematic identification of risk factors for Alzheimer's disease through shared genetic architecture and electronic medical records", Pac Symp Biocomput, pp. 224-235, 2013.

[3] H. Braak, E. Braak, "neuropathological stageing of Alzheime-related changes", Acta Neuropathol (Berl), vol. 82, pp. 239-259, 1991.

[4] M. Hashimoto, H. Kazui, K. Matsumoto, Y. Nakano, M. Yasuda, E. Mori, "Does donepezil treatment slow the progression of hippocampal atrophy in patients with Alzheimer's disease?", Am J Psychiatry, vol. 162, no. 4, pp. 676-682, 2005.

[5] C. Pietrzik, C. Behl, "concepts for the treatment of Alzheimer's disease: molecular mechanisms and clinical application", Int J Exp pathol, vol. 86, no. 3, pp.173-185, 2005.

[6] F. V. J. Dauwels, A. Cichocki, "Diagnosis of Alzheimer's Disease from EEG signals: where are we standing?", Current Alzheimer Research, vol. 7, no. 6, pp. 487-505, 2010.

[7] W. Woon, A. Cichocki, F. Vialatte, T. Musha, « Techniques for early detection of Alzheimer's disease using spontaneous EEG recordings », Physiological Measurement, vol. 28, no. pp. 335-347, 2007.

[8] J. Jeong, "EEG dynamics in patients with Alzheimer's disease," Clin. Neurophysiol., vol. 115, no. 7, pp. 1490-1505, 2004.

[9] P. Grassberger, "Generalized dimensions of strange attractors", physics Letters A, vol 97, no. 6, pp. 227-230, 1983.

[10] Y. B. Pesin, "Characteristic Lyapunov exponents and smooth ergodic theory", Russian Math. Surveys, vol 32, no. 4, pp. 55-114, 1977.

[11] B. B. Mandelbrot, "the fractal geometry of nature", W. H. Freeman and Comapgny.

[12] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability", Am. J. Physiol. Regul. Integr. Comp. Physiol., vol. 283, pp. R789-R797, 2002.

[13] D. Abasolo, R. Hornero, P. Espino, D. Alvarez, J. Poza, "Entropy analysis of the EEG background activity in Alzheimer's disease patients", Physiological Measurement, vol. 27, no. 3, pp. 241-253, 2006.

[14] T. De Bock, S. Das, M. Mohsin, N. B. Munro, L. M. hively, Y. Jiang, C. D. Smith, D. R. Wekstein, G. A. Jicha, A. Lawson, J. Lianekhammy, E. Walsh, S. Kiser, C. Black, "Early Detection of Alzheimer's Disease Using Nonlinear Analysis of EEG via Tsallis Entropy", Biomedical Sciences and Engineering Conference, 2010.

[15] D. Abásolo, R. Hornero, P. Espino, J. Poza, C. I. Sánchez, and R. de la Rosa, "Analysis of regularity in the EEG background activity of Alzheimer's disease patients with approximate entropy," Clin. Neurophysiol., vol. 116, no. 8, pp. 1826-1834, 2005.

[16] S. Pincus, "Approximate entropy as a measure of irregularity for psychiatric serial metrics," Bipolar Disord., vol. 8, no. 5 pt 1, pp. 430-440, Oct. 2006.

[17] J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López, "Analysis of electroencephalograms in Alzheimer's diseasepatients with multiscale entropy," Physiol. Meas., vol. 27, pp.1091-1106, 2006.

[18] J. Jeong, J. Gore, B. Peterson, "Mutual information analysis of the EEG in patients with Alzheimer's disease", Clinical Neurophysiology 112, pp. 827-835, 2001.

[19] D. Abásolo, R. Hornero, C. Gómez, M. García, and M. López,"Analysis of EEG background activity in Alzheimer's disease patients with Lempel-Ziv complexity and Central Tendency Measure," *Med. Eng. Phys*., vol. 28, pp. 315-322, 2006.

[20] A. Y. Kaplan, A. A. Fingelkurts, A. A. Fingelkurts, S. V. Borisov, B. S. Darkhovsky, "Nonstationary nature of the brain activity as revealed by EEG/MEG: Methodological, practical and conceptual challenges", Neuronal Coordination in the Brain: A Signal Processing Perspective, vol. 85, no. 11, pp. 2190-2212, 2005.

[21] W. J. Freeman, "A cinematographic hypothesis of cortical dynamics in perception", Int. J. of Psychophysiology, vol. 60, pp.149–161, 2006.

[22] T. M. Cover and J. A. Thomas, "Elements of Information Theory", Second Edition, John Wiley & Sons, 2006.

[23] L. Rabiner and B.H. Juang, "Fundamentals of Speech Recognition", Prentice Hall Signal Processing Series, 1993.

[24] C. Goh, E. Ifeachor, G. Henderson, C. Latchoumane, J. Jeong, C. Bigan, M. Besleaga, N. Hudson, P. Capotosto and S. R. Wimalaratna, "Characterisation of EEG at different stages of Alzheimer's disease (AD)", Clinical Neurophysiology, Vol. 117, pp. 138-139, 2006.

[25] C. F. Latchoumane , F. B. Vialatte, J. Solé-Casals, M. Maurice, S.R. Wimalaratna, N. Hudson, J. Jeong, A. Cichocki, "Multiway array decomposition analysis of EEGs in Alzheimer's disease", J Neurosci Methods, vol. 207, no. 1, pp. 41-50, 2012.

[26] M. Elgendi, B. Rebsamen, A.Cichocki, F. Vialatte, J. Dauwels, "Real-Time Wireless Sonification of Brain Signals", Advances in Cognitive Neurodynamics (III), pp. 175-181, 2013.

[27] A. Brun & E. Englund, "Regional pattern of degeneration in Alzheimer's disease: neuronal loss and histopathological grading", Histopathology, 5, 459-564, 1981.