

Clustering and Modeling of EEG Coherence Features of Alzheimer's and Mild Cognitive Impairment Patients

Kwaku Akrofi, *Student Member, IEEE*, Mary C. Baker, *Member, IEEE*, Michael W. O'Boyle,
Randolph B. Schiffer

Abstract— Using multiple discriminant analysis (MDA) and k-means clustering, coherence features extracted from the EEGs of a group of 56 subjects were analyzed to assess how feasible an automated coherence-based pattern recognition system that detects Alzheimer's disease (AD) would be. Sixteen of the subjects were AD patients, 24 were mild cognitive impairment (MCI) patients while 16 were age-matched controls. With MDA, an overall classification rate (CR) of 84% was obtained for AD vs. MCI vs. Controls classifications. The high CR implies that it is possible to distinguish between the three groups. The coherence features were also statistically analyzed to derive a neural model of AD and MCI, which indicated that patients with AD may have a greater number of damaged cortical fibers than their MCI counterparts, and furthermore, that MCI may be an intermediary step in the development of AD.

I. INTRODUCTION

COHERENCE features calculated from the Electroencephalogram (EEG) have been successfully used to distinguish Alzheimer's disease (AD) and/or mild cognitive impairment (MCI) patients from age-matched controls [1]–[3]. Specifically, AD patients show lower coherence than MCI patients and controls [3]–[5]. And, MCI patients also show reduced coherence relative to controls [1], [2].

The *magnitude squared coherence function* (or “coherence”) between two EEG channels x and y is defined as [9]:

$$C_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)},$$

where $P_{xy}(f)$ is the cross power spectral density of x and y , $P_{xx}(f)$ and $P_{yy}(f)$ are the power spectral densities of x and y , respectively, and f is the frequency in Hertz. Coherence, at a given frequency, is a measure of the synchrony between x and y ; the higher the synchrony, the higher the coherence. In EEG studies, the mean coherence over a given frequency range is a feature often used to evaluate differences between

patient groups.

AD is believed to be a “disconnection syndrome”, resulting in (or the result of) damage to long cortical fibers in the cortex of the brain [5]. Coherence can estimate the extent of this damage if the EEG channels roughly match the termination points of such a cortical fibers. In this case, a low coherence would indicate possible damage or “disconnection” of that fiber tract, which may also explain why AD patients often times exhibit lower EEG coherence.

MCI patients suffer memory or cognitive loss that is somewhat greater than controls, but not to the profound level of patients with AD or other forms of dementia [1]. MCI is considered a transitional state to AD, since as many as 25% of such patients per year develop AD [2]. While AD can develop from MCI, not all patients with MCI progress to AD. Hence, identification of MCI is vital to early AD detection.

In this paper, the mean EEG coherences of a group of AD, MCI, and age-matched controls are analyzed using a k-means clustering algorithm. Our aims were threefold: Firstly, we applied multiple discriminant analysis to the mean coherences before k-means clustering, this in an attempt to improve classification rates; Secondly, we provided a statistical analysis of the coherences and derived a model showing the extent of potential damage to selected cortical fibers; And lastly, we propose a model outlining possible relationships between AD and MCI in terms of their EEG coherence patterns.

II. PROCEDURES

A. Participants

56 subjects participated in the study. 16 were AD patients (6 males, 10 females, mean age 75.07 years, 15 right-handed and 1 left-handed), 24 MCI patients (10 males, 14 females, mean age 75.36 years, all right-handed), and 16 controls (6 males and 10 females, mean age 75.53 years, all-right-handed). All volunteered to participate in the study and gave written consent as part of their involvement in the Memory Clinic at Texas Tech University Health Sciences Center.

B. EEG Details

For all EEG recordings, 16 electrodes were positioned according to the 10-20 international system. The chosen electrode sites were: Fp1, Fp2, F7, F3, F4, F8, T3, C3, C4, T4, T5, P3, P4, T6, O1 and O2. Impedance of the electrode-skin interface was maintained below 5k Ω at each electrode

Manuscript received April 7, 2008.

K. Akrofi is a graduate student at Texas Tech University (TTU) in Electrical and Computer Engineering (ECE) Department, Lubbock, TX 79404 USA (Phone: 806-742-3533; e-mail: kwaku.akrofi@ttu.edu).

M. C. Baker is with the TTU ECE Dept., Lubbock, TX 79404 USA (e-mail: mary.baker@ttu.edu).

M. W. O'Boyle is with the TTU Human Development and Family Studies Dept., Lubbock, TX 79404 USA (e-mail: michael.oboyle@ttu.edu).

R. B. Schiffer is with the TTU Health Sciences Center, Department of Neuropsychiatry and Behavioral Sciences, Lubbock, TX 79430 USA (e-mail: psyrb@ttuhsc.edu).

site via conducting gel. In all, there were 16 channels (see Fig. 1 for details, e.g., the Fp1–F3 channel is labeled Channel 1 in the diagram and so forth), and were referenced in accordance with a longitudinal bipolar montage.

Each subject was asked to remain as still as possible and to keep their eyes closed while their EEG was recorded, using a Medelec Valor digital amplifier with a sampling rate of 256 Hz. A 30 sec artifact-free epoch was selected from each channel recording, filtered using a 0.5 to 50 Hz bandpass filter, and normalized to the total power present within a given channel. The latter helps to minimize any differences between amplifier and impedance settings across channels and participants. These filtered and normalized data were then used in all coherence calculations.

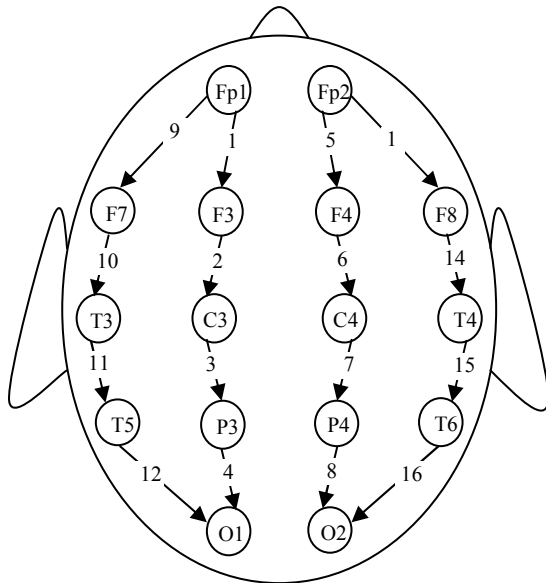


Fig. 1. 16-channel longitudinal bipolar montage.

III. FEATURE SELECTION AND CLUSTERING

A. Feature Selection

The coherences of selected pairs of channels were computed. These channel pairs are the same ones used in [5] and their locations are roughly the termination points of (projections of) key fiber tracts in the cortex of the brain [5]. Table I shows the list of channel pairs. The coherence of each was averaged over six frequency bands: 0.5 – 4, 4 – 8, 8 – 12, 12 – 16, 16 – 20 and 20 – 24 Hz. And as a result, there were 18 channel pairings across 6 frequency bands for a total of 108 EEG *features* for each subject.

To select relevant and statistically significant EEG features showing important differences between any two of the groups (AD, MCI and controls), the **Wilcoxon rank-sum test** [6] was used. This test assesses the significance of the differences between any two groups for a given feature, and Table II shows those features that were significantly different. The probability of such a difference being significant by chance alone was maintained at ≤ 0.03 .

For each of the sets of features in Table II, **multiple discriminant analysis** [7] was used to transform mean

TABLE I
CHANNEL PAIRINGS FOR BIPOLAR MONTAGE [1], [5].

Channel Electrodes	Channel No.	Channel Electrodes	Channel No.	Channel Pair (CP)
P3–O1	4	F3–C3	2	42
P3–O1	4	Fp1–F3	1	41
P3–O1	4	Fp1–T1	9	49
P4–O2	8	Fp2–F4	5	85
P4–O2	8	Fp2–T2	13	813
P4–O2	8	F4–C4	6	86
Fp1–F3	1	C3–P3	3	13
Fp1–F3	1	T5–O1	12	112
Fp2–F4	5	C4–P4	7	57
Fp2–F4	5	T6–O2	16	516
P3–O1	4	P3–C3	3	43
P3–O1	4	T3–T5	11	411
P4–O2	8	P4–C4	7	87
P4–O2	8	T4–T6	15	815
Fp1–F3	1	F3–C3	2	12
Fp1–F3	1	Fp1–F7	9	19
Fp2–F4	5	F4–C4	6	56
Fp2–F4	5	Fp2–F8	13	513

coherences into a vector space whose eigenvectors best highlight the differences between the groups; this technique produced a reduction in the feature set, specifically eliminating those with small eigenvalues.

B. K-means Clustering

A k-means clustering algorithm was applied to the three groups of features in Table II, yielding three classifications:

TABLE II
WILCOXON TEST RESULTS, SIGNIFICANCE LEVEL = 0.03.

Classes being compared	Frequency Bands			
	0.5 – 4 Hz	4 – 8 Hz	12 – 16 Hz	20 – 24 Hz
AD vs. Controls	41	49	86	13
	49		13	815
	813		57	12
	112		12	
	57			
	513			
AD vs. MCI	41		41	13
	49		13	19
	813		57	513
	112		513	
	19			
	513			
MCI vs. Controls	57	86		49

Entries in the table are features (channel pairing and frequency band) that show appreciable separation between AD and Controls, AD and MCI, and MCI and Controls.

AD vs. Controls, AD vs. MCI, and MCI vs. Controls. This process was repeated, this time with **multiple discriminant analysis** transformation of the features taking place before k-means clustering. Table III shows these classification rates.

Clustering algorithms search for “natural groupings” or “clusters” in the feature vectors. They do not require any training. The k-means algorithm [7], [8], starting with user-provided cluster mean vectors, iteratively calculates the mean vector of each cluster until there is little or no change in successively calculated mean vectors. This algorithm features fast convergence and provides high classification rates as long as the number of clusters is known, reasonable initial mean estimates are used, and the clusters are comparable in size and fairly well-separated.

We combined all the 21 features in Table II into one feature set and applied the k-means algorithm to perform a

TABLE III
TWO-CLASS CLASSIFICATION RESULTS

Two-Class Comparison	Classification Rates (CR)	
	Without MDA	With MDA
AD vs. Controls	AD	AD
	81.25% (13/16)	87.5% (14/16)
	Controls	Controls
	68.75% (11/16)	100% (16/16)
	Overall: 74.73%	Overall: 93.54%
AD vs. MCI	AD	AD
	81.25% (13/16)	87.5% (14/16)
	MCI	MCI
	58.33% (14/24)	87.5% (21/24)
	Overall: 68.84%	Overall: 87.5%
MCI vs. Controls	MCI	MCI
	62.5% (15/24)	66.67% (16/24)
	Controls	Controls
	93.75% (15/16)	93.75% (15/16)
	Overall: 76.55%	Overall: 79.06%

The CR is the percentage of the members of a class that were correctly classified. The overall CR percentage of all feature vectors within a comparison that were correctly classified, irrespective of class.

three-class pattern recognition. As done previously, we repeated this procedure, using the transformed feature vectors. The classification rates obtained are shown in Table IV. Note that just as in the two-class analyses, the

TABLE IV
THREE-CLASS CLASSIFICATION RESULTS

MDA Status	Classification Rates (CR)		
	AD	Controls	MCI
No MDA	81.25% (13/16)	43.75% (7/16)	50% (12/24)
	-3 misclassified as MCI	-5 misclassified as AD	-6 misclassified as AD
		-4 misclassified as MCI	-6 misclassified as Controls
		Overall: 56.22%	
With MDA	87.5% (14/16)	81.25% (13/16)	83.33% (20/24)
	-1 misclassified as MCI	-1 misclassified as AD	-1 misclassified as AD
	-1 misclassified as Controls	-2 misclassified as AD	-3 misclassified as Controls
		Overall: 83.99%	

introduction of the multiple discriminant analysis transformation brought about substantial improvement in classification performance.

IV. THE EID MODEL

Using the 4-step process below, we derived a model of the cortical fibers in the brain:

1. Calculate the median and inter-quartile range of each feature in Table II, for each class.
2. For every feature, normalize the three medians to that of the controls class to obtain a **relative coherence median** (RCM). All features of the Controls would have an RCM of 1.
3. Find the mean RCM of all channel pairings.
4. If the mean RCM of a channel pairing is greater than 1.1, give it an *E* rating. Give it a *D* rating if it is less than 0.9. Channel pairings with a mean RCM between 0.9 and 1.1 get a rating of *I*.

Each channel pairing is essentially a model of a cortical fiber. Hence the RCMs and ratings provide a measure of the damage to a set of these fibers. A fiber rated *E* has been enhanced by AD or MCI; One rated *I* has been left *intact*; and one rated *D* has been *damaged*. Note that all fibers in controls are rated *I*. Fig. 2 plots the fibers of the average AD and MCI patient. (Fig. 2 is not to scale regarding the anatomy of the fibers. For example, the fibers modeled by channel pairings 41 and 49 lie close to each other and are part of the **superior longitudinal fasciculus** [5], but have been drawn as separated in Fig. 2 to illustrate their differential ratings.)

V. DISCUSSION

Our aim for k-means clustering was to determine whether the coherence features as obtained from reasonably well-separated clusters representing AD, MCI and Control groups. Such a separation between the groups would suggest that EEG coherence can be used as the basis for a pattern recognition system that automatically and reliably detects the potential for AD in MCI patients.

Even without enhancing classification rates using multiple discriminant analysis, the overall rates of classifications in Tables III and IV were acceptable, ranging from 56% in the three-class group to nearly 77% for the two group MCI-Controls classification. Note that when the feature vectors were transformed, the overall classification rate increased to as high as 94% and never was below 79%. Thus, it is possible to distinguish EEG coherence feature vectors of AD patients, MCI patients, and controls using a fairly easily and relatively automatic detection algorithm.

The aforementioned Wilcoxon test found 14 features with significant separation between AD and Controls, and 13 for AD and MCI, but only 3 for MCI versus Controls. This suggests two things about the coherence features: Firstly, it is easier to distinguish between AD patients and controls,

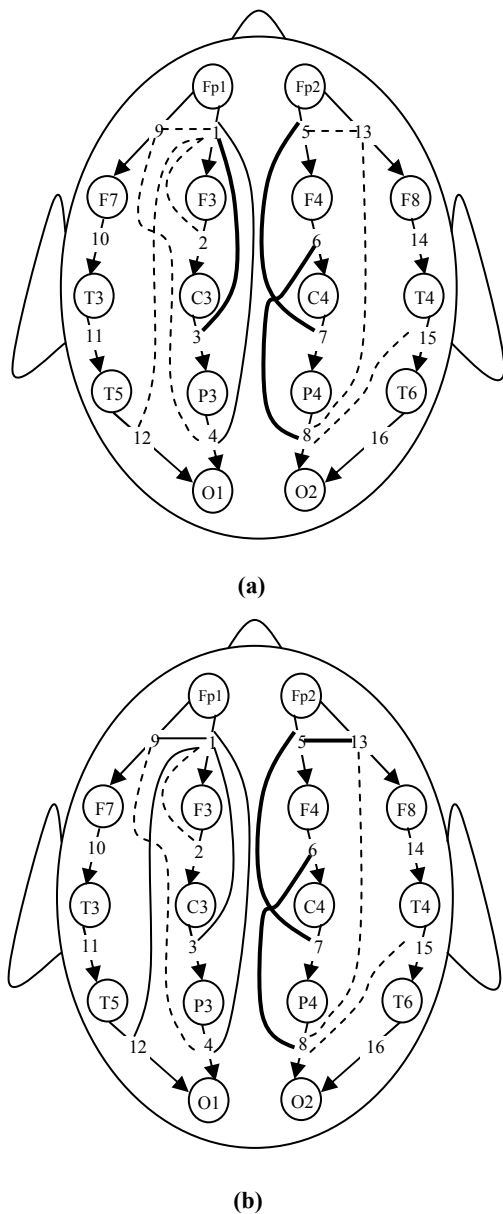


Fig. 2. EID model of the average (a) AD and (b) MCI patients. *I*-rated fibers are solid, *E*-rated ones thick and solid, and *D*-rated ones dashed.

and AD and MCI patients than it is to distinguish between MCI and controls. Table III confirms this result: note that with the multiple discriminant analysis transformation, the MCI-controls classification showed the lowest overall classification rate (79%). MCI patients consistently produced the lowest classification values during k-means clustering. Secondly, if it is difficult to tell MCI patients and controls, then they must be similar in terms of their EEG coherence which is an estimate of the integrity of their cortical fibers. This similarity implies that the latter have less damage to their fibers or have fewer fibers damaged as compared to AD patients. Our EID model supports this suggestion: note that Fig. 2a shows more *I*-rated fibers.

The EID model predicts some characteristics of AD and MCI patients. Firstly, *D*-rated fibers in MCI patients are also

D-rated in AD. This suggests that AD may indeed be a progression from the MCI state. Secondly, the differences between AD and MCI appear primarily on the left side of the brain, and currently, we do not fully understand the neurological importance of this finding. Lastly, the coherence of some fibers is enhanced by a given disease states, presumably as compensation for other damaged connections.

VI. CONCLUSION

K-means clustering of EEG coherence features obtained from a group of AD, MCI and age-matched controls yielded an overall classification rate of 66%. After multiple discriminant analysis transformation of the features, overall classification rates improved to 84%. This high percentage suggests that a reliable and automatic method of detecting AD and MCI patients based on pattern recognition of EEG coherence is indeed possible.

Statistical analysis of the coherence features led to the development of a model based on the integrity of cortical fibers in AD and MCI patients: AD patients had more damaged fibers than MCI patients. And, interestingly the EEG coherences of some fibers were enhanced in each group, presumably to compensate for the loss coherence in other cortical fibers. Further studies need to be conducted to verify such enhancement.

REFERENCES

- [1] J. S. Dugar, "Effects of dementia on EEG coherence," Master's thesis, Dept. Elect. Comp. Eng., Texas Tech Univ., Lubbock, TX, 2005.
- [2] Z.-Y. Jiang and L.-L. Zheng, "Inter- and intra-hemispheric EEG coherence in patients with mild cognitive impairment at rest and during working memory task," *JSUZ-B*, vol. 7, no. 5, pp. 357–364, July 2006.
- [3] H.-Y. Tao and T. Xin, "Coherence characteristics of gamma-band EEG during rest and cognitive task in MCI and AD," in *Proc. 27th Annu. IEEE-EMBC*, Shanghai, China, 2005, pp. 2747–2750.
- [4] J. Jeong, "EEG dynamics in patients with Alzheimer's disease," *Clin Neurophys*, vol. 115, pp. 1490–1505, July 2004.
- [5] A. F. Leuchter, T. F. Newton, I. A. Cook, D. O. Walter, S. Rosenberg-Thomas, and P. A. Lachenbruch, "Changes in brain functional connectivity in Alzheimer-type and multi-interfaced dementia," *Brain*, vol. 115, pp. 1543–1561, October 1992.
- [6] J. L. Devore, *Probability and Statistics for Engineering and the Sciences*. Belmont, CA: Brooks/Cole, 3rd Ed., 1991, pp. 608–623.
- [7] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*. New York, NY: John Wiley & Sons, 2nd Ed., 2001.
- [8] T. Kanungo, D. M. Mount, N. S. Netanyahu et al, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intelligence*, vol. 24, pp. 881–892, July 2002.
- [9] E. M. Glaser and D. S. Ruchkin. *Principles of Neurobiological Signal Analysis*. New York, NY: Academic Press, 1976, pp 168–175.