Wavelet analysis for EEG feature extraction in deception detection

Anna Caterina Merzagora, Scott Bunce, Meltem Izzetoglu and Banu Onaral

Abstract— Deception detection has important clinical and legal implications. However, the reliability of methods for the discrimination between truthful and deceptive responses is still limited. Efforts to improve reliability have examined measures of central nervous system function such as EEG. However, EEG analyses based on either time- or frequency-domain parameters have had mixed results. Because EEG is a nonstationary signal, the use of joint time-frequency features may yield more reliable results for detecting deception. The goal of this study was to investigate the feasibility of deception detection based on EEG features extracted through wavelet transformation. EEG was recorded from 4 electrode sites (F3, F4, F7, F8) during a modified version of the Guilty Knowledge Test (GKT) in 5 subjects. Wavelet analysis revealed significant differences between deceptive and truthful responses. These differences were detected in features whose frequency range roughly corresponds to the EEG beta rhythm and within a time window which coincides with the P300 component. These preliminary results indicate that joint time-frequency EEG features extracted through wavelet analysis may provide a more reliable method for detecting deception than standard ERPs.

I. INTRODUCTION

THE ability to detect deception has important legal, moral and clinical implications, and has recently received revived interest from the scientific community. Currently, the most widely used technique for the quantitative discrimination between deceptive and truthful responses is polygraphy, which relies on measures of autonomic nervous system response, i.e. emotional efferents. Traditional polygraphy, however, has been criticized for having unacceptable level of reliability [1]. Consequently, a number of other recording modalities have recently been investigated for the possible application to deception detection, including functional magnetic resonance (fMRI) [2]-[9] and electroencephalography (EEG) [10]-[12].

Until recently, the reported EEG studies in deception detection have primarily focused on the topography and time-domain analyses of the P300 component. However, there is still disagreement about the features that may best discriminate between deceptive and truthful responses [13];[14]. This study provides a preliminary investigation of the ability to use joint time-frequency features extracted using wavelets to differentiate truth from deception. EEG is known to be a nonstationary signal and its frequency components have been functionally related to information processing and behavior. Wavelet analysis capitalizes on these properties, and has been successfully utilized in a number of different EEG applications [15]-[18].

A primary emphasis in this study was to examine the capacity for physiological measures to differentiate among the cognitive elements of truth and deception, i.e. the knowledge that one is lying. The task was designed to elicit high motivation to escape detection, but to minimize participants' anxiety about being deceptive. To accomplish this end, the task was framed as a form of poker-like card game in which it is socially acceptable to "bluff," or to lie, minimizing feelings of anxiety about lying [19]. The task was also based on a modified Guilty Knowledge Test (GKT) because of the solid empirical evidence for its statistical reliability.

II. MATERIALS AND METHODS

Participants: The participants of this study were 5 right-handed individuals. They had no previous history of neurological or mental abnormalities and were non-smokers, as based on self-reports. Participants were asked to refrain from alcohol the night before the experiment. All participants signed an informed consent statement approved by the Drexel University College of Medicine Institutional Review Board.

Protocol: Participants were given a total of 5 cards, four of which (one from each suit) were face-up on the computer screen (the 'hand'). Participants were informed that the identities of these four face-up cards, as in some forms of poker, were known by the participants, as well as the researchers. Participants were then asked to choose a fifth card from among three sealed envelopes, each of which contained a playing card which they kept in their hand ('target' card) and $50. Participants were informed that only they knew the identity of this card, and the experimenter would be attempting to learn the identity of this card by alternately presenting a series of cards, asking the question "Do you have this card?", and examining their brain responses. They were told that if they were successful in concealing the identity of the card, that would be able to

A.C. Merzagora is with the School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA 19104 USA (phone: 215-8951988; fax: 215-8954983; e-mail: a.merzagora@drexel.edu).

S. Bunce is with Department of Psychiatry, Drexel University College of Medicine, Philadelphia. PA 19104 USA. (e-mail: sbunce@drexelmed.edu).

M. Izzetoglu is with the School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA 19104 USA (e-mail: meltem@cbis.ece.drexel.edu).

B. Onaral is with the School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA 19104 USA (e-mail: banu.onaral@drexel.edu).
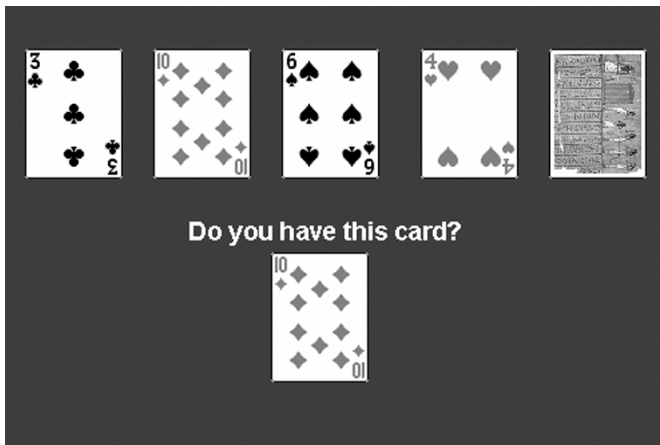
Fig. 1. Example of card presentation on the computer screen during the modified GKT.

keep the $50, in addition to their participation remuneration ($25).

Four categories of cards were presented to participants. The 'target' card (either the 5 of clubs or 2 of hearts) was presented 16 times. The "correct" response to the question "Do you have this card?" for this card was "no." The 'truth' card (either the 2 of hearts or 5 of clubs, reciprocal to the 'target' card) was also presented 16 times with a correct response of "no." 'Control' cards consisted of 4 presentations each of the 4 'hand' cards, with a correct response of "yes." The fourth category of cards presented were 'non-target' cards, single presentations of 16 other cards from the deck.

In Fig. 1 an example of presentation of a 'control' card is shown. Cards were presented for 3 s with an inter-stimulus interval of 12 s.

EEG system and preprocessing: EEG recordings were made using Ag-AgCl electrodes from 4 scalp positions in the prefrontal area, as assigned according to the International 10-20 System: F3, F4, F7, F8. Vertical and horizontal electrooculograms (VEOG and HEOG) were also recorded, simultaneously. Data were acquired using a Neuroscan Synamps amplifier running Acquire software. The data were bandpass filtered between 0.1-100 Hz, with a sampling rate of 500 Hz.

For each subject, EEG data were corrected for ocular artifacts based on VEOG and HEOG using Stim software (Neuroscan, Inc.). The recordings were then divided into 1200 ms epochs time locked to the stimulus presentation, with a 200 ms pre-stimulus baseline. Noisy trials were rejected through visual inspection. Each trial was baseline corrected using the mean of the 200 ms prestimulus. Averages of trials corresponding to 'target' and 'truth' cards were computed for each channel for each subject.

For the averages, P300 amplitude and latency were identified. Also, their spectra were computed and the peak values in the different EEG bands were obtained. These separate time- and frequency-domain features were obtained in order to compare their performances in deception detection with the performance of the features obtained through wavelet analysis.

Wavelet Analysis for feature extraction: In this study, quadratic B-spline wavelets were used in the wavelet analysis due to their near optimal time-frequency localization properties. Moreover, their waveform is similar to the waveforms to be detected in the EEG signal; hence extraction of EEG components is more likely to be successful.

In continuous time, B-spline functions of order n form a basis for the subspace of all piecewise continuous polynomial functions of degree n with derivatives up to $n-1$ that are continuous everywhere on the real axis. Specifically,, quadratic B-spline scaling function is obtained from

$$B_{(2)}(x) = \begin{cases} \frac{1}{2}x^2 & x \in [0,1) \\ -x^2 + 3x - \frac{3}{2} & x \in [1,2) \\ \frac{1}{2}x^2 - 3x + \frac{9}{2} & x \in [2,3) \\ 0 & otherwise \end{cases} \quad (1)$$

which is, in discrete time, $b_{(2)}(k) = B_{(2)}(k)$. Moreover, the quadratic B-spline function acts as a low pass filter and the related quadratic B-spline wavelet acts as a bandpass filter. With all these characteristics, B-spline functions can serve as a mother wavelet and hence the filter coefficients of the quadratic B-spline wavelet can be obtained and used for the multiresolution analysis.

The B-spline wavelet transform was computed for the averaged trials corresponding to the 'truth' or 'target' cards for each subject and each channel, separately .In this study, the signals were evaluated up to the 6th scale, so that some of the bands related to the WT detail coefficients (d) roughly corresponded to the well known EEG frequency components. In particular:

- d4 (~15–31 Hz) roughly corresponds to beta;
- d5 (~7.5–15 Hz) roughly corresponds to alpha;
- d6 (~3.75–7.5 Hz) roughly corresponds to theta;

Because going from one level of detail to the next lower one involved a downsampling by a factor 2, the number of WT coefficients between the two levels of detail decreased. In particular, d4 consists of 31 coefficients, d5 consists of 15 coefficients and d6 consists of 7 coefficients. Each of these WT coefficients sets spanned the post-stimulus segment of the averaged EEG trials.

III. RESULTS AND DISCUSSION

Analysis of features in the time-domain (P300 amplitude or latency, areas under the curves of the averages) did not reveal any significant difference between channels or responses. Frequency-domain features (peaks in the different EEG bands) did not show any differences as well.

WT coefficients in the EEG bands of interest (d4, d5 and d6 as explained in Section II) were analyzed for statistically significant differences. For each coefficient in these three bands, i.e. d4, d5 and d6, a repeated measures ANOVA was performed. The within group effect was the stimulus, with 2 levels ('target and 'truth' card). The between group effect was the EEG channel, with 4 levels (F3, F4, F7 and F8). The confidence level was set at 95%; the Bonferroni correction for the p-values was used.

Coefficients in the detail levels corresponding to lower frequency ranges (i.e. d5 and d6, which roughly correspond respectively to alpha and theta bands in EEG) failed to show any significant difference. No variation was observed between stimuli or in the topographic distribution of activation.

Also coefficients in d4 (~15–31 Hz) did not present any difference between channels. Nevertheless, 10 out of the 31 coefficients in d4 showed statistically significant differences between 'target' and 'truth' cards. Specifically, the coefficients that showed such differences were: d4_8, d4_10, d4_16, d4_17, d4_20, d4_21, d4_24, d4_27, d4_29, d4_31.

Fig. 2 shows the bar plots of the d4 coefficients for each channel. Values obtained from averages of 'target' and 'truth' card responses are compared (black bars for 'target' and gray bars for 'truth'); stars denote statistically significant differences at p<0.05. As observable from the plots, statistical significance is consistent across the 4 channels used in the experimental setup (F3, F4, F7, F8) and no lateralization is observed, thus suggesting similar involvement of the whole frontal region.

Fig.2 shows both the coefficient numbers (going from 1 to 31) and the time axis (in ms) spanned by the wavelet coefficients. As can be observed from the figure, significant differences in the beta waves (level d4) are located in time ranges around 300 ms, 500 ms, 650 ms and after 750 ms. This localization in time for the differences between deceptive and truthful responses suggests that dissimilarities in cognitive processes are concomitant with the elicitation of the P300 component, but also appear somewhat later in the EEG recordings.

IV. CONCLUSION

The aim of this preliminary study was to investigate the ability of wavelet domain features obtained from the EEG differentiate truth from deception during a low anxiety task. The results revealed that wavelet coefficients corresponding to beta waves were found to differentiate when subjects were telling the truth versus when they were lying. These results suggest that wavelet coefficients could potentially be used as new or additional features in order to detect deception using EEG. The need for joint time-frequency analysis was further substantiated by the fact that traditional time- and frequency-domain analyses failed to show any statistically significant difference between truthful and deceptive responses for this experimental protocol. In contrast, wavelet coefficients with a joint time-frequency distribution corresponding to the beta rhythm were able to discriminate truth from lie in time windows from 300 to 1000 ms post-
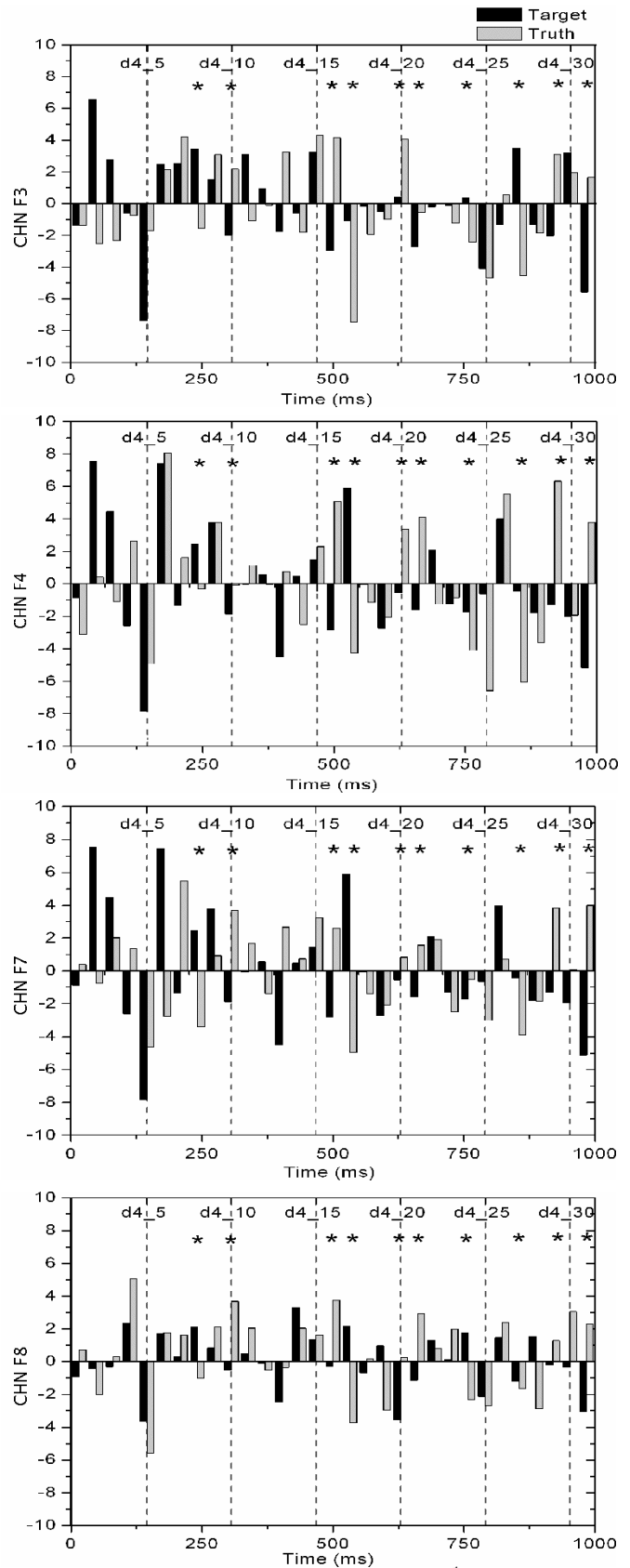

Fig. 2. Bar plots of the WT detail coefficients of the 4th level, which spans frequencies in the 15-31 Hz range. For each channel, coefficients of the averages for 'target' and 'truth' cards are shown (respectively in black and gray). Stars denote significant differences at 95% of confidence.

stimulus.

Based on these preliminary results, the application of WT-based features to deception detection appears promising. Larger sample sizes, along with more diverse protocols, will be necessary to examine generalizability of the discriminating features identified here. Additionally, longer response ranges as well as response-locked averages, could be investigated, as most of the significant differences in these analyses were found at longer latencies.

REFERENCES

[1] Office of Technology Assessment (1983). Scientific validity of polygraph testing: A research review and evaluation–A technical memorandum. U.S. Office of Technology Assessment, Washington, DC..

[2] D. D. Langleben, L. Schroeder, J. A. Maldjian, R. C. Gur, S. McDonald, J. D. Ragland, C. P. O'Brien, and A. R. Childress, "Brain activity during simulated deception: an event-related functional magnetic resonance study," *NeuroImage*, vol. 15, no. 3, pp. 727-732 2002.

[3] D. D. Langleben, J. W. Loughead, W. B. Bilker, K. Ruparel, A. R. Childress, S. I. Busch, and R. C. Gur, "Telling truth from lie in individual subjects with fast event-related fMRI," *Hum. Brain Mapp.*, vol. 26, no. 4, pp. 262-272, 2005.

[4] F. A. Kozel, K. A. Johnson, Q. Mu, E. L. Grenesko, S. J. Laken, and M. S. George, "Detecting Deception Using Functional Magnetic Resonance Imaging," *Biol. Psychiatry*, vol. 58, no. 8, pp. 605-613, Oct.2005.

[5] K. Luan Phan, A. Magalhaes, T. J. Ziemlewicz, D. A. Fitzgerald, C. Green, and W. Smith, "Neural correlates of telling lies: A functional magnetic resonance imaging study at 4 Tesla1," *Acad. Radiol.*, vol. 12, no. 2, pp. 164-172, 2005.

[6] J. M. Nunez, B. J. Casey, T. Egner, T. Hare, and J. Hirsch, "Intentional false responding shares neural substrates with response conflict and cognitive control," *NeuroImage*, vol. 25, no. 1, pp. 267-277, Mar.2005.

[7] F. B. Mohamed, S. H. Faro, N. J. Gordon, S. M. Platek, H. Ahmad, and J. M. Williams, "Brain Mapping of Deception and Truth Telling about an Ecologically Valid Situation: Functional MR Imaging and Polygraph Investigation--Initial Experience," *Radiology*, vol. 238, no. 2, pp. 679-688, Feb.2006.

[8] S. A. Spence, M. D. Hunter, T. F. D. Farrow, R. D. Green, D. H. Leung, C. J. Hughes, and V. Ganesan, "A cognitive neurobiological account of deception: evidence from functional neuroimaging," *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, vol. 359, no. 1451, pp. 1755-1762, 2004.

[9] K. Izzetoglu, G. Yurtsever, A. Bozkurt, B. Yazici, S. Bunce, K. Pourrezaei, and B. Onaral, "NIR spectroscopy measurements of cognitive load elicited by GKT and target categorization," 2003, p. 6.

[10] M. Soskins, J. P. Rosenfeld, and T. Niendam, "Peak-to-peak measurement of P300 recorded at 0.3 Hz high pass filter settings in intraindividual diagnosis: complex vs. simple paradigms," *International Journal of Psychophysiology*, vol. 40, no. 2, pp. 173-180, Mar.2001.

[11] J. P. Rosenfeld, J. W. Ellwanger, K. Nolan, S. Wu, R. G. Bermann, and J. Sweet, "P300 scalp amplitude distribution as an index of deception in a simulated cognitive deficit model," *Int. J Psychophysiol.*, vol. 33, no. 1, pp. 3-19, July1999.

[12] J. P. Rosenfeld, A. M. Reinhart, M. Bhatt, J. Ellwanger, K. Gora, M. Sekera, and J. Sweet, "P300 correlates of simulated malingered amnesia in a matching-to-sample task: topographic analyses of deception versus truthtelling responses," *Int. J Psychophysiol.*, vol. 28, no. 3, pp. 233-247, May1998.

[13] J. P. Rosenfeld, M. Soskins, G. Bosh, and A. Ryan, "Simple, effective countermeasures to P300-based tests of detection of concealed information," *Psychophysiology*, vol. 41, no. 2, pp. 205-219, 2004.

[14] A. R. Miller, C. Baratta, C. Wynveen, and J. P. Rosenfeld, "P300 latency, but not amplitude or topography, distinguishes between true and false recognition," *J Exp. Psychol. Learn. Mem. Cogn*, vol. 27, no. 2, pp. 354-361, Mar.2001.

[15] R. Polikar, M. H. Greer, L. Udpa, and F. Keinert, "Multiresolution wavelet analysis of ERPs for the detection of Alzheimer's disease,", 3 ed 1997, pp. 1301-1304.

[16] V. J. Samar, K. P. Swartz, and M. R. Raghuveer, "Multiresolution analysis of event-related potentials by wavelet decomposition," *Brain Cogn.*, vol. 27, no. 3, pp. 398-438, 1995.

[17] T. Demiralp and A. Ademoglu, "Decomposition of event-related brain potentials into multiple functional components using wavelet transform," *Clin. Electroencephalogr.*, vol. 32, no. 3, pp. 122-138, 2001.

[18] T. Demiralp, A. Ademoglu, M. Comerchero, and J. Polich, "Wavelet Analysis of P3a and P3b," *Brain Topography*, vol. 13, no. 4, pp. 251-267, June2001.

[19] D. T. Lykken, "Why (some) Americans believe in the lie detector while others believe in the guilty knowledge test," *Integr. Physiol Behav. Sci*, vol. 26, no. 3, pp. 214-222, July1991.