

Detection of spontaneous class-specific visual stimuli with high temporal accuracy in human electrocorticography

Kai J. Miller, *Member, IEEE*, Dora Hermes, Gerwin Schalk, *Member, IEEE*, Nick F. Ramsey, Bharathi Jagadeesh, Marcel den Nijs, Jeffrey G. Ojemann, Rajesh P.N. Rao

Abstract—Most brain-computer interface classification experiments from electrical potential recordings have been focused on the identification of classes of stimuli or behavior where the timing of experimental parameters is known or pre-designated. Real world experience, however, is spontaneous, and to this end we describe an experiment predicting the occurrence, timing, and types of visual stimuli perceived by a human subject from electrocorticographic recordings. All 300 of 300 presented stimuli were correctly detected, with a temporal precision of order 20ms. The type of stimulus (face/house) was correctly identified in 95% of these cases. There were ~20 false alarm events, corresponding to a late 2nd neuronal response to a previously identified event.

I. INTRODUCTION

NEURONAL populations in early visual cortex process different classes of visual information in the same cortical surface area, dependent on the spatial location of the stimulus in the visual field, rather than the class of visual input. Higher order inferotemporal visual areas instead represent different classes of visual stimuli with distinct cortical surface regions. For example, face and house visual stimuli are represented in distinct, adjacent inferotemporal areas [1]. Using ECoG, we measured from both of these visual areas simultaneously and attempted to spontaneously identify visual stimuli and assign them to one of two classes, faces or houses. Previous brain surface ECoG studies have demonstrated class specific representation (particularly faces) by averaging raw potential measurements from many within-class visual stimulus presentations [2]. Rather than examine fluctuations in the raw potential, we employed a new technique for extracting broadband temporal power spectral estimates which recently has been shown to capture robust local neuronal population activity [3, 4], and demonstrate that it can be used to capture, with high temporal fidelity, the cortical responses to spontaneous face or house visual stimuli.

Manuscript received April 20, 2009. This work was supported by NSF Grants 0622252 & 0130705. Contact: kjmiller@u.washington.edu
 KJ Miller (Neurobiology & Behavior, Physics), Bharathi Jagadeesh (Physiology and Biophysics), Marcel den Nijs (Physics), Jeffrey G Ojemann (Neurological Surgery) and Rajesh PN Rao (Computer Science and Engineering) are with the University of Washington, Seattle, WA 98195. Dora Hermes and Nick Ramsey are at the Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, department of Neurology and Neurosurgery, section of Brainfunction and Plasticity, Utrecht, The Netherlands. Gerwin Schalk is with the Wadsworth Institute, New York State Department of Health, Albany, NY.

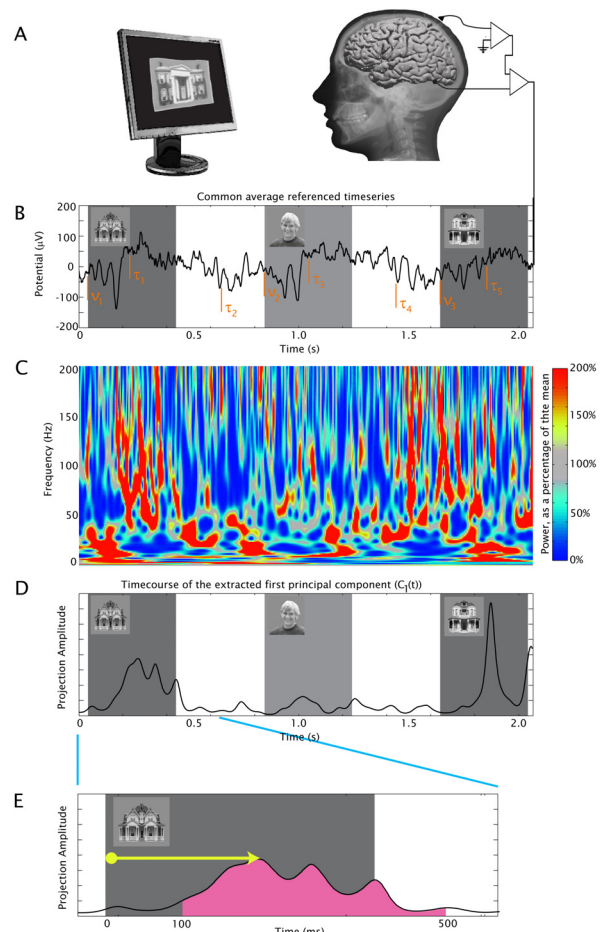


Fig. 1. Electroencephalographic arrays were placed on the subtemporal cortical surface of an epilepsy patient (A). Simple, luminance and contrast matched, grayscale faces and houses that were displayed in random order for 400 ms each, with inter stimulus intervals of 400 ms. The ECoG signal was re-referenced to the common average (B) and for each electrode 1 second epochs centered at the middle of each stimulus presentation, τ_q , were selected to calculate the power spectral density (PSD; $P(f, n, \tau_q)$). Based upon many samples of the PSD, a covariance matrix between frequencies was estimated, diagonalized using a principle component technique to isolate broadband spectral change. As shown in (C), a continuous time-frequency approximations of the PSD, $\tilde{P}(f, n, t)$, was calculated using a wavelet approach. The continuous time frequency approximation was then projected into the 1st principal component (corresponding to broadband change), associated with the time-varying neuronal population activity, $C_1(t)$. (D). Following an estimated (v'_k) or actual (v_k) stimulus event time, the total response magnitude (R_k or R'_k , in pink) could be estimated, as shown in (E).

II. METHODS

A. Experimental setting

A 32 year old male epileptic patient was implanted with a surface electrode array in his left hemisphere (Fig 1) to localize a seizure focus prior to surgical resection. While this electrode array was in place, he participated in a visual experiment. Pictures of simple, luminance and contrast matched, grayscale faces and houses were displayed in random order for 400ms each, with 400ms inter-stimulus interval between (Fig 1). He was asked to report a simple target (a single upside-down house), while the electrical potential was measured from each platinum electrode (each has 2.3mm diameter exposed, separated by 1cm. *Ad Tech*, Racine WI), with location determined from x-ray [5].

B. Signal processing

The ECoG potentials were measured with respect to an electrode on the scalp, and then were re-referenced with respect to the common average reference across all electrodes. From each electrode, n , samples of power spectral density (PSD; $P(f, n, \tau_q)$) were calculated from 1 second epochs of the raw potential, $V(n, t)$ (Fig 1), centered at the middle of each stimulus presentation, τ_q . These $P(f, n, \tau_q)$, were normalized in two steps: each spectral sample was divided by the average PSD, and then the log was taken.

A principal component method was applied to these normalized PSD samples to identify motifs of spectral change [3]. For each electrode, the eigenvalues, λ_k , and eigenvectors, \vec{e}_k , of a covariance matrix $C(f, \tilde{f})$ between

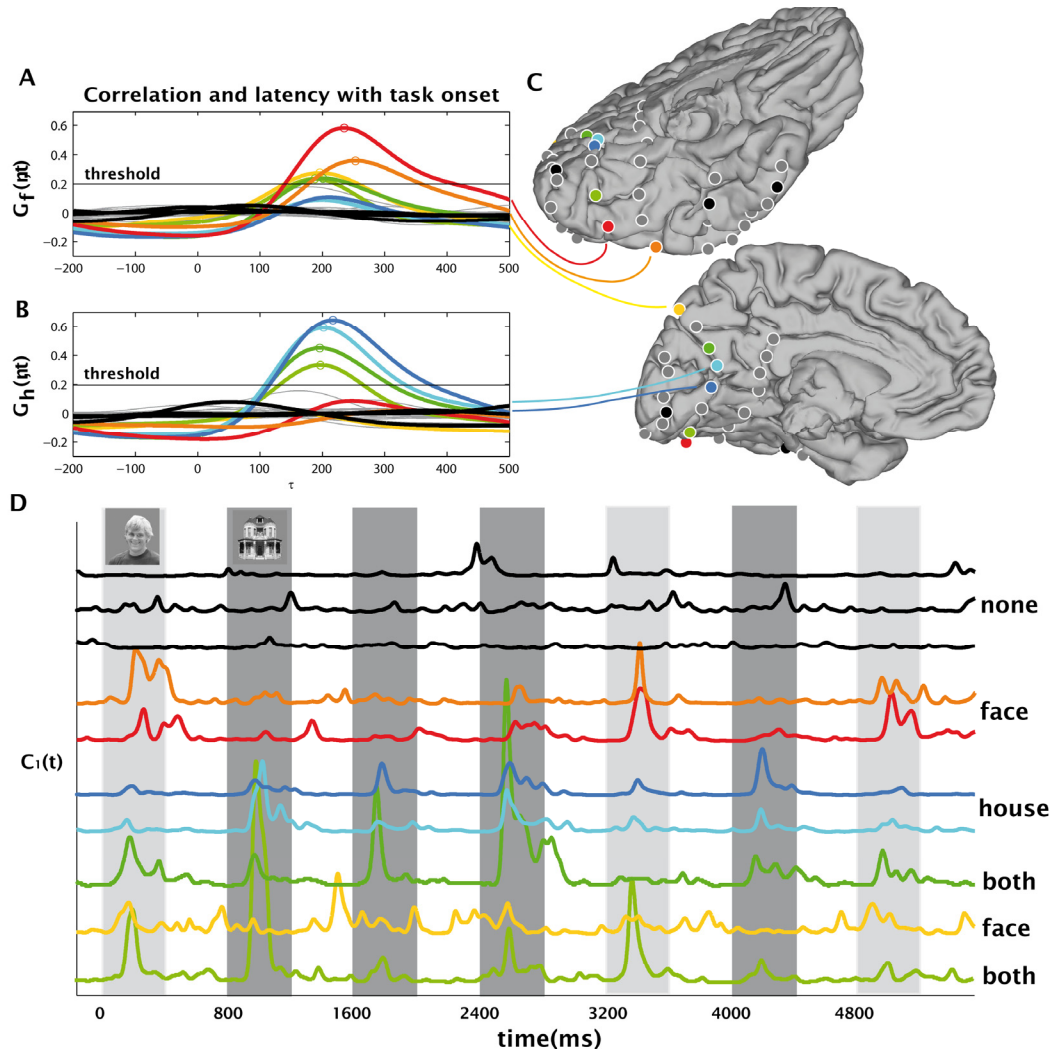


Fig. 2. Correlation ($G(n, \tau)$) as a function of latency from task onset (τ) for faces (A) and houses (B). When the maximum (circle at the peak) correlation of neuronal population activity ($C_1(t)$) with the stimulus function ($S_{\{f,h,b\}}(t)$) was above the threshold of 0.2, an electrode was said to be significant. Electrodes significant for faces are yellow-orange-red, for houses, blue and electrodes that are significant for both are green. Black and gray electrodes were not significant. The position of the electrodes on template brain (C). Neuronal population activity ($C_1(t)$) at a function of time (ms) during face (light gray) and house (dark gray) stimulus presentation. (D). Faces and houses were presented for 400 ms, with an inter-stimulus interval of 400 ms. Electrodes are colored as in A and B, colored electrodes are ordered by peak latency from stimulus onset with the shortest latency at the bottom. Non-responsive black electrodes could not sensibly be ordered.

frequencies reveal the robust common features during visual perception. The most significant of these eigenvectors (\vec{e}_1 , with largest eigenvalue λ_1 , and roughly equal contribution from all frequencies) has been linked to local population activity (mean firing rate of the underlying population [4]). Continuous time-frequency approximations of the PSD (dynamic spectra, $\tilde{P}(f, n, t)$) were calculated using a wavelet approach (complex Morlet filter, with 7 cycles, at each Hz). The projection into the 1st principal component can then be estimated at each point in time, after first normalizing the continuous spectra (first by dividing through by the mean spectrum, and then by taking the log). Then, the projection of the continuous spectra, $e_1(f) * \tilde{P}(f, t)$ is smoothed (with a Gaussian of SD 12ms), and re-exponentiated to obtain the continuous quantity, $C_1(t)$, associated with the time-varying neuronal population activity (Fig 1).

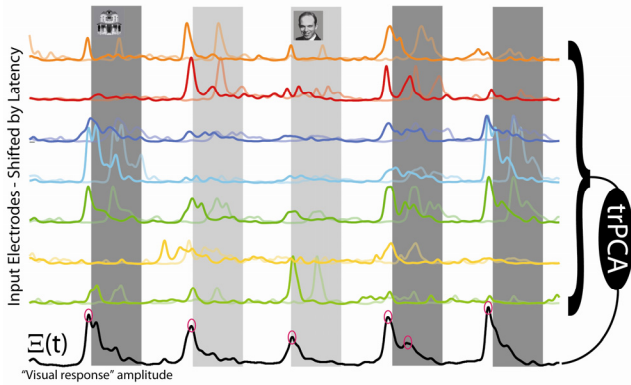


Fig.3. Neuronal population activity ($C_1(t)$ light colors) in visually relevant electrodes was shifted in time to align peak latencies ($C_1(t - \tau)$, dark colors). Electrode colors correspond to Fig 2c. A task related principal component analysis (trPCA) collapsed the set of $C_1(t - \tau)$ into a single visual response channel ($\Xi(t)$, black), whose peaks (pink) reflect estimated visual onset times, v'_k .

C. Relevant electrodes, and response latencies

The electrodes which covered important visual regions, along with their associated response latencies, were determined in the following fashion: “Stimulus functions” $\{S_f(t), S_h(t), S_b(t)\}$ were created by placing Gaussians of width $SD=60ms$ centered at the onset of {face, house or both face and house nonspecifically} stimuli. The correlation between each of these stimulus functions and the n^{th} electrode were calculated:

$$G_{\{f,h,b\}}(n, \tau) = \int dt \tilde{S}_{\{f,h,b\}}(t) \tilde{C}_1(n, t + \tau)$$

These are shown in Fig 2. Here $\tilde{X}(t) = \frac{X(t) - \bar{X}}{\sigma_X}$, where overline bar denotes mean, and σ denotes standard deviation. If $G_f(n, \tau)$ is below a set threshold (0.2 in this case) for all τ , then the neuronal population beneath electrode n does not respond to face stimuli. The value of τ for which $G_f(n, \tau)$ is maximized reveals the latency from stimulus presentation to cortical response in electrode n .

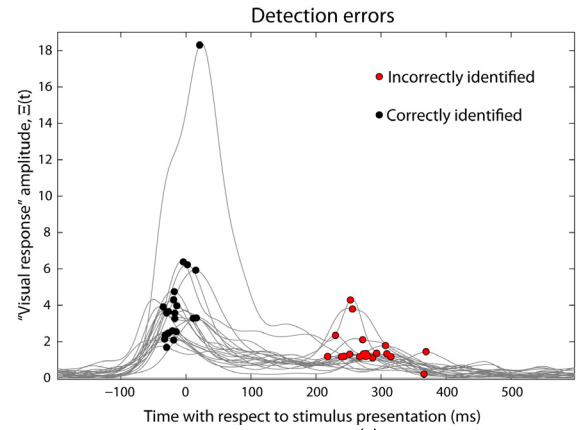


Fig.4. Visual response amplitude ($\Xi(t)$) in 21 of 22 incorrect detected events (red) These incorrect detected events reflect second responses, after correctly identified events (black).

D. Event detection and stimulus type classification

These methods were trained on 2/3 of the data, tested on the remaining 1/3, and subsequently cross-folded 3 times for validation.

1) *Event detection*: The visual response was collapsed by projecting latency shifted $C_1(t - \tau)$ from all visually relevant electrodes (colored electrodes in Fig 2c) into to a single “response channel” ($\Xi(t)$) for detection using task-related principal component analysis (Fig 3, [6]). The times of peak “visual response” were assumed to be most likely to correspond to stimulus onset times, because all electrodes had been shifted by their latency to peak prior to projection into $\Xi(t)$. Identification of these peaks in testing data were determined by smoothing $\Xi(t)$ (with an $SD=20ms$ Gaussian) and then calculating predicted visual event times v'_k for which $\Xi(v'_k)$ was a local maximum for 200ms in either direction and greater than 1 SD from baseline (baseline and SD from training period, see Fig 3). When correctly identified, the predicted times of visual stimulus onset v'_k could be compared with times of actual stimulus onset v_k .

2) *Event classification*: The “total response” to a stimulus at time v_k/v'_k was determined as the time integral from 100-600ms following the actual/predicted visual onset time (Fig 1, denoted R_k/R'_k). These R'_k were then classified as following face or house stimuli, using 3-fold cross-validation, based upon training on R_k (Linear Fisher Discriminant Analysis).

III. RESULTS

A. Correlations with stimuli are shown in table 1

B. Event detection

There were 300 visual stimuli presented. All of these were correctly identified, but 22 incorrect events were also detected. Further inspection revealed that 21 of these 22 were late, “second responses” (Fig 4) following a correct response prediction. The last error was truly incorrect, clearly uncorrelated with the onset of any stimulus.

C. Presentation time prediction

Of the 300 correctly identified events, the mean error (average absolute prediction time error) was 22.98 ± 23.23 ms (mean \pm SD) (faces = 21.71 ± 15.69 ms, houses = 24.25 ± 28.87 ms, Fig 5). The systematic error (average discrepancy keeping +/- sign of error) was 9.29 ms (faces = 11.65 ms, houses = -6.92 ms). The distribution of temporal errors for houses and faces were not statistically distinguishable from each other ($p = 0.19$, unpaired t-test).

D. Classification accuracy

The classification accuracy of correct stimulus-type (faces vs. houses) was 94.7% (faces: 96.0%; houses: 93%).

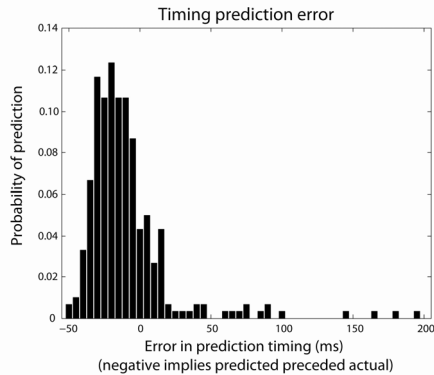


Fig.5. Timing prediction errors

IV. DISCUSSION

The correlation calculation illustrated in Fig 2 and table 1 is obviously less than perfect, but that is because arbitrary 60ms SD Gaussian kernels were used as model for neuronal population activity response functions. They were chosen *a priori* thus were clearly not physically motivated nor sophisticated, although they were sufficient to identify relevant electrodes – and a useful metric to determine correlation with the stimulus.

Although the experiment was periodic, and therefore accuracy in estimated stimulus presentation events and timing could potentially be explained trivially, the method was explicitly blind to any timing between events and all prior testing result. Furthermore, analysis of resulting event detection and timing reveals that errors in timing were not systematic with respect to prior stimulus accuracy or timing.

Except for one case, incorrectly identified events were due to a “late 2nd response”. The remaining one occurred at the end of an experimental block and may be behavioral (the subject glancing at an examiner, etc). There was a ~10ms systematic error in the estimated time of stimulus. This was not due to outliers (see Fig 5), and perhaps was due to fact that latencies were compared with symmetric (Gaussian) response function, but the responses were heavy-tailed to later times from peak. Therefore, latencies in electrode responses were overestimated (by net order ~10ms), and the subsequently estimated stimulus times were systematically ~10ms too early.

TABLE I
CORRELATION WITH “STIMULUS FUNCTIONS”








$C_1(t)$	Specificity	Peak correlation	Latency to peak (ms)
	Face	0.3598	252
	Face	0.5825	234
	House	0.6456	216
	House	0.5918	201
	Both	0.5218	197
	Face	0.2760	195
	Both	0.4298	192
$\Xi(t)$	Both	0.7250	0

Table of peak correlations, $G_{\{f,h,b\}}(n, \tau)$; “Specificity” denotes most appropriate stimulus {face, house, both face and house}; “Peak correlation denotes maximum value of G ”; “Latency to peak” denotes the value of τ at which G was maximum.

V. CONCLUSION

This study demonstrates conclusively that the timing and type of visually stimuli can be robustly predicted from the continuously measured electrocorticogram from the surface of visual regions in human cortex.

ACKNOWLEDGMENT

We would like to thank the patient, and the staff and physicians at Harborview hospital, in Seattle, WA.

REFERENCES

- [1] A. Ishai, L. G. Ungerleider, A. Martin, J. L. Schouten, and J. V. Haxby, "Distributed representation of objects in the human ventral visual pathway," *Proc Natl Acad Sci U S A*, vol. 96, pp. 9379-84, Aug 3 1999.
- [2] T. Allison, G. McCarthy, A. Nobre, A. Puce, and A. Belger, "Human extrastriate visual cortex and the perception of faces, words, numbers, and colors," *Cerebr Cortex*, vol. 4, pp. 544-54, Sep-Oct 1994.
- [3] K. J. Miller, S. Zanos, E. E. Fetz, M. den Nijs, and J. G. Ojemann, "Decoupling the Cortical Power Spectrum Reveals Real-Time Representation of Individual Finger Movements in Humans," *Journal of Neuroscience*, vol. 29, p. 3132, 2009.
- [4] K. J. Miller, L. B. Sorensen, J. G. Ojemann, and M. den Nijs, "ECoG observations of power-law scaling in the human cortex," *In Review (pre-print at http://arxiv.org/PS_cache/arxiv/pdf/0712/0712.0846v1.pdf)*, 2009.
- [5] K. J. Miller, S. Makeig, A. O. Hebb, R. P. N. Rao, M. denNijs, and J. G. Ojemann, "Cortical electrode localization from X-rays and simple mapping for electrocorticographic research: The “Location on Cortex”(LOC) package for MATLAB," *Journal of Neuroscience Methods*, vol. 162, pp. 303-308, 2007.
- [6] K. J. Miller, A. O. Hebb, J. G. Ojemann, R. P. N. Rao, and M. denNijs, "Task-Related Principal Component Analysis: Formalism and Illustration," *IEEE Eng Med Biol Soc.*, pp. 5469-5472, 2007.