and the model predictions for the work variable was significantly better than for position (paired T-test $p < 0.05$).

The data used in Fig. 1 came from data set 17. To allow comparison between the individual animals the data set/animal numbers were as follows: data sets 1–4, 5–7, 8–9, 10, 11–12, 13–17, 18–20 were from animals 1–7, respectively. In Fig. 2(b) we have plotted the force field variable as a binomial variable in time (black line) as well as the model prediction (gray line). Note that the force level could be predicted with 100% accuracy by using a simple threshold (dotted line).

IV. DISCUSSION

In this paper, we present the first demonstration that a BMI can successfully predict a range of movement related variables while a subject works against at least two dynamical constraints. These conditions were chosen to reflect those that one would experience in a real world environment. In order for a brain-controlled neuroprosthetic arm to function correctly, the user must be able to control the arm while transporting a wide variety of objects, not only including objects of differing mass, but a great range of natural and human-made objects with unpredictable dynamics. Our current demonstration that neuronal populations in the rat motor cortex can be used to predict two alternating force fields represents a first step towards designing BMIs that can handle a universal set of load dynamics. Thus, a BMI using the same population of cortical neurons could be used to predict not only the kinematics of a reaching movement, but also the forces associated with moving the manipulandum endpoint. In the data presented here, our prediction of work was significantly better than that of position, and thus using control variables that incorporate both position and force such as work may provide better and more natural BMI control.

It should be noted that it is possible to use positional data to control a robotic arm, but this may not be natural and could lead to difficulties. However, if the BMI user is actually controlling a variable such as work, they would have the ability to set the strength of their own movements rather than relying on preset values of a robotic position controller. In general, previous BMI demonstrations have acted as position controllers and have not included any force information. We propose that BMIs that can handle both position and force variables (plus the composite work variable) provide an approach to developing the next generation of natural and efficient robotic prostheses.

REFERENCES

[1] J. Ashe, "Force and the motor cortex," *Behav. Brain Res.*, vol. 87, pp. 255–269, 1997.
[2] A. Riehle and E. Vaadia, *Motor Cortex In Voluntary Movements*. Boca Raton, FL: CRC Press, 2005.
[3] D. M. Taylor, S. I. Tillery, and A. B. Schwartz, "Direct cortical control of 3D neuroprosthetic devices," *Science*, vol. 296, pp. 1829–1832, 2002.
[4] M. D. Serruya, N. G. Hatsopoulos, L. Paninski, M. R. Fellows, and J. P. Donoghue, "Instant neural control of a movement signal," *Nature*, vol. 416, pp. 141–142, 2002.
[5] J. K. Chapin, K. A. Moxon, R. S. Markowitz, and M. A. Nicolelis, "Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex," *Nat. Neurosci.*, vol. 2, pp. 664–670, 1999.
[6] J. M. Carmena, M. A. Lebedev, R. E. Crist, J. E. O'Doherty, D. M. Santucci, D. F. Dimitrov, P. G. Patil, C. S. Henriquez, and M. A. Nicolelis, "Learning to control a brain–machine interface for reaching and grasping by primates," *PLoS Biol.*, vol. 1, p. E42, 2003.
[7] F. Gandolfo, C. Li, B. J. Benda, C. P. Schioppa, and E. Bizzi, "Cortical correlates of learning in monkeys adapting to a new dynamical environment," *Proc. Nat. Acad. Sci. USA*, vol. 97, pp. 2259–2263, 2000.
[8] P. L. Gribble and S. H. Scott, "Overlap of internal models in motor cortex for mechanical loads during reaching," *Nature*, vol. 417, pp. 938–941, 2002.
[9] J. T. Francis and J. K. Chapin, "Force field apparatus for investigating movement control in small animals," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 963–965, Jun. 2004.

Cortically Coupled Computer Vision for Rapid Image Search

Adam D. Gerson, Lucas C. Parra, and Paul Sajda

Abstract—We describe a real-time electroencephalography (EEG)-based brain–computer interface system for triaging imagery presented using rapid serial visual presentation. A target image in a sequence of nontarget distractor images elicits in the EEG a stereotypical spatiotemporal response, which can be detected. A pattern classifier uses this response to reprioritize the image sequence, placing detected targets in the front of an image stack. We use single-trial analysis based on linear discrimination to recover spatial components that reflect differences in EEG activity evoked by target versus nontarget images. We find an optimal set of spatial weights for 59 EEG sensors within a sliding 50-ms time window. Using this simple classifier allows us to process EEG in real time. The detection accuracy across five subjects is on average 92%, i.e., in a sequence of 2500 images, resorting images based on detector output results in 92% of target images being moved from a random position in the sequence to one of the first 250 images (first 10% of the sequence). The approach leverages the highly robust and invariant object recognition capabilities of the human visual system, using single-trial EEG analysis to efficiently detect neural signatures correlated with the recognition event.

Index Terms—Brain–computer interface (BCI), cortically coupled computer vision, electroencephalography (EEG), image triage, rapid serial visual presentation (RSVP).

I. INTRODUCTION

The human visual system is exquisitely adept at parsing a scene and recognizing objects at a glance. Such an ability for rapid processing of visual information is even more impressive in light of the fact that neurons are relatively slow processing elements, compared to digital computers where individual transistors can switch 10^6 times faster than a neuron can spike. Rapid recognition is not only critical for our survival but is also necessary given the architectural constraints of our visual pathways.

Computer vision systems have proven to be less effective, in general, than human image analysts. An important problem, in disciplines ranging from radiology to satellite reconnaissance, is optimizing image throughput. Often there are a limited number of human analysts available to exploit the enormous amount of imagery being collected. Given that the "human visual processor" is the most general purpose and robust vision system ever "constructed," how can we optimally utilize the available human visual processors for searching through the immense amount of imagery?

A. D. Gerson and P. Sajda are with the Department of Biomedical Engineering, Columbia University, New York, NY 10027 USA (e-mail: adg71@columbia.edu; ps629@columbia.edu).
L. C. Parra is with the Department of Biomedical Engineering, City College of New York, New York, NY 10031 USA (e-mail: parra@ccny.cuny.edu).

Fixation (2 seconds)   **RSVP Sequence (100 ms / image)**   **Target**

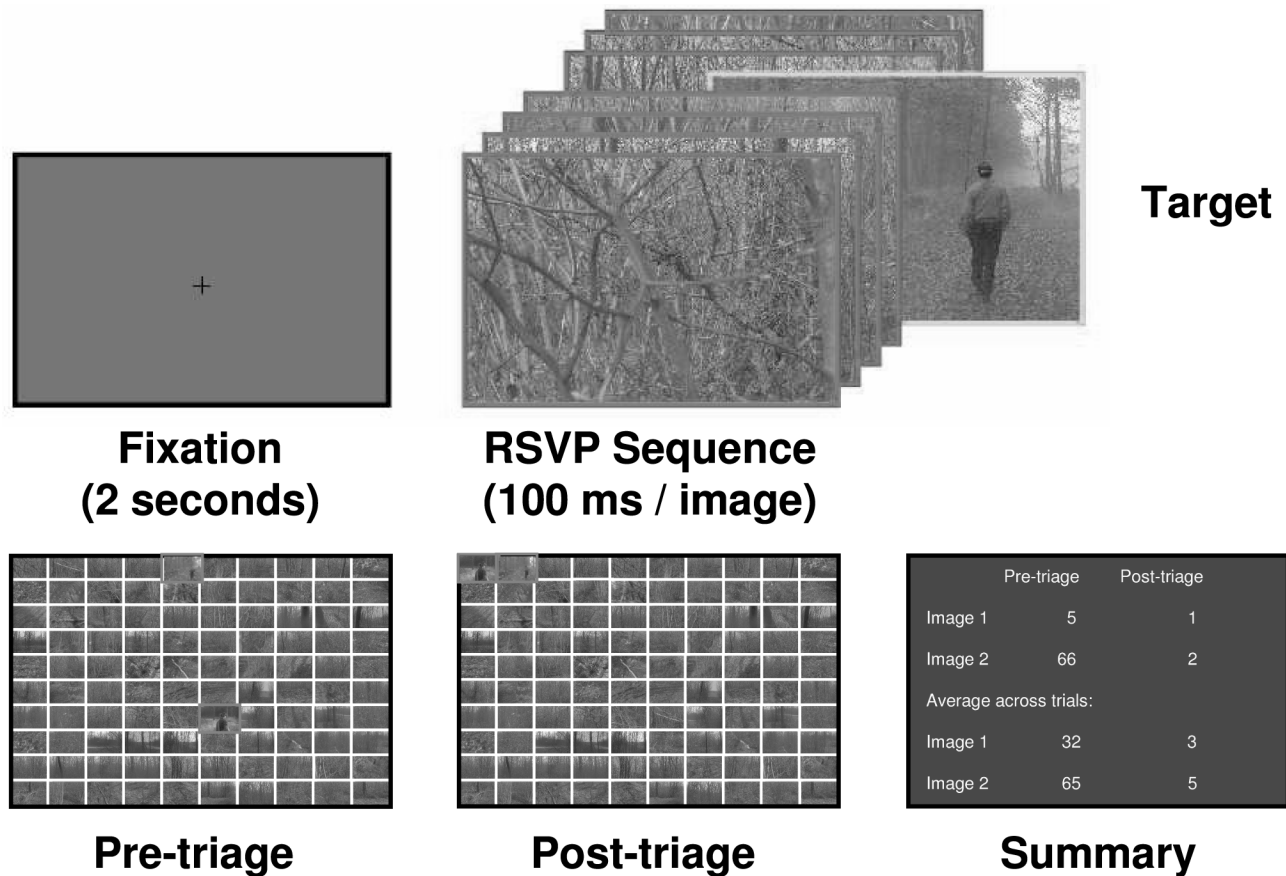**Pre-triage**   **Post-triage**   **Summary**

Fig. 1. Example RSVP triage trial. Fixation cross lasting 2 s is followed by sequence of 100 images. Each sequence contains two target images. Targets can appear at any position within sequence. Following image sequence, 10×10 matrix is displayed showing sequence of images. Target images are outlined. After pressing space-bar, this matrix is sorted according to EEG. Ideally, target images move to first two positions after resorting based on EEG. Participants press space-bar again to bring up summary slide showing position of target images before and after triage. Next trial begins after subject presses space-bar.

In this paper, we describe an electroencephalography (EEG) system capable of detecting neural signatures of visual recognition events evoked during rapid serial visual presentation (RSVP). The system is capable of using these signatures to triage sequences of images, reordering them so that target images are placed near the beginning of the sequence. We term our system "cortically coupled computer vision" since we leverage the robust recognition capabilities of the human visual system (e.g., invariance to pose, lighting, scale, etc.) and use a noninvasive cortical interface (e.g., EEG) to intercept signatures of recognition events; the visual processor performs perception and recognition and the EEG interface detects the result (e.g., decision) of that processing.

II. MATERIALS AND METHODS

A. Participants

Five participants (three females, age 21 years) from the Columbia University community participated. All had normal or corrected-to-normal vision and reported no history of neurological problems. Informed consent was obtained from all participants in accordance with the guidelines and approval of the Columbia University Institutional Review Board.

B. Data Acquisition

EEG data was acquired in an electrostatically shielded room (ETS-Lindgren, Glendale Heights, IL) using a Sensorium EPA-6 Electrophysiological Amplifier (Charlotte, VT) from 59 Ag/AgCl scalp electrodes mounted in a standard electrode cap (Electro-Cap, Eaton, OH) at locations based on the International 10–20 system and from three periocular electrodes placed below the left eye and at the left and right outer canthi. All channels were referenced to the left mastoid with input impedance < 15 kΩ and chin ground. Data was sampled at 1000 Hz with an analog pass band of 0.01–300 Hz using 12-dB/octave high pass and eighth-order Elliptic low-pass filters. Subsequently, a software-based 0.5-Hz high-pass filter was used to remove dc drifts and 60- and 120-Hz (harmonic) notch filters were applied to minimize line noise artifacts. These filters were designed to be linear phase to prevent delay distortions. Motor response and stimulus events recorded on separate channels were delayed to match latencies introduced by the digital filtering of the EEG.

C. Behavioral Paradigm

During the RSVP triage task, participants were presented with a continuous sequence of natural scenes. Participants completed two blocks of 50 sequences with a brief rest period lasting no more than five minutes between blocks. Each sequence consisted of 100 images and contained two target images with one or more people in a natural scene. These target images could appear at any position within each 100-image sequence. The remaining natural scenes without a person are referred to as distractor images. Each image was presented for 100 ms. As illustrated in Fig. 1, following the image sequence a series of self-paced feedback slides were presented indicating the position of target images within the sequence before and after EEG-based triage. A fixation cross was displayed for 2 s between trials.

During the second block, participants were instructed to quickly press the left button of a generic three-button mouse with their right index finger as soon as they recognized target images. They were

instructed to press the button twice, as quickly as possible, if one target image immediately followed the other. Participants did not respond with a button press during the first block.

D. Stimuli

Images were selected randomly with replacement from a queue of 251 nontarget and 33 target gray scale images. Images were obtained with a Kodak DCS420 digital camera with a 28-mm camera lens (Rochester, NY) [1]. Images were deblurred and the size of the images was reduced from 1536×1024 pixels to 640×426 pixels to decrease processing requirements of the stimulus computer and ensure precise timing of stimulus presentation. The images were nonlinearly transformed via gamma correction to match their mean luminance. The fixation cross display had the same mean luminance as the images. A Dell Precision 530 Workstation (Round Rock, TX) with nVidia Quadro4 900XGL graphics card (Santa Clara, CA) and E-Prime software (Psychological Software Tools, Pittsburgh, PA) controlled stimulus display. An LCD projector (InFocus LP130, Wilsonville, OR) projected stimuli through an RF shielded window onto a front projection screen. Stimuli subtended $33° \pm 3° \times 25° \pm 3°$ of visual field. Target images were visually inspected to ensure that target objects did not comprise more than 25% of the area (fraction of pixels) in the scene.

E. Online Artifact Reduction

Immediately prior to the RSVP task, participants completed an eye motion calibration experiment during which they were instructed to blink repeatedly upon the appearance of a white on black fixation cross and then make several horizontal and vertical eye movements according to the position of a fixation cross subtending $1° \times 1°$ of the visual field. Horizontal eye movements subtended $33° \pm 3°$ and vertical eye movements subtended $25° \pm 3°$. The timing of these visual cues was recorded simultaneously with EEG. This enabled determination of linear components associated with eye blinks and eye movements that were subsequently projected out of EEG recorded during the RSVP triage task. This procedure is described in detail in [2].

F. Real-Time Spatial Integration for Image Classification

The RSVP task is effectively an oddball task, eliciting a P300 response which has been used extensively to develop BCIs for communication [3]. In order to classify EEG online, we use a Fisher linear discriminator [4], [5] to estimate a spatial weighting vector $\mathbf{w}_{\tau,\delta}$ which maximally discriminates between sensor array signals for two conditions, $c \in (0, 1)$, where $c = 1$ represents a target trial and $c = 0$ a nontarget trial. This weighting vector is specific to a training window starting at a relative onset time τ, with a duration of δ. The spatial weighting vector generates a component[1]

$$y_c(t) = \mathbf{w}_{\tau,\delta}^T \mathbf{x}_c(t) + b_{\tau,\delta} \tag{1}$$

that best separates the EEG signals by making $y_0(t) < y_1(t)$ for as many samples in the training window as possible. The result is a discriminating component specific to target recognition activity while minimizing activity correlated with both task conditions such as early visual processing.

Following each trial, the mean $\boldsymbol{\mu}_c$ and covariance matrix $\boldsymbol{\Sigma}_c$ are updated for the condition associated with the trial. The spatial weighting vector $\mathbf{w}_{\tau,\delta}$ and bias $b_{\tau,\delta}$ are updated as

$$\mathbf{w}_{\tau,\delta} = \boldsymbol{\Sigma}_{\text{pool}}^{\#} \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{\text{pool}}^{\#} \boldsymbol{\mu}_0 \tag{2}$$

$$b_{\tau,\delta} = \frac{1}{2} \left[\boldsymbol{\mu}_0 \boldsymbol{\Sigma}_{\text{pool}}^{\#} \boldsymbol{\mu}_0 - \boldsymbol{\mu}_1 \boldsymbol{\Sigma}_{\text{pool}}^{\#} \boldsymbol{\mu}_1 \right] \tag{3}$$

[1]We use the term "component" instead of "source" to make it clear that this is a projection of all the activity correlated with the underlying source.

where the pooled covariance $\boldsymbol{\Sigma}_{\text{pool}} = ((N_0 \boldsymbol{\Sigma}_0 + N_1 \boldsymbol{\Sigma}_1)/(N_0 + N_1))$, and N_0 and N_1 are the number of samples acquired for nontarget and target conditions. We use the pseudoinverse of $\boldsymbol{\Sigma}_{\text{pool}}$, $\boldsymbol{\Sigma}_{\text{pool}}^{\#}$ to ensure the stability of our online algorithm, particularly at the beginning of training when only a few samples have been observed and $\boldsymbol{\Sigma}$ is rank deficient.

Given our linear model, determination of sensor projections from the discriminating activity is straightforward, namely

$$\mathbf{a} = \frac{\langle \mathbf{x}(t), y(t) \rangle}{\langle y(t), y(t) \rangle} \tag{4}$$

where $\langle \cdot, \cdot \rangle$ denotes an inner product. Equation (4) describes the projection $\mathbf{a}$ of the discriminating component $y(t)$ that explains most of the activity $\mathbf{x}(t)$. A strong projection indicates low attenuation. Therefore, the intensity of sensor projections $\mathbf{a}$ indicates proximity of the component to the sensors.

During each experimental condition (with and without motor response), 5000 images were presented to the subject in sequences of 100 images. EEG evoked by the first 2500 images (50 targets, 2450 nontargets) was used to train the classifier. During the experimental sessions, a training window 400–500 ms following stimulus onset was used to extract training data. This training window was selected based on the timing of the P300 for this task. Weights were updated adaptively with each trial during the training period. These weights were frozen at the end of the training period and blindly applied to the subsequent testing dataset (images 2501–5000).

1) Multiple Classifiers: After the experiment, multiple classifiers with different training window onsets τ were used to boost triage performance. The training window onsets ranged from 0 to 900 ms in steps of 50 ms. The duration of the training windows δ was 50 ms. Once these classifiers were trained, the optimal weighting of these classifier outputs was found using logistic regression to discriminate between target and nontarget images

$$z_c = \boldsymbol{\omega}^T \boldsymbol{\psi}_c + \beta \tag{5}$$

making $z_0 < z_1$ for as many trials as possible. $\boldsymbol{\psi} = [\langle y(t)_{\tau_1} \rangle \langle y(t)_{\tau_2} \rangle \cdots \langle y(t)_{\tau_J} \rangle]^T$ where the sample average

$$\langle y(t)_{\tau_j} \rangle = \frac{1}{\delta} \sum_{t=\tau_j}^{\tau_j+\delta-1} y(t) \tag{6}$$

is taken over corresponding training windows with onsets $\tau_1 = 0$ ms, $\tau_2 = 50$ ms, $\cdots$, $\tau_J = 900$ ms. Note that while the structure of our classifier resembles a multilayer perceptron, rather than adapting the coefficients for all classifiers simultaneously, classifiers were trained separately to avoid overtraining and improve generalization. Again, only EEG data evoked by the first 2500 images was used to train the classifiers and then find the interclassifier weights. These weights were then applied to the testing data set evoked by the second set of 2500 images (images 2501–5000).

Similarly as with the sensor projection, $\mathbf{a}$, we can find projections of the discriminating activity

$$\boldsymbol{\alpha} = \frac{\langle \boldsymbol{\psi}, z \rangle}{\langle z, z \rangle}. \tag{7}$$

The projection $\boldsymbol{\alpha}$ describes the relative strength of the temporal discriminating component over time.

TABLE I
TRIAGE PERFORMANCE AND BEHAVIORAL RESULTS

Subject	EEG (no motor)	EEG (motor)	Button	EEG (motor) and Button	RT (training) (ms)	RT (testing) (ms)	% Correct (training)	% Correct (testing)	Response bias c, (testing)
1	0.92	0.91	0.87	0.94	418 ± 133	413 ± 101	88	86	-0.8
2	0.94	0.96	0.86	0.97	412 ± 64	450 ± 64	94	74	-1.2
3	0.90	0.87	0.96	0.96	445 ± 79	423 ± 59	86	94	-0.7
4	0.91	0.92	0.98	0.98	433 ± 74	445 ± 59	98	98	-0.5
5	0.91	0.93	0.98	0.98	398 ± 86	402 ± 58	96	96	-0.9
Group	0.91 ± 0.02	0.92 ± 0.03	0.93 ± 0.06	0.97 ± 0.02	421 ± 91	426 ± 71	92 ± 5	90 ± 10	-0.8 ± 0.3

G. Image Triage

After each image sequence (two targets and 98 nontargets), image matrices (10×10) were presented to the subject to show the position of each image within the sequence (Fig. 1). Target images were outlined in red. The image sequence was then resorted based on the output of a single classifier with one training window, $\langle y(t) \rangle$ ($\tau = 400$ ms, $\delta = 100$ ms). Ideally, the two target images are triaged to the first two positions in the resorted image sequence.

Following the experiment, all image sequences were concatenated to create training and testing sequences that each contain 2500 images (50 targets and 2450 nontargets). These image sequences are resorted according to the output of our classifier with multiple training windows z, for EEG evoked by every image.

1) Button-Based Image Triage: For comparison, sequences were triaged based on the button response. Images were resorted according to

$$p(\mathrm{target}|RT)$$
$$= \frac{p(RT|\mathrm{target})p(\mathrm{target})}{p(RT|\mathrm{target})p(\mathrm{target})+p(RT|\mathrm{nontarget})p(\mathrm{nontarget})} \quad (8)$$

where RT is the onset of a button response that occurs within one second of image onset. $p(\mathrm{target}|RT) = 0$ when no response occurred within 1 s of image onset. The priors $p(\mathrm{target}) = 0.02$ and $p(\mathrm{nontarget}) = 0.98$. $p(RT|\mathrm{target})$ is a Gaussian distribution with a mean and variance determined from the response times from the training sequences. Since more than one response is likely to follow a target image if the two target images are presented within 1 s of each other, for training sequences response times were assigned to target images based on the position of the target image within the sequence. In other words, if the target appeared first in the sequence and two button responses occurred within 1 s of this target's onset, the first response was assigned to that target image and the second response was assigned to the second target image. For testing sequences, if two or more responses occur within 1 s of the onset of any image, the response with the greatest $p(\mathrm{target}|RT)$ is assigned to the image. $p(RT|\mathrm{nontarget})$ is a mixture of 13 Gaussians, each with the same variance as that used for $p(RT|\mathrm{target})$ and with means assigned by shifting the mean from $p(RT|\mathrm{target})$ 600 ms in the past to 700 ms in the future in increments of 100 ms, excluding the actual mean of $p(RT|\mathrm{target})$. This mixture model contains a sufficient number of Gaussians so that the mixture is consistent within the 1-s interval following image onset. $p(RT|\mathrm{nontarget})$ was designed to model responses occurring within 1 s of the onset of a nontarget image that is presented within 1 s prior to or following a target image.

III. RESULTS

A. Task Performance

In the RSVP task, five participants correctly responded to $92 \pm 5\%$ and $90 \pm 10\%$ of targets during training and testing sequences, respectively. Response times for correctly identified target across five subjects

had a mean of 421 ± 91 ms and 426 ± 71 ms for training and testing sequences, respectively. Behavioral statistics for five subjects are shown in Table I. Mean and standard deviation of response times for training and testing sequences are listed with the corresponding percent of correctly identified targets. The response bias [6], [7], c is also listed for testing sequences

$$c = \frac{\Phi^{-1}(H) + \Phi^{-1}(F)}{2} \quad (9)$$

where H and F are the hit rate and the false alarm rate, respectively, and Φ^{-1} is the inverse of the normal cumulative distribution function used to convert probabilities into z scores. Negative values of c indicate a bias toward not responding to images.

B. Discriminating Components

Following the experiment, 19 classifiers were trained using training windows with onsets ranging from 0 to 900 ms in increments of 50 ms. The duration of each training window was 50 ms. Fig. 2 shows the mean scalp projections for discriminating components from these classifiers across five participants. Scalp projections **a** of discriminating components were normalized prior to averaging across subjects. Components are shown for both the motor and nonmotor tasks.

Bilateral occipital activity seen about 150 ms following stimulus onset is followed by strong negatively correlated activity over frontal electrodes. This frontal activity occurs slightly later for the motor condition (250 versus 200 ms) and is followed by positively correlated activity over parietal electrodes about 350 ms after stimulus onset. This parietal activity spreads over central electrodes until about 500 ms for both motor and nonmotor tasks. This frontal-parietal progression of activity is consistent with models of the P3 complex indicating that anterior cingulate, inferior-temporal lobe, and hippocampus are primarily responsible for generation of observed potentials [2], [8]–[11]. The scalp maps diverge for motor and nonmotor tasks from 550 through 700 ms. The motor condition shows sustained positively correlated activity over frontal and central electrodes during this period, while the nonmotor condition elicits only negatively correlated activity bilaterally over occipital electrodes. Subsequent components show positively correlated activity over frontal electrodes for both motor and nonmotor conditions. Curiously, activity specific to the motor conditions occurs after the response. While it is possible that this activity may be associated with a somatosensory response, given the symmetry of this bilateral response this seems unlikely. Fig. 2 also shows the time course of the multiple classifier system $\boldsymbol{\alpha}$. A strong temporal projection indicates there is strong correlation in time. This time course is stronger for the earlier components of the nonmotor condition versus motor condition.

C. Triage Performance

Triage results for one subject (subject 2) are shown in Fig. 3. Fig. 3(a) shows number of targets as a function of the number of distractor images both before and after triage based on button press and EEG. The
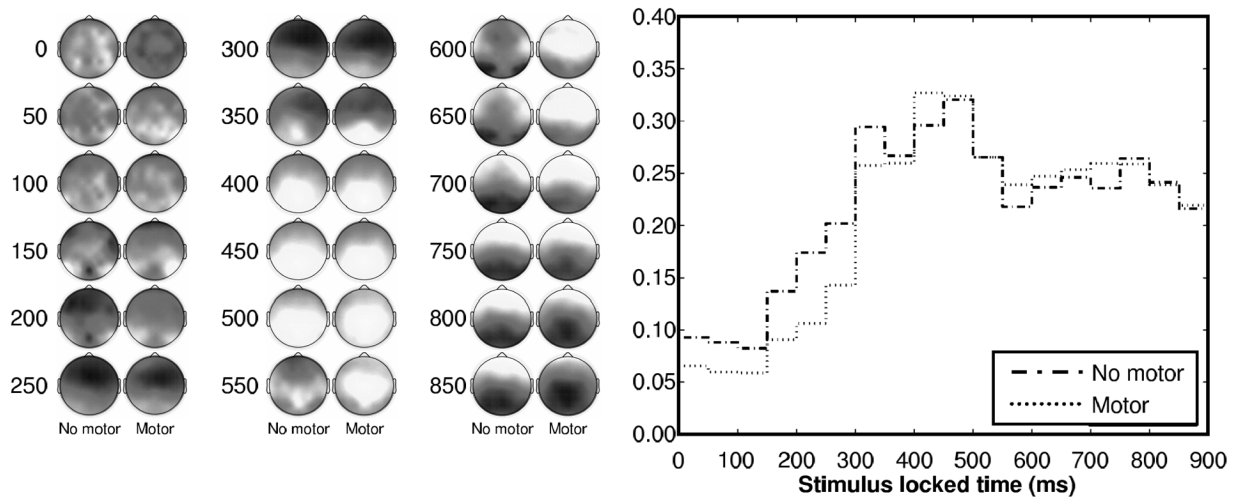
Fig. 2. Group results over five subjects. Scalp maps on left show sensor projections **a** of discriminating activity $y(t)$ averaged over all subjects for nonmotor (left columns) and motor (right columns) experimental tasks. Numbers to left of each scalp map pair correspond to onset, relative to stimulus presentation, of temporal window selected to train each classifier, in milliseconds. Duration of each classifier is 50 ms. Note that scalp maps for motor and nonmotor conditions begin to diverge about 550 ms following stimulus onset. Plot on right shows projected time course **α** of discriminating activity for nonmotor (dashed–dotted curve) and motor (dotted curve) tasks.
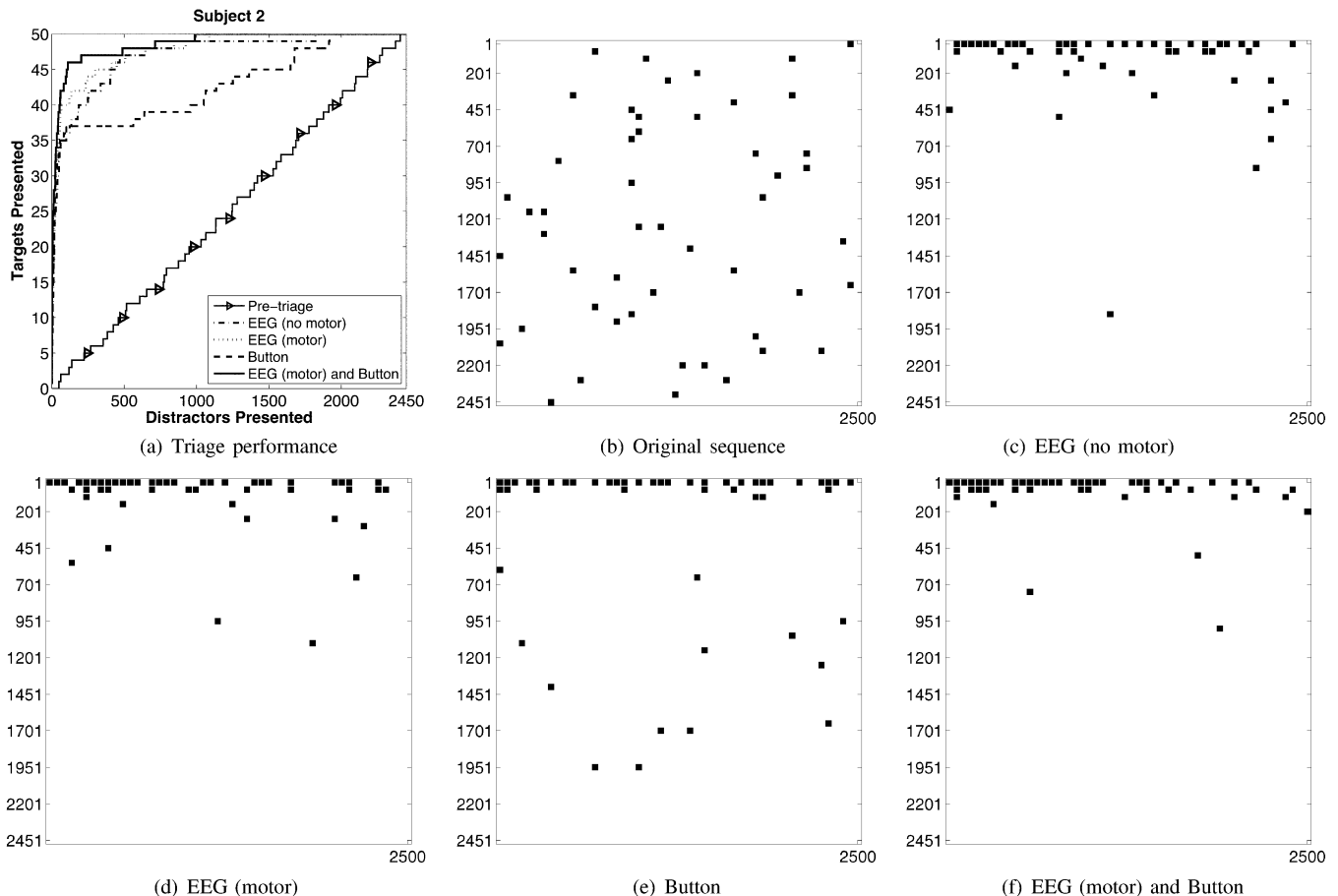

Fig. 3. Triage performance for subject 2 using multiple classifiers. (a) Number of target images presented as function of number of distractor images presented. Ideal triage system will place 50 (100%) of target images before all 2450 distractor images. Curve marked by triangles shows original sequence. Button-based triage is shown by dashed curve. Dashed–dotted curve shows EEG-based triage during experiment without motor response. Dotted curve shows EEG-based triage during experiment with motor response and thick black curve shows triage based on EEG (motor) and button response. (b)–(f) Rasters showing position of nontarget (white areas) and target (black squares) within (b) original image sequence, (c) EEG (no motor)-based triage sequence, (d) EEG (motor)-based triage sequence, (e) button-based triage sequence, and (f) combined EEG (motor) and button-based triage sequence. First and last images in each sequence are shown by squares in upper left and lower right of each raster, respectively.

area under the curve generated by plotting fraction of targets as a function of the fraction of distractor images presented is used to quantify triage performance. Triage performance of the multiple classifier system for five subjects is listed in Table I. This area is 0.50 for all unsorted image sequences since target images are randomly distributed throughout the sequences. Ideal triage performance results in an area of

1.00. There is no significant difference in performance between button-based and EEG-based triage $(0.93 \pm 0.06, 0.92 \pm 0.03, p = 0.69, N = 5)$. Interestingly, there is no significant difference in performance between EEG-based triage for the motor and no motor response conditions $(0.92 \pm 0.03, 0.91 \pm 0.02, p = 0.81, N = 5)$.

Figs. 3(b)–3(f) are rasters showing the position of the target images (black squares) and nontarget images (white areas) in the concatenated image sequence. Based on these rasters and the EEG and button-based triage performance for five subjects listed in Table I, it is clear that both EEG and button-based triage systems are capable of a high level of performance. The button-based triage performance begins to fail, however, when subjects do not consistently respond to target stimuli and response times exceed 1 s. Subject 2, for instance, correctly responded to only 74% of targets during the testing session and had a stronger bias than other subjects not to respond to images as reflected by the response bias measure c listed in Table I. In fact, this subject did not respond to 12 of 50 target images and the response time for one target image exceeded 1 s. Excessively late responses cannot effectively be classified using our Bayesian methods since it is not clear whether these button presses were in response to the target image or a subsequent nontarget image. The EEG response evoked by images with either no response or a late response is, however, still consistent with EEG evoked by the target images with predictable response times. The EEG-based triage system is therefore capable of detecting the recognition of these target images and subsequently resorting these target images appropriately. For this reason, we exploit the information provided by both EEG and button press using another perceptron to boost triage performance. This approach is effective for increasing triage performance for subjects that either did not respond or had a delayed motor response to a significant number of target images (e.g., subjects 1 and 2).

IV. CONCLUSION

We have demonstrated a real-time system for triaging sequences of images based on EEG signatures extracted using linear classifiers. While there were no significant differences in triage performance based on EEG or button press, button-based triage performance was superior for subjects that correctly responded to a high percentage of target images. However, for two subjects that responded to fewer images correctly (subjects 1 and 2), the EEG-based triage system offered better performance. We find that integrating information from EEG and behavioral response offers the best strategy. Using this method, we can capture images that lack an overt behavioral response and would have otherwise been missed by a button-based triage system.

Our on-line image triage system was initially designed to classify EEG based on a single training window $(\tau = 400 \text{ ms}, \delta = 100 \text{ ms})$. We realized once the experiments were completed that the performance of our EEG-based triage system is vastly improved using the multiple classifier system, and we plan to use these methods for future BCI experiments. We do not present results from the single classifier but rather note the improvement in performance using the multiple classifier system. It is clear from the performance improvement that the evoked response contains information that supplements the P3 for classification. Data was presented sequentially to the multiple classifier system in order to simulate online training conditions. This classifier was implemented for use online.

Following each image sequence, feedback was presented showing the position of target images before and after triage. The influence of feedback on behavioral performance and evoked EEG is unknown since we did not conduct a control experiment without feedback. This certainly warrants further study. It would be of interest, for instance, to determine whether feedback evokes activity in reward centers such as the anterior cingulate.

It is not clear why some subjects consistently respond to more target images correctly. With this in mind, we are investigating the impact of target salience within the scene (e.g. clutter, camouflage, and noise), image presentation rate, and user expertise (e.g., novice versus expert) on both behavioral and neural responses. We expect that the density of target images within each sequence will have a significant impact on evoked EEG activity due to attentional blink effects [12], [13] and the strong dependence of the P300 response on target probability and target-to-target interval (TTI) [14]. Since the strength of this response is also modulated by the level of surprise, we also expect contextual cues to affect EEG evoked by target images. For this reason, we are also exploring the effects of spatio-temporal correlation between consecutive video frames.

REFERENCES

[1] J. van Hateren and A. van der Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex," *Proc. Roy. Soc. London B*, vol. 265, pp. 359–366, 1998.

[2] A. Gerson, L. Parra, and P. Sajda, "Cortical origins of response time variability during rapid discrimination of visual objects," *NeuroImage*, vol. 28, no. 2, pp. 326–341.

[3] E. Donchin, K. M. Spencer, and R. Wijesinghe, "The mental prosthesis: Assessing the speed of a p300-based brain-computer interface," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 174–179, Jun. 2000.

[4] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. New York: Wiley, 2001.

[5] L. C. Parra, C. D. Spence, A. Gerson, and P. Sajda, "Recipes for the linear analysis of EEG," *NeuroImage*, vol. 28, no. 2, pp. 342–353.

[6] J. G. Snodgrass and J. Corwin, "Pragmatics of measuring recognition memory: Applications to dementia and amnesia," *J. Expe. Psychol.: General*, vol. 117, p. 34050, 1988.

[7] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*. New York: Wiley, 1966.

[8] D. Friedman, "Cognition and aging: A highly selective overview of event-related potential (ERP) data," *J. Clinical Exper. Neuropsychol.*, vol. 25, no. 5, pp. 702–720, 2003.

[9] J. Polich, Ed., *Detection of Change: Event-Related Potential and fMRI Findings*. Boston, MA: Kluwer, 2003.

[10] S. Makeig, A. Delorme, M. Westerfield, J. Townsend, E. Courchense, and T. Sejnowski, "Electroencephalographic brain dynamics following visual targets requiring manual responses," *PLoS Biol.*, vol. 2, no. 6, pp. 747–762, Jun. 2004.

[11] S. Makeig, M. Westerfield, T.-P. Jung, J. Covington, J. Townsend, T. Sejnowski, and E. Courchesne, "Independent components of the late positive response complex in a visual spatial attention task," *J. Neurosci.*, vol. 19, pp. 2665–2680, 1999.

[12] K. K. Evans and A. Treisman, "Perception of objects in natural scenes: is it really attention free?," *J. Exper. Psychol.: Human Perception Performance*, vol. 31, no. 6, pp. 1476–1492, 2005.

[13] C. Kranczioch, S. Debener, and A. K. Engel, "Event-related potential correlates of the attentional blink phenomenon," *Cognitive Brain Res.*, vol. 17, pp. 177–187, 2003.

[14] C. J. Gonsalvez and J. Polich, "P300 amplitude is determined by target-to-target interval," *Psychophysiol.*, vol. 39, pp. 388–396, 2002.