

# Collaborative Brain-Computer Interfaces for Target Detection and Localisation in Rapid Serial Visual Presentation

A. Matran-Fernandez and R. Poli  
Brain-Computer Interfaces Laboratory  
School of Computer Science and Electronic Engineering  
University of Essex  
{amatra,rpoli}@essex.ac.uk

*School of Computer Science and Electronic Engineering  
University of Essex, UK  
Technical Report CES-531  
ISSN: 1744-8050  
July 2014*

## Abstract

The rapid serial visual presentation protocol can be used to show images sequentially on the same spatial location at high presentation rates. We used this technique to present aerial images to participants looking for predefined targets (airplanes) at rates ranging from 5 to 12 Hz. We used linear support vector machines for the single-trial classification of event-related potentials from both individual users and pairs of users (in which case we averaged either their individual classifiers' analogue outputs before thresholding or their electroencephalographic signals associated to the same stimuli) with and without the selection of compatible pairs.

We considered two tasks — the detection of targets and the identification of the visual hemifield in which targets appeared. While single users did well in both tasks, we found that pairs of participants with similar individual performance provided significant improvements. In particular, in the target-detection task we obtained median improvements in the area under the receiver operating characteristic curve (AUC) of up to 8.3% w.r.t. single-user BCIs, while in the hemifield classification task we obtained AUCs up to 7.7% higher than for single users. Furthermore, we found that this second system allows not just to say if a target is in on the left or the right of an image, but to also recover the target's approximate horizontal position.

# 1 Introduction

Brain-Computer Interfaces (BCIs) convert electroencephalographic (EEG) signals from the brain into commands that allow users to control devices without the help of the usual peripheral pathways. Traditionally, BCIs have been developed with the aim of helping people with limitations in their motor control or their ability to communicate [1, 2, 3]. However, some forms of BCIs have recently started focusing on the augmentation of human abilities (e.g., speed) of able-bodied users, both individually and in groups by means of collaborative or cooperative BCIs (cBCIs) [4, 5, 6, 7, 8]. The latter work by merging EEG signals (or the corresponding control commands) from multiple users with the aim of controlling a single device.

Some of these forms of BCI focus on augmenting visual perception capabilities to speed up the process of finding pictures of interest in large collections of images [9, 10, 6, 11]. These systems would find applications, for instance, in counter intelligence and policing, where large amounts of images need to be viewed and classified daily by analysts looking for possible threats or, more generally, targets [10]. Apart from detecting such targets accurately and at high speeds, it stands to reason that current triage systems would benefit from techniques, such as the one we will present in this paper, that could establish the position of targets within the images.

It has been shown that the combination of the Rapid Serial Visual Presentation (RSVP) protocol (which sequentially displays images in the same spatial location at high presentation rates, e.g., 10 Hz [12]) with BCIs can effectively reduce triage time without a detriment in target detection accuracy [7, 13, 11]. Even at such a rapid pace of presentation, observers can detect target configurations and these elicit Event-Related Potentials (ERPs) in the brain. In particular, if targets are reasonably rare, a P300 ERP (a large positive wave typically peaking 300–600 ms after stimulus onset) is likely to be produced in response to them as conditions are effectively those of the “oddball” paradigm (i.e., in a situation where a low proportion of stimuli of interest are placed within a sequence of uninteresting or distractor stimuli) [14, 15]. Together, the selective differences in the potentials produced by target and non-target stimuli are sufficient to build BCIs that are capable of identifying images containing targets [9] with a reasonable accuracy — particularly when multiple observers are pooled together with *collaborative* forms of BCI [6, 5].

The P300 is one of the most widely used ERPs for controlling BCIs (both in traditional and the newer paradigms mentioned above), together with event-related synchronization and desynchronization [16] and Steady-State Visual Evoked Potentials (SSVEP) [2, 17], but it is just one the many components that have been identified in EEG signals. Another ERP that can be exploited in BCIs [18] and of particular interest for this work is the N2pc (a small negative asymmetric component preceding the P300) which, in the literature, has predom-

inantly been related to processes associated with selective attention [19, 20, 21]. The N2pc ERP is elicited when participants are given a search template or object to look for and the search display shows at least one distractor (i.e., non-target) item apart from the target.

The usual approach to increase the signal-to-noise ratio in BCIs, which are highly contaminated by noise and artifacts, is to average signals from different trials to isolate the ERP of interest [22]. For example, in their N2pc-driven BCI, Awni *et al* performed averages across 3 repetitions of the stimuli (trials). They reported large variations in classification accuracy across participants when discriminating between left and right targets (different-colored numbers in a circle) [18].

However, it is not always possible to average across multiple trials (e.g., a person cannot make the same decision several times), or it might not be practical (e.g., when designing BCIs for healthy users, where speed is a key factor). In this type of situations, aggregating signals from a number of users has proven to be useful, thus creating a “multi-brain” or cBCI (e.g., [4, 23]).

The field of collaborative BCIs is relatively new, and the issue of what is the best way to form groups has not been considered yet. The general opinion based on studies about group decision making is that bigger groups lead to better or more accurate decisions [24]. However, Kao and Couzin [25] showed that in many contexts where this “crowd wisdom” effect is not present, small groups can maximize decision accuracy, depending on correlations between the behavior of the members. In visual perception experiments, Bahrami *et al* found that observers performed better in pairs, provided that they had similar visual sensitivities and were able to communicate freely [26].

With respect to collaborative BCIs, it has been shown that groups are able to accelerate responses w.r.t. non-BCI decisions (e.g., key presses), and bigger groups lead to higher accuracies [4, 23, 27, 8]. However, when compared to non-BCI users, non-BCI decisions might prove to be more accurate than those reached by means of cBCIs [23, 28].

As we mentioned above, BCIs have been used for the automatic detection of targets in images by means of the EEG with reasonably good results [10, 29]. In [5], we reported on preliminary work on the collaborative classification of aerial images by means of the RSVP paradigm at different presentation rates (5–15 Hz) and varying the number of targets that participants were asked to look for. By pairing observers (in all possible ways from a pool of five), we were able to speed up the process of revising the images and obtained noteworthy higher accuracies than with single observers. This work was extended in [6], where we tested 10 participants and used them to form groups of 2 and 3 observers. We found statistically significant differences between groups and single-user BCIs. However, no form of selection was applied when forming groups.

Several forms of combining evidence from multiple individuals have been considered [30, 31]. Whether one form or another performs best may depend on

the field of application. In [5, 6], we found that the best way of integrating information from multiple participants for our P300-based cBCI is to average the outputs of individual support-vector machines (SVMs), each specialised to classify the data of one participant. Tests with directly averaging the ERPs from each participant suggested that this is a suboptimal strategy. However, we did not test this strategy in a left vs right classification (the N2pc has a relatively low latency jitter, so averaging this ERP across participants might work better than for the P300).

The work presented in this report uses the stimulation protocol proposed in [5, 6]. Also, a subset of the participants used for this study were originally tested in such prior work. However, as we indicated above, in this paper we have applied the concept of collaborative BCIs (in its two possible embodiments) to the localisation of targets within images. In addition, we will explore the effects of selecting the participants which form the groups in collaborative BCIs, both for left vs right classification and target vs non-target classification. Selection will be done on the basis of performance similarity.

## 2 Methods

### 2.1 Participants and setup

Due to the nature of RSVP, participants were screened for any personal or family history of epilepsy. We gathered data from 9 volunteers with normal or corrected-to-normal vision (age  $24.7 \pm 3.9$ , three females). They all read, understood and signed an informed consent form approved by the Ethics Committee of the University of Essex.

Participants were comfortably seated at approximately 80 cm from an LCD screen where the stimuli were presented. EEG data were acquired with a BioSemi ActiveTwo system with 64 electrodes mounted in a standard electrode cap following the international 10-20 system plus one electrode on each earlobe (all impedances  $< 20 \text{ k}\Omega$ ). The EEG was referenced to the mean of the electrodes placed on the earlobes. The initial sampling rate was 2048 Hz. Signals were band-pass filtered with cutoff frequencies of 0.15 and 25 Hz before downsampling to 64 Hz. A form of correction for eye-blinks and other ocular movements was performed by applying the standard subtraction algorithm based on correlations [32] to the average of the differences between channels Fp1 and F1 and channels Fp2 and F2.

### 2.2 Experimental design

The images for our experiments consisted of 2,400 aerial pictures of London. Images were converted to grayscale and their histograms were equalised. Pic-

ture size was  $640 \times 640$  pixels. Pictures were shown to participants in sequences (or bursts) of 100 images with no gaps between two consecutive stimuli (Inter-Stimulus Interval, ISI=0 ms). Within a burst, 10 images were “target” images, and they differed from the non-targets in that a randomly rotated and positioned plane had been (photo realistically) superimposed as exemplified in figure 1(left). Non-target pictures did not contain planes, as illustrated in figure 1(right).

Approximately 60% (144 out of 240) of our target images contained a lateral target (i.e., a target that appeared on the left or right side of the picture). More specifically, we had 59 Left Visual Field (LVF) target pictures and 85 Right Visual Field (RVF) target pictures. The epochs associated with these images were analysed with particular attention, as lateral-target images were expected to generate N2pc components as well as P300s. We hope that the former would allow the system to localise the plane within target-containing pictures. Targets that did not appear on either side of an image were considered central targets.

We tested RSVP protocols with 4 different “levels of difficulty” which differed in the presentation rate. Each level consisted of 24 bursts which were presented in order of increasing presentation rate at 5, 6, 10 and 12 Hz. Hence, bursts of 100 images lasted between 20 seconds (for the slowest presentation rate) and 8.33 seconds (for the fastest).

Participants were instructed to try to minimise eye blinks and general movements during a burst in order to obtain EEG signals with as few artifacts as possible. They were assigned the task of mentally counting the planes they saw within each burst and were instructed to report the total at the end of a burst (to encourage them to stay focused on the task). Participants could rest after bursts and were free to decide when to start the next sequence. Bursts started upon the participant clicking on a mouse button. Experiments lasted no more than 90 minutes.

## 2.3 Feature selection and classification

We mainly focused on two types of ERPs: the N2pc and the P300. Of course, we expected both to be rarer (or have a reduced amplitude) in response to non-targets than in the case of targets. Also, we did not expect them to be always present together even for targets.

The experimental protocol we used for the two classification tasks (left vs right and target vs non-target) was the same (see section 2.2). However, given the differences in the known characteristics of P300s and N2pcs, we used different feature-sets in order to best detect and exploit each ERP. These will be described in the following two sub-sections.

Collaborative classification with and without group-member selection will then be discussed in sections 2.3.3 and 2.3.4.

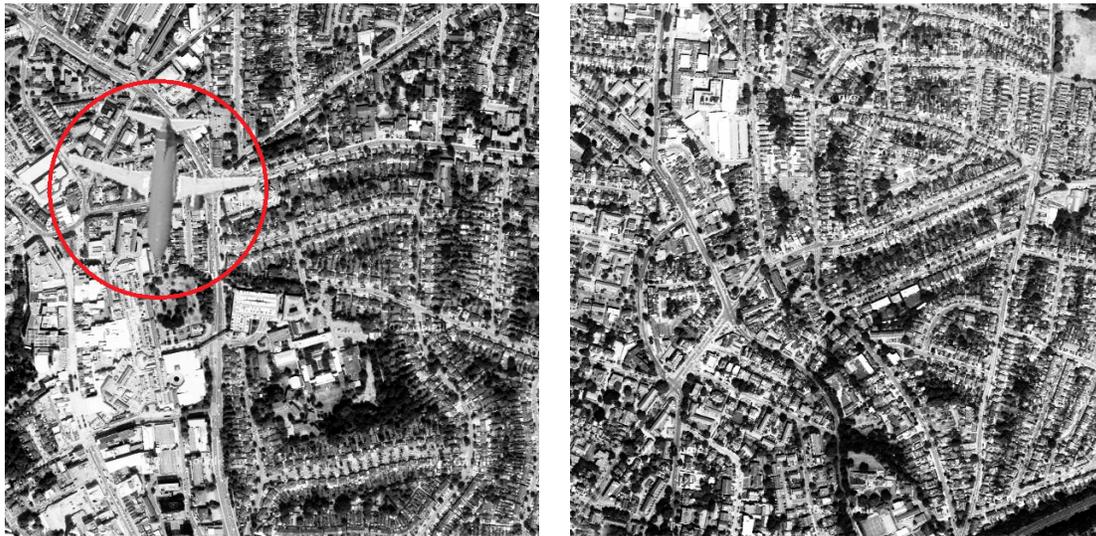


Figure 1: Examples of target (left) and non-target (right) images used in our experiments. The target plane in the image on the left has been highlighted for presentation purposes.

### 2.3.1 Electrodes for left vs right classification

Following the onset of each lateral-target picture on the screen, we extracted 200 ms epochs of EEG signal from approximately 200 ms to 400 ms after stimulus onset (the temporal window where the N2pc most often occurs according to the literature). Including samples at the epoch’s limits (200 and 400 ms, respectively), this resulted in 14 samples per channel at the 64 Hz sampling rate used. The data were referenced to the mean value of the 200 ms interval before stimulus onset.

Following previous literature on the N2pc component (e.g., [21]) and due to the small size of the set of lateral-target images (with the associated potential overfitting risks), we decided to use only four differences between pairs of electrodes (PO7–PO8, P7–P8, PO3–PO4 and O1–O2) for left vs right discrimination. Concatenating these electrode differences yields the feature-vector representation of each epoch used for classification. This includes only  $14 \times 4 = 56$  elements.

We divided the epochs in our set of 144 LVF and RVF pictures into two: 65% of the epochs (corresponding to 55 RVF and 38 LVF images) were used as a training set — which itself was used for 10-fold cross-validation to find the optimal C value when training the SVMs — and the remaining 35% were used as an independent test set.

### 2.3.2 Electrodes for target vs non-target classification

For the classification of target and non-target images we extracted 300 ms epochs of EEG signals (from 300 ms to 600 ms after stimulus onset), resulting in 20 samples per channel at the 64 Hz sampling rate used.

We used only centro-posterior-occipital electrode sites, as these are typically where P300s are most prominent. Based on our previous studies, we used 28 electrodes (Oz, POz, Pz, CPz, CP1–6, TP7–8, P1–10, P7–8, PO3–4, O1–2), so, for the purpose of classification, epochs were represented with 560 features. As before, epochs were referenced to the average voltage in the 200 ms interval before stimulus onset.

For each participant, we used the training set (including 1,200 trials) to find the optimal C parameter for a linear SVM classifier. This was done via 10-fold cross-validation. Each classifier was trained to distinguish between the target and non-target conditions. Once the optimal SVM had been found, this was tested on the epochs from the independent test set.

### 2.3.3 Collaborative classification

We used two methods to merge signals from multiple participants in our cBCIs. First, we averaged the feature vectors across pairs of participants and trained a new SVM classifier for each group (Single Classifier cBCI, SC-cBCI). Since we had already tried this method in previous work and found it to be sub-optimal for P300-based classification [6], we used it only for N2pc detection. In our second method, we averaged the outputs of the individually tailored classifiers from each member of the group, thus creating a Multiple Classifier cBCI (MC-cBCI). We report the results for this method for both the P300-based cBCI and the N2pc-based cBCI.

In order to assess the performance and behaviour of the classifiers, we recorded the *analogue* output scores of the SVMs, with which we then computed the Receiver Operating Characteristic (ROC) curve for each participant. Finally, we condensed the information contained in each ROC curve into a single performance figure: the Area Under the Curve (AUC) [33, 34].

### 2.3.4 Group-member selection

In relation to the selection of group members, we used a method where pairs are formed according to the similarity in performance of individual participants, using different levels of similarity. More specifically we allowed pairs of participants to work as a group if the absolute difference of their AUC values — a value that we term *dissimilarity index* — was below a threshold  $\delta$ . More formally, participants  $i$  and  $j$  formed a pair if

$$|AUC_i^f - AUC_k^f| \times 100\% \leq \delta,$$

where  $AUC_x^f$  represents the AUC value for participant  $x$  (with  $x = 1, \dots, 9$ ) at a presentation rate of  $f$  Hz (with  $f = 5, 6, 10, 12$ ). We created groups by setting the threshold  $\delta$  at 5, 10, 15 and 20% and considered only the cBCIs from pairs of subjects for which the dissimilarity index was below the threshold.

Of course, this selection process reduces the number of groups that can be included in the analysis (from the 36 possible groups of two participants). However, given that cBCIs are conceived with the aim of augmenting human capabilities, it is reasonable to select participants based on their individual performance when forming groups. For comparison, we have included the results when no group selection is performed and all pairs are considered ( $\delta = 100\%$ ).

## 3 Results

In this section we will report results quantifying the performance and behaviour of single-user BCIs (sBCIs) and cBCIs for the single-trial classification of LVF vs RVF (for images that are already known to contain a target, the location of which is unknown) and for target detection. In Section 4 we discuss these results.

### 3.1 Left vs right classification

#### 3.1.1 Performance on lateral targets

In N2pc literature, it is common to refer to the two brain hemispheres as “contralateral” (i.e., the opposite hemisphere to the visual field where the target appears) or “ipsilateral” (i.e., the same hemisphere to the visual field where the target appears) with respect to the appearance of a lateral target. In particular, when plotting grand averages (averages of participant-by-participant averages), these show the differences between contralateral and ipsilateral ERPs. Figure 2 shows these grand-averaged differences across all lateral-target epochs from the training set, for different presentation rates, measured at electrode sites PO7 and PO8. In order to compute it, we obtained the contralateral waveform as the average of the epochs recorded from channel PO7 (on the posterior-occipital *left* region of the scalp) for all RVF targets with the epochs recorded from channel PO8 (on the posterior-occipital *right* region of the scalp) for all LVF targets. Similarly, the ipsilateral epoch consists of the average of the epochs recorded from channel PO7 for LVF targets with the epochs recorded from channel PO8 for all RVF targets for each participant.

Also following the N2pc conventions, we plotted these using an *inverted ordinate axis*, so higher means more negative. When the presentation rate is increased (up to 10 Hz), the latency of the N2pc (as measured by the time when it reaches its peak) is shortened. We can also see from this figure how peak amplitudes decrease as presentation rates increase.

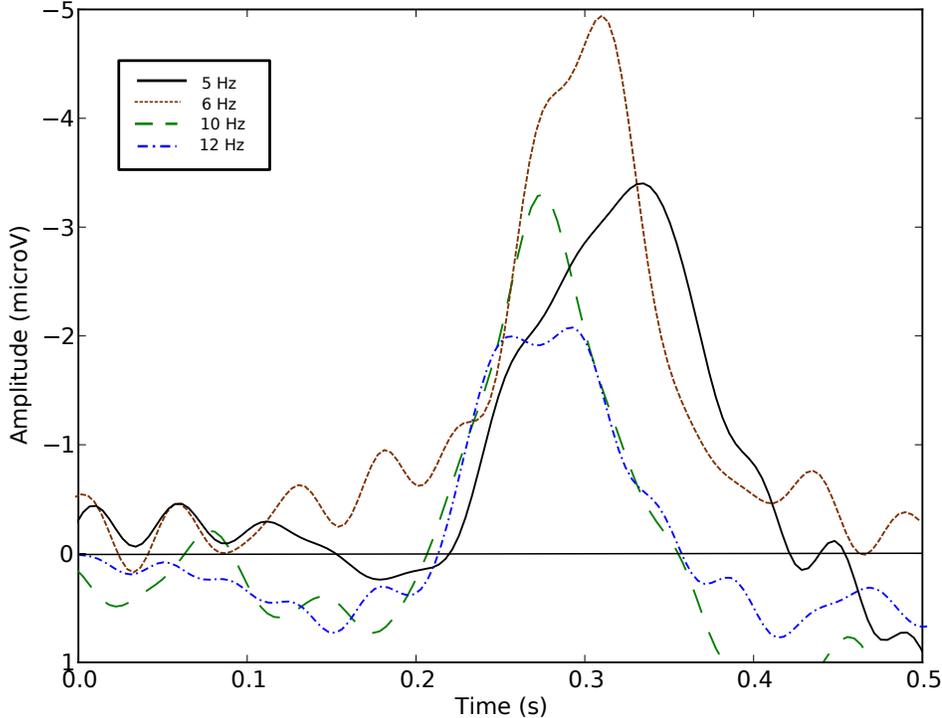


Figure 2: Difference plot of the contralateral minus the ipsilateral grand-averages at channels PO7 and PO8 across all lateral targets from the training set. Amplitudes are measured in  $\mu V$ .

The first row of table 1 shows the median AUC values obtained for left vs right classification for single-user BCIs for each level of difficulty. Consistently with the ERP plots from figure 2, performance decreases at the highest presentation rate. This figure shows how at 12 Hz the difference between contralateral and ipsilateral electrode sites is much reduced in both amplitude and duration, making it more difficult to be detected by the BCI. This decrease in amplitude might be caused by either the uncertainty of the participant at such high presentation rate<sup>1</sup> or the temporal proximity of lateral targets within a burst for high speeds, which might cause subsequent targets to fall within a possible refractory period for this ERP.

The remaining rows of the table report the median gains in performance over the better participant of each pair for each stimulation frequency, separately for our two types of collaborative BCIs — SC-cBCIs and MC-cBCIs — for different values of the dissimilarity-index threshold  $\delta$ .

With 9 participants, in principle we can form up to 36 distinct pairs, but

<sup>1</sup>Indeed, the reported number of planes for 12 Hz was lower than for slower rates, showing that many targets were missed by participants, and those that do not fall within the foveated area are more likely to be missed.

Table 1: Median AUC values for single-user BCIs and median improvement over the best participant in the group when using collaborative BCIs, as a function of presentation rate and the dissimilarity-index threshold  $\delta$ .

<i>Method</i>	$\delta$	<i>5 Hz</i>	<i>6 Hz</i>	<i>10 Hz</i>	<i>12 Hz</i>
sBCI	N/A	77.6%	76.8%	79.8%	66.5%
SC-cBCI	5%	+7.7%	+3.8%	-1.1%	+1.2%
	10%	+7.6%	+0.2%	+1.1%	+0.2%
	15%	+5.2%	+0.2%	+1.0%	-0.3%
	20%	+5.2%	-1.6%	+1.0%	-1.1%
	100%	+2.2%	-4.0%	-1.6%	-1.6%
MC-cBCI	5%	+6.5%	+3.2%	+4.1%	+3%
	10%	+6.5%	+2.2%	+3.6%	+1.6%
	15%	+5.6%	+1.9%	+2.6%	+0.3%
	20%	+5.6%	+1.0%	+2.6%	0.0%
	100%	-0.5%	-6.8%	0.0%	0.0%

Table 2: Percentages of groups that are accepted by our selection mechanism for different values of the stimulation frequency and the dissimilarity-index threshold  $\delta$ .

$\delta$	<i>5 Hz</i>	<i>6 Hz</i>	<i>10 Hz</i>	<i>12 Hz</i>
5%	41.7%	19.4%	30.6%	22.2%
10%	47.2%	41.7%	41.7%	44.4%
15%	66.7%	50.0%	50.0%	63.9%
20%	66.7%	63.9%	50.0%	80.6%
100%	100.0%	100.0%	100.0%	100.0%

when using a pair selection strategy one can only accept fewer pairs. In table 2 we quantify the effects that different values for the threshold  $\delta$  have on the fraction of pairs that can be accepted.

### 3.1.2 Performance on central targets

Since our SVMs can give an analogue output score for each epoch, we fed the cBCIs with epochs from the non-lateral targets (that the classifiers had not seen before) in order to check how their behaviour changed when presented with targets that are closer to the centre of the screen than those used for training.

As an illustration, in figure 3 we plotted the horizontal position of the target (as indicated by the abscissa of its centroid) for each target picture (lateral and central targets) against the raw output from the MC-cBCI for a presentation rate of 5 Hz and a dissimilarity index of 5%. In this plot, using the values from all the groups included, we obtained a correlation coefficient of -0.44, showing that the BCI can give an indication of where the target is located.

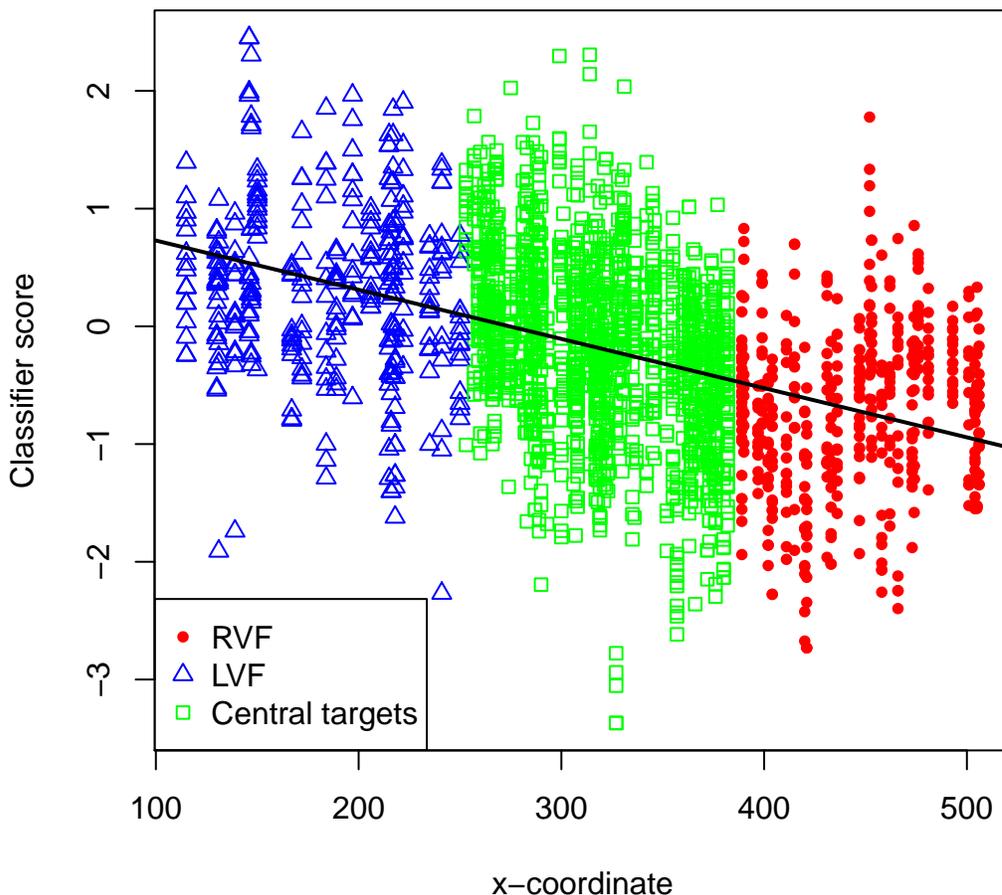


Figure 3: Plot of target positions (in pixels, for  $640 \times 640$  px images) vs MC-cBCI SVM output scores for LVF, RVF and central targets, at 5 Hz, for groups formed with a dissimilarity index of 5% (correlation coefficient =  $-0.44$ ). The linear regression line is also shown.

### 3.2 Target vs non-target classification

The first row of table 3 shows the median AUC values obtained for target vs non-target classification for single-user BCIs for each level of difficulty. Consistently with the literature, the performance of this BCI decreases with increasing presentation rates. We also observed that some individuals perform better at 6 Hz than they do at 5 Hz, a phenomenon that had also been previously reported [29]. The performance for higher rates is probably related to a mixture of the effect of the refractory period of the P300 and RSVP-related issues, such as the attentional blink. As we pointed out above, at high speeds of presentation the reported number of planes dramatically decreases for all participants, so lower AUCs for such presentation rates should not be surprising.

Table 3: Median AUC values for single-user BCIs (top), median AUC value for MC-cBCIs (middle) and median improvement over the better participant in the group when using collaborative BCIs (bottom), as a function of presentation rate and the dissimilarity-index threshold  $\delta$ .

<i>Method</i>	$\delta$	<i>5 Hz</i>	<i>6 Hz</i>	<i>10 Hz</i>	<i>12 Hz</i>
sBCI	N/A	87.2%	86.5%	74.6%	61.6%
MC-cBCI	5%	94.6%	92.8%	82.9%	65.1%
	10%	93.8%	92.6%	81.8%	65.7%
	15%	93.6%	92.1%	80.6%	67.3%
	20%	93.0%	91.1%	78.2%	69.0%
	100%	92.7%	90.9%	79.3%	69.0%
Improvement over better performer	5%	+3.1%	+3.4%	+3.9%	+3.5%
	10%	+2.7%	+3.1%	+3.3%	+1.1%
	15%	+2.0%	+2.2%	+0.9%	+0.7%
	20%	+1.7%	+1.9%	0.0%	0.0%
	100%	+1.2%	+1.5%	0.0%	0.0%

Table 4: Percentages of groups that are accepted by our selection mechanism for different values of the stimulation frequency and the dissimilarity-index threshold  $\delta$ .

$\delta$	<i>5 Hz</i>	<i>6 Hz</i>	<i>10 Hz</i>	<i>12 Hz</i>
5%	25.0%	33.3%	38.9%	30.6%
10%	52.8%	58.3%	55.6%	61.1%
15%	69.4%	72.2%	66.7%	77.8%
20%	80.6%	86.1%	86.1%	94.4%
100%	100.0%	100.0%	100.0%	100.0%

The second block of the table reports the actual median AUC values obtained by pairs for the MC-cBCIs, for different values of the dissimilarity-index threshold  $\delta$  and for each stimulation frequency. The third block, instead, reports the median difference between the performance of an MC-cBCI and the performance of the better participant of the corresponding pair. Note that we report these values to illustrate the benefits of a pair over the better participant in it, and that they are *not* the differences between the AUCs of an MC-cBCI and a corresponding sBCI.

Finally, Table 4 quantifies the effects that different values for the threshold  $\delta$  have on the fraction of pairs that can be accepted (out of the possible 36).

## 4 Discussion

### 4.1 Left vs right classification

#### 4.1.1 Single-user BCI

As we have seen in figure 2, in the left vs right classification task, the N2pc ERP changes when varying the presentation rate of our RSVP paradigm (whilst keeping the ISI = 0).

In particular, we observed a reduction in the duration of this ERP and a decrease in its latency at the two highest presentation rates. For a presentation frequency of 10 Hz, the shape and timing of the N2pc was consistent with those reported in the literature. However, the N2pc amplitude dropped significantly at 12 Hz. This decrease in amplitude might be caused by either the uncertainty of the participant at such high presentation rate (as we mentioned above, the reported number of planes was much lower than for slower rates, showing that many targets were missed by participants) or the close temporal proximity of lateral targets within a burst, which might cause some of them falling within a possible refractory period for this ERP.

The amplitude of the N2pc has been linked to subject engagement, and, so, we expected it to vary as a function of the presentation rate. The higher amplitude observed for 6 Hz over that at 5 Hz might be linked to participants being more attentive for this second level of difficulty, as the task’s demands increased.

Let us now turn our attention to the left vs right classification results in the top row of table 1. These indicate that the N2pc can reliably be detected by an sBCI in the single-trial conditions of our experiments for presentation rates of up to 10 Hz (the median AUC value is almost 80%). In fact, performance seems to increase in the interval 5–10 Hz to then start decreasing for higher speeds. Still, even for the rates as high as 12 Hz, most participants are well above chance levels with the top quartile of our participants showing AUCs  $\geq 72.2\%$ .

#### 4.1.2 Collaborative BCI

In this paper, we also showed that collaborative BCIs can outperform “traditional” single-user BCIs when only similar performers are allowed to form pairs. Since our BCI systems are designed for able-bodied users, as opposed to traditional BCIs, participants could conceivably be selected based on performance and neural responses so as to best match the requirements of our BCIs. Thus, while performance variance across participants is a traditional worry for BCI, it is less so for our systems, both in the individual and in the collaborative forms. In any case, as illustrated in table2, our selection method is not unreasonably stringent, typically accepting 40+% pairs.

Consistently with our findings in [6], by increasing the required performance similarity when pairing participants, we were able to increase further the perfor-

mance of groups both w.r.t. that of the better participant individually and w.r.t. sBCIs.

This is reflected by the results in table 1, where lower values of the dissimilarity index obtain higher improvements over the better of the two members than higher values. Also, when  $\delta = 100\%$  we see that cBCIs are almost always either worse or on par with corresponding sBCIs.

This seems reasonable, considering that when a participant with a high AUC is paired with a low scorer, the extra information of the latter w.r.t. the former is not enough to translate into an improvement in the performance of the better one.

If we now compare the absolute improvements across our two types of cBCIs (also reported in table1), we can see that at the two lowest presentation rates the reported improvements are similar for SC-cBCI and MC-cBCI. However, for presentation rates of 10 Hz or more, MC-cBCIs perform better than corresponding SC-cBCIs.

Finally, the correlation between horizontal position of targets and classifiers' outputs (see figure 3) revealed that the N2pc can not only be used to distinguish between LVF and RVF targets, but it can also tell to what degree a target is lateral. Moreover, separate groups can be better at locating targets than the overall performance shown on figure 3. Even though we have not studied this in depth here, some groups showed correlation coefficients greater (in absolute value) than 0.5.

## 4.2 Target vs non-target classification

### 4.2.1 Single-user BCI

Classification results for the single-trial sBCI for target detection indicate that the P300 can reliably be detected in the conditions of our experiments for presentation rates of 5 and 6 Hz. However, beyond this, performance rapidly degrades. At 10 Hz, the AUC is more than 10% lower than for 5–6 Hz, and at 12 Hz, it is 25% lower. Interestingly, and somehow surprisingly given the smaller amplitude of the N2pc w.r.t. the P300, this did not happen in the left vs right classification discussed in the previous sections, where AUCs for 5, 6 and 10 Hz are almost indistinguishable, and they only drop by approximately 10% at 12 Hz.

We should note that here we obtained lower median AUCs than those reported by others for sBCIs (e.g., [11]). However, in our framework the number of targets within a burst is much higher than those used by others, and at high presentation rates, the temporal separation between two targets is likely to fall well within the time frame of the attentional blink and/or the refractory period of the P300, thus causing a drop in performance.

We feel that this is a reasonably price to pay for a more realistic environment in which the rate of targets vs non-targets is compatible not only with intelligence

analysis but also with other screening frameworks, e.g., for medical applications. Also, as discussed below, performance can improve very substantially by using cBCIs.

#### 4.2.2 Collaborative BCI

In this paper, we expanded on previous work [6] by including a method for participant selection when forming groups of two for a cBCI.

Consistent with our previously reported results, we have shown that by pairing users into an MC-cBCI, but without performing any selection, we can improve the median AUC by about 5% with respect to single-user BCIs when all pairs are considered.

However, by creating pairs using participants with similar AUC scores, we were able to reach AUC scores well above 90% and improve the BCI performance by up to 8% (for a presentation rate of 10 Hz and with a dissimilarity-index threshold  $\delta=5\%$ ).

It should be noted that by pairing subjects according to the dissimilarity index, we are not excluding the worst performers (i.e. the higher median AUCs are not due to the fact that we are excluding observers with low AUCs). Rather, we are making sure that groups are formed by performance-matched members, where the closer their performance (the lower the threshold  $\delta$ ), the better the cBCI AUCs.

## 5 Conclusions

In this paper, we used an RSVP protocol to present aerial images of an urban environment to participants looking for predefined targets (airplanes) at rates ranging from 5 to 12 Hz. We considered two tasks (detecting targets and approximately establishing their horizontal position within the pictures, a task that we have proposed here for the first time), two BCI approaches (single-user and collaborative BCIs), two forms of collaborative BCI (using a single classifier to process the averages of the raw signals from users and thresholding the average of multiple single-user classifiers), and two forms of membership selection for groups (all-pairs allowed and performance-matched pairs).

Our results conclusively indicate that cBCIs, particularly when pairs are formed by individuals with similar performance, offer a 5 to 10% performance improvement (as evaluated by the AUC scores) over the corresponding single-user BCIs. Furthermore, we found that there is a significant correlation between the features of the N2pc (as represented by the SVM’s output score) and the horizontal position of targets within images, which suggests a whole spectrum of possible BCI applications for this ERP in the future.

Future research should explore ways of combining our P300 and N2pc classifiers, which is an obvious next step once it has been shown that it is possible to detect both ERPs independently. With the lessons learnt from this work, we can now envision a cascade of the two classifiers: the first one would decide whether a given image contains or not a target (P300 detection); the second (left vs right classifier) would help limit the area of search within a given image when a target has been detected in the first step. Thus, it would be possible to improve current visual search RSVP systems by roughly locating targets after detection, which would in turn reduce the workload of an external observer that had to manually check the images classified by the system as targets.

Moreover, in future research we will need to extend the work to different targets and types of images, to see to what extent it is possible to build BCIs that can be used for target detection and localisation across a range of target types. We would also like to study the appearance and form of the P300 and N2pc components during videos, e.g., for video surveillance.

## Acknowledgements

The authors would like to thank the UK's Engineering and Physical Sciences Research Council (EPSRC) for financially supporting the early stages of this research (grant EP/K004638/1, entitled "Global engagement with NASA JPL and ESA in Robotics, Brain Computer Interfaces, and Secure Adaptive Systems for Space Applications"). Dr Caterina Cinel is also warmly thanked for contributions to the early stages of this research.

## References

- [1] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, Dec 1988. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0013469488901496>
- [2] R. Scherer and G. Muller, "An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 979–984, 2004. [Online]. Available: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1300792](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1300792)
- [3] L. Citi, R. Poli, C. Cinel, and F. Sepulveda, "P300-based BCI mouse with genetically-optimized analogue control," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 16, no. 1, pp. 51–61, Feb. 2008.

- [4] Y. Wang and T.-P. Jung, “A Collaborative Brain-Computer Interface for Improving Human Performance,” *PLoS ONE*, vol. 6, no. 5, May 2011. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0020422>
- [5] A. Stoica, A. Matran-Fernandez, D. Andreou, R. Poli, C. Cinel, Y. Iwashita, and C. W. Padgett, “Multi-brain fusion and applications to intelligence analysis,” in *Proceedings of SPIE Volume 8756*, Baltimore, Maryland, USA, 30 April – 1 May 2013.
- [6] A. Matran-Fernandez, R. Poli, and C. Cinel, “Collaborative brain-computer interfaces for the automatic classification of images,” in *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*. San Diego (CA): IEEE, 6–8 November 2013, pp. 1096–1099.
- [7] P. Yuan, Y. Wang, W. Wu, H. Xu, X. Gao, and S. Gao, “Study on an online collaborative BCI to accelerate response to visual targets,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society*, 2012, pp. 1736–1739.
- [8] R. Poli, C. Cinel, F. Sepulveda, and A. Stoica, “Improving Decision-making based on Visual Perception via a Collaborative Brain-Computer Interface,” in *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*. San Diego (CA): IEEE, February 2013.
- [9] A. D. Gerson, L. C. Parra, and P. Sajda, “Cortically coupled computer vision for rapid image search,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 14, no. 2, pp. 174–179, Jun. 2006. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/16792287>
- [10] A. Kruse and S. Makeig, “Phase I analysis report for UCSD / SoCal NIA team,” Institute for Neural Computation, University of California San Diego, La Jolla, Tech. Rep. January, 2007.
- [11] Y. Huang, D. Erdogmus, M. Pavel, S. Mathan, and K. E. Hild II, “A framework for rapid visual image search using single-trial brain evoked responses,” *Neurocomputing*, vol. 74, no. 12, pp. 2041–2051, 2011.
- [12] K. Forster, “Visual perception of rapidly presented word sequences of varying complexity,” *Perception & Psychophysics*, vol. 8, no. 4, pp. 215–221, 1970. [Online]. Available: <http://dx.doi.org/10.3758/BF03210208>
- [13] S. Mathan, D. Erdogmus, Y. Huang, M. Pavel, P. Ververs, J. Carciofini, M. Dorneich, and S. Whitlow, “Rapid image analysis using neural signals,”

- in *CHI'08 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2008, pp. 3309–3314.
- [14] N. K. Squires, K. C. Squires, and S. A. Hillyard, “Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man,” *Electroencephalography and clinical neurophysiology*, vol. 38, no. 4, pp. 387–401, 1975.
- [15] L. Farwell and E. Donchin, “Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials,” *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, Dec 1988.
- [16] G. Pfurtscheller and F. H. Lopes da Silva, “Event-related eeg/meg synchronization and desynchronization: basic principles,” *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [17] H. Cecotti, “Spelling with non-invasive brain-computer interfaces - current and future trends,” *Journal of Physiology - Paris*, vol. 105, pp. 106–114, 2011.
- [18] H. Awni, J. J. Norton, S. Umunna, K. D. Federmeier, and T. Bretl, “Towards a brain computer interface based on the N2pc event-related potential,” in *6th Annual International IEEE EMBS Conference on Neural Engineering*. San Diego (CA): IEEE, 6–8 November 2013.
- [19] S. J. Luck and S. A. Hillyard, “Spatial filtering during visual search: evidence from human electrophysiology,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 20, no. 5, pp. 1000–1014, 1994.
- [20] M. Eimer, “The N2pc component as an indicator of attentional selectivity,” *Electroencephalography and clinical neurophysiology*, vol. 99, no. 3, pp. 225–234, 1996.
- [21] S. Luck, “Electrophysiological correlates of the focusing of attention within complex visual scenes: N2pc and related ERP components,” *Oxford Handbook of ERP components*, 2012.
- [22] E. Donchin, K. M. Spencer, and R. Wijesinghe, “The mental prosthesis: assessing the speed of a p300-based brain-computer interface,” *Rehabilitation Engineering, IEEE Transactions on*, vol. 8, no. 2, pp. 174–179, 2000.
- [23] M. P. Eckstein, K. Das, B. T. Pham, M. F. Peterson, C. K. Abbey, J. L. Sy, and B. Giesbrecht, “Neural decoding of collective wisdom with multi-brain computing.” *NeuroImage*, vol. 59, no. 1, pp. 94–108, 2012.
- [24] J. Surowiecki, *The wisdom of crowds*. Random House LLC, 2005.

- [25] A. B. Kao and I. D. Couzin, “Decision accuracy in complex environments is often maximized by small group sizes.” *Proceedings of the Royal Society B: Biological Sciences*, vol. 1, 2014.
- [26] B. Bahrami, K. Olsen, P. E. Latham, A. Roepstorff, G. Rees, and C. D. Frith, “Optimally interacting minds,” *Science*, vol. 329, no. 5995, pp. 1081–1085, 2010. [Online]. Available: <http://www.sciencemag.org/cgi/content/abstract/329/5995/1081>
- [27] P. Yuan, Y. Wang, W. Wu, H. Xu, X. Gao, and S. Gao, “Study on an online collaborative BCI to accelerate response to visual targets,” in *Proceedings of 34nd IEEE EMBS Conference*, 2012.
- [28] P. Yuan, Y. Wang, X. Gao, T.-P. Jung, and S. Gao, “A collaborative brain-computer interface for accelerating human decision making,” in *Proceedings of the 7th International Conference on Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for eInclusion (UAHCI) 2013*, ser. LNCS, C. Stephanidis and M. Antona, Eds., vol. 8009. Las Vegas, NV, USA: Springer, July 2013, pp. 672–681, part I.
- [29] G. Healy, P. Wilkins, A. F. Smeaton, D. Izzo, M. Rucinski, C. Ampatzis, and E. M. Moraud, “Curiosity cloning: Neural modelling for image analysis, technical report,” Dublin City University; European Space and Technology Research Center (ESTEC), Dublin, Ireland; The Netherlands, Tech. Rep. 0, 2010.
- [30] H. Cecotti and B. Rivet, “Subject combination and electrode selection in cooperative brain-computer interface based on event related potentials,” *Brain Sciences*, vol. 4, no. 2, pp. 335–355, 2014. [Online]. Available: <http://www.mdpi.com/2076-3425/4/2/335>
- [31] H. Cecotti, B. Rivet *et al.*, “Performance estimation of a cooperative brain-computer interface based on the detection of steady-state visual evoked potentials,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014)*, 2014, pp. 2078–2082.
- [32] P. Quilter, B. MacGillivray, and D. Wadbrook, “The removal of eye movement artefact from EEG signals using correlation techniques,” in *Random Signal Analysis, IEEE Conference Publication*, vol. 159, 1977, pp. 93–100.
- [33] J. Hanley and B. McNeil, “The meaning and use of the area under a receiver operating characteristic (ROC) curve.” *Radiology*, vol. 143, pp. 29–36, 1982.
- [34] A. P. Bradley, “The use of the area under the ROC curve in the evaluation of machine learning algorithms,” *Pattern recognition*, vol. 30, no. 7, pp. 1145–1159, 1997.