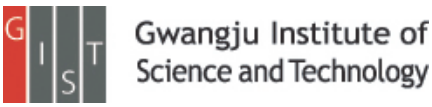
An auditory brain–computer interface evoked by natural speech

M A Lopez-Gordo, E Fernandez, S Romero, F Pelayo and Alberto Prieto

Journal of Neural Engineering (2012)

Presenter : Younghak Shin

GIST, Dept. of Information and Communication, INFONET Lab.



Introduction

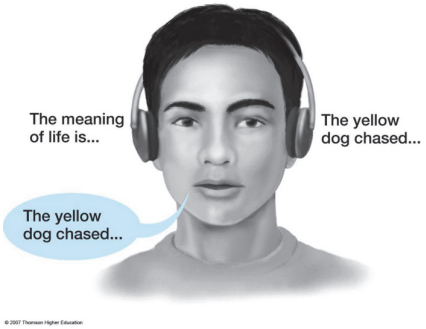
- The vast majority of BCIs use visual interaction with the subjects, either as stimulation, biofeedback or as visual support for the paradigm.
- Examples are the BCIs based on steady-state visual evoked potentials or the P300.
- However, some patients (in an advanced stage of ALS with visual impairment or patients under the condition of unresponsive wakefulness syndrome, who still preserve the necessary level of consciousness to process auditory stimuli and follow verbal instructions) can not use the above BCIs.
- In this regard, an auditory BCI seems to be a more appropriate solution.
- There have not been many attempts to use the auditory modality instead of the visual one.

Introduction

- The auditory BCIs present some inconveniences that have not been suitably addressed so far.
- For instance, the number of channels used (e.g. 67 and 16 channels in [10, 27] respectively), because it can be a limitation in the usability and increases the time for the initial setup.
- This BCI requires considerable cognitive effort in terms of training (up to several weeks [33]), length of the trials or number of stimuli
- The use of synthetic sounds instead of natural ones (e.g. tone beeps [10]) or complex auditory paradigms [27, 34] - since they can affect the level of interest and cognitive effort of the participants.
- In this paper, they postulate that an auditory BCI based on natural sound (e.g. human voice), a simple and low-demanding task (e.g. selective attention), a small number of electrodes and a simple paradigm (e.g. binary class) with little or no training.

Objective

- The main goal of this study is to determine the viability of a BCI based on a dichotic listening task with natural speech
  - **Dichotic listening** is a psychological test commonly used to investigate selective attention within the auditory system
  - pay attention to the stream of stimuli delivered to one ear and to ignore the stream delivered to the other one



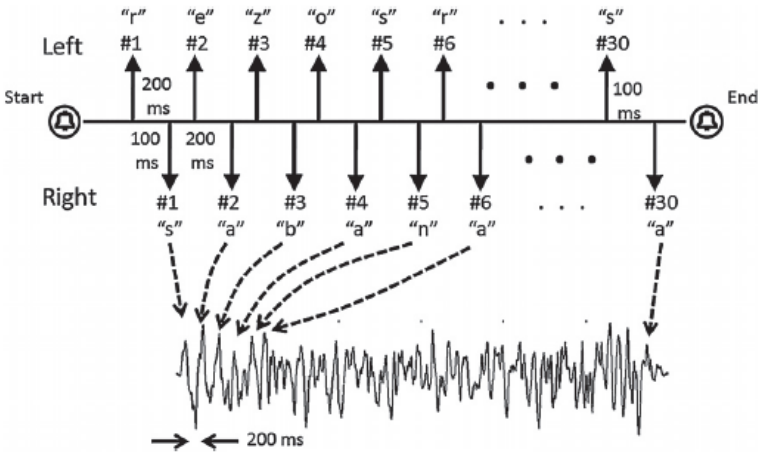
- The usability is expected to be improved due to a careful design based on:
  - first, a minimum number of EEG channels, namely, just one, which makes this BCI suitable for subjects attached to electronic devices in a hospital or lying on a bed;
  - second, minimum cognitive effort by means of short trials;
  - third, no need of formal training, which was substituted for a calibration run and the use of the native-to-human selective attention in a dichotic listening paradigm.

Methods

- The experiment is based on the dichotic listening paradigm
- The essence of dichotic listening is to pay attention to the stream of stimuli delivered to one ear and to ignore the stream delivered to the other one.
- A dichotic listening paradigm was established by means of two streams of stimuli, one per ear, which were presented simultaneously to the subject through earphones
- Five stimuli are given per second (ISI = 200 ms) and an interleaving interval (ILI) is 100 ms between streams.
- The beginning and the end of each trial was signaled by means of a high beep and a low beep, respectively.

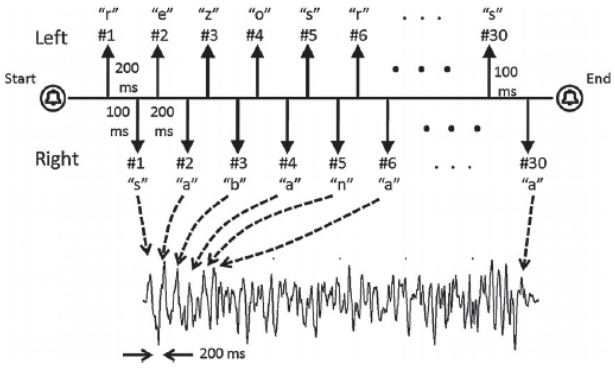
Methods

- The figure shows the typical structure of a evaluation trial
- After the beep, the trial starts with two streams of 30 stimuli per ear.
- Each stimulus corresponds to a letter from words with five and six letters (left and right respectively) that are repeated six and five times respectively.
- In this example the words ‘rezos’ and ‘sabana’ were used for both streams.
- The lower graph shows the typical steady response obtained after the average of the calibration trials for one of the streams.



Methods

- It is established that early components of the human auditory evoked potentials, namely N1 and P2 are larger when the stimulus is attended compared to when it is ignored [38, 39].
- This effect can be observed in the dichotic listening paradigm.
- Because the ISI and ILI of each stimulation were fixed to a value (respectively, 200 and100 ms)
- for the two streams of stimuli, to two counter-phased responses of the same frequency (5 Hz =ISI−1) whose amplitudes can be cognitively modulated.
- The net effect is a sinusoidal-shaped response with most of the energy located around the frequency of 5 Hz



Subjects and recording

- A total of 12 healthy subjects (7 men, 5 women; aged 31.1 ± 6.0 years) participated in the experiment.
- No type of auditory or cognitive disease was reported by the subjects
- The experiment was conducted in a quiet room isolated from external disturbances throughout.
- As visual stimulation was not needed, the subjects were recommended to close their eyes during the whole experiment.
- They configured a single channel with an active electrode placed on the vertex (Cz) and referenced to the mean value of the mastoids.
- The ground electrode was placed between the Fpz and the Fz.
- The recordings were acquired on Neuroscan, were band-pass filtered between 1 and 30 Hz and were sampled at a rate of 1 kHz.

Stimulation

- Two conditions were established:
  - first, a dichotic listening condition wherein two distinct streams of letters were read out simultaneously (condition 1)
  - second, the same as condition 1 except that complete sentences were used rather than letters (condition 2).
- Letters in condition 1 that were randomly picked from a dictionary with more than 13 000 entries.
- The letters were previously sampled (16 bits per sample at 44100 Hz and 200 ms length) from a male voice that read the Spanish alphabet.
- a total number of 30 stimuli (6 × 5 or 5 × 6) were delivered to each ear
- The total duration of the stimulation was 6.1 s.
- For condition 2, two different speeches were delivered simultaneously, one per ear
- The sentences were previously recorded with the same characteristics as depicted for condition 1.
- The subjective perception of the stimuli of condition 2 is similar to a low-quality cellular phone call

Experimental design

- Because the experiment was purely auditory without any visual stimulation, an assistant was needed to give verbal statements to the participants.
- The experiment consisted of two sessions, namely the spelling session (condition 1) and the speech session (condition 2).
- Each session consisted of five runs that, in turn, consisted of ten trials.
- The first run of each session was used for calibration and the other four were used for evaluation.
- Because selective attention is a native skill for humans, no training was needed to perform the dichotic listening task
- only the first ten trials (first run) of each session dedicated to calibrate the system (calibration for classifier).
- During the calibration, only one stream was delivered at a time to one of the ears.
- The result was that a very stable and reliable pattern was extracted in just one run.

Experimental design

- During the evaluation, and previous to the stimulation onset, an auditory sentence was read by an assistant to the participant
- After beep, the participant had to take a binary decision, either 'yes/true' or 'no/false', for that sentence.
- The sentence was randomly picked from a list and sentences were composed in such a way that the number of positive and negative answers was the same.
- e.g., 'I am Spanish', 'I am not Spanish' or 'two plus three equals five', 'two plus three equals six'
- If the correct answer was 'yes/true', the participant had to pay attention to the stimulus delivered to the left ear and vice versa if the decision was 'no/false'.

Experimental design

- In general, the communication was intended as a conversation initiated by the assistant in which he asks binary response questions to which the participant answers
- e.g., 'Do you need pain relief?', or 'Would you prefer a warmer room?'.
- The correct word was given to the participants as feedback at the end of each trial
- The feedback was solely intended to help them to find a more efficient strategy to focus their attention.

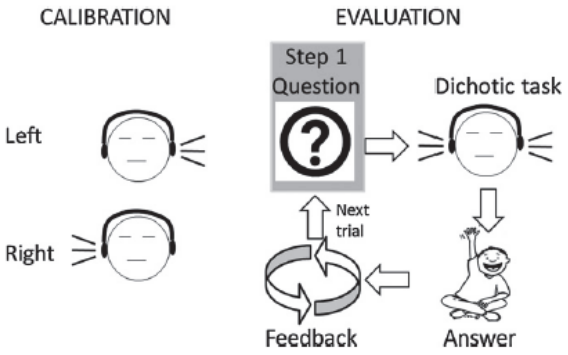
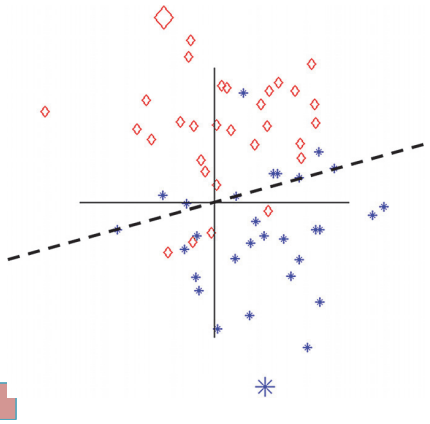


Figure 2. Synoptic representation of the procedure followed in the experiment.

Feature extraction and Classification

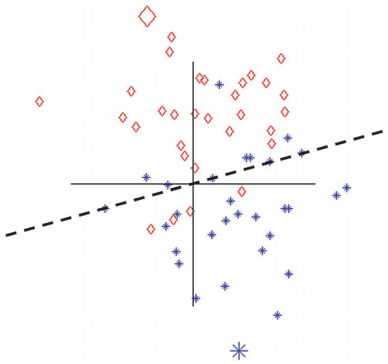
- The data collected during the calibration trials were separated into the left ear and right ear data; these two types of data were averaged separately.
- The DFT was computed on each and the DFT coefficients at the frequency of repetition (5 Hz) were extracted and used as patterns for classification.
- These coefficients are complex numbers that convey information about the amplitude and phase of the event-related responses.
- As both streams were presented at the same frequency but with an ILI of half period (100 ms), the DFT coefficients differed only in the phase (180◦).
- The typical representation of the DFT coefficients in polar coordinates



(stars and diamonds represent the left and right, respectively; the big star and diamond for the patterns).

Feature extraction and Classification

- Given that during the evaluation both sequences were delivered simultaneously, the amplitudes of the features extracted are smaller than the patterns.
- Because the responses to both sequences are counter-phased (180◦ phase shift), they partially cancel each other, thus giving rise to lower levels of signal.
- During the evaluation, the extracted features of each trial were compared to the pattern and were classified under the criterion of the minimum Euclidean distance: $S_t = \min_i |\psi_t - \psi_i|$,
- where S_t is the classification decision for trial *t*, ψ_i are the DFT coefficients of the two patterns extracted, ψ_t is the DFT coefficient extracted from trial *t*



Results

	Spelling session				
Subject	acc (%)	ITR (bits m^{-1})	# of stimuli	acc (%)	ITR (bits m^{-1})
S01	73	1.55	26	68	1.13
S02	83	3.31	13	78	5.33
S03	80	2.78	9	73	5.05
S04	68	0.90	28	68	0.97
S05	75	1.89	18	75	3.15
S06	65	0.66	17	60	0.51
S07	75	1.89	5	73	9.09
S08	75	1.89	15	70	2.37
S09	50	0.00	30	50	0.00
S10	80	2.78	14	75	4.04
S11	50	0.00	30	50	0.00
S12	60	0.29	30	60	0.29
GAV1	69	1.49	20	67	2.66
GAV2	73	1.79	18	70	3.19

	Speech session				
Subject	acc (%)	ITR (bits m^{-1})	# of stimuli	acc (%)	ITR (bits m^{-1})
S01	63	0.23	46	70	0.77
S02	93	3.08	20	73	2.27
S03	78	1.15	9	70	3.96
S04	85	1.95	27	65	0.73
S05	85	1.95	31	70	1.15
S06	65	0.33	60	65	0.33
S07	75	0.94	30	65	0.66
S08	73	0.76	48	70	0.74
S09	53	0.01	60	48	0.00
S10	80	1.39	18	83	5.52
S11	45	0.00	60	53	0.02
S12	83	1.65	20	65	0.99
GAV1	73	1.12	36	66	1.43
GAV2	78	1.34	31	70	1.71

- The experiment did not work for two of the twelve participants (S09, S11).
- Columns two and three denote the full-length trials and last three columns denote the optimum number of stimuli.
- Two grand averages across the subjects are presented, one across all of the participants (GAV1) and the other excluding S09 and S11 (GAV2).

Results

	Spelling session				
Subject	acc (%)	ITR (bits m^{-1})	# of stimuli	acc (%)	ITR (bits m^{-1})
S01	73	1.55	26	68	1.13
S02	83	3.31	13	78	5.33
S03	80	2.78	9	73	5.05
S04	68	0.90	28	68	0.97
S05	75	1.89	18	75	3.15
S06	65	0.66	17	60	0.51
S07	75	1.89	5	73	9.09
S08	75	1.89	15	70	2.37
S09	50	0.00	30	50	0.00
S10	80	2.78	14	75	4.04
S11	50	0.00	30	50	0.00
S12	60	0.29	30	60	0.29
GAV1	69	1.49	20	67	2.66
GAV2	73	1.79	18	70	3.19

	Speech session				
Subject	acc (%)	ITR (bits m^{-1})	# of stimuli	acc (%)	ITR (bits m^{-1})
S01	63	0.23	46	70	0.77
S02	93	3.08	20	73	2.27
S03	78	1.15	9	70	3.96
S04	85	1.95	27	65	0.73
S05	85	1.95	31	70	1.15
S06	65	0.33	60	65	0.33
S07	75	0.94	30	65	0.66
S08	73	0.76	48	70	0.74
S09	53	0.01	60	48	0.00
S10	80	1.39	18	83	5.52
S11	45	0.00	60	53	0.02
S12	83	1.65	20	65	0.99
GAV1	73	1.12	36	66	1.43
GAV2	78	1.34	31	70	1.71

- S09 and S11 were not able to transmit any amount of information (ITR = 0.0, accuracy = 50%).
- However, when both are excluded to compute GAV2, accuracies exceed 70%,with an ITR of 3.19 (70%- 1.71 for speech session) for the optimal number of stimuli.

Results

- For the optimal number of stimuli, incremental fractions of trials in steps of 200 ms (each including unitary incremental of responses)were extracted and classified.
- The optimal length of trials was defined, for each subject, as the mean of two numbers.
- The first number is calculated as the number of stimuli for which fewer stimuli always cause an accuracy equal to or below 70%.
- The second number is the number of stimuli for which a higher number of stimuli always yield an accuracy equal to or above 70%.
- According to [45], 70% corresponds to the minimum accuracy that guarantees meaningful communication.

Results

- Figure 4 (left) shows the averaged performance across subjects versus the number of stimuli used for the classification of the spelling session.
- Both the ITR and the accuracies smoothly converge to 1.49 bits min−1 and 69%, respectively, as the number of stimuli increases. (see table1, GAV1)
- The ITR is 6.3 bits min−1, obtained for just two stimuli, corresponding to an accuracy of 57%, far below 70%.

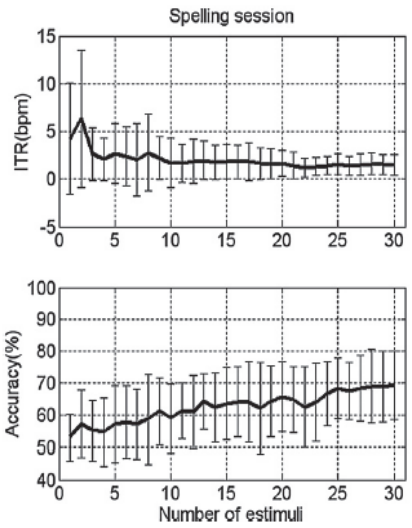


Figure 4(left). Mean performance (thick lines) and standard deviation (vertical bars) across subjects versus incremental number of stimuli for the spelling and speech sessions (left and right respectively).

Results

- Figure 4 (right) shows the averaged performance across subjects achieved during the speech session.
- The ITR–accuracies converge to 1.12–73%(see table 2, GAV1)
- The maximum ITR obtained corresponds to 2.7, which was achieved with an accuracy of 52% and just one stimulus
- However, this accuracy is considerably under 70% and chance level.

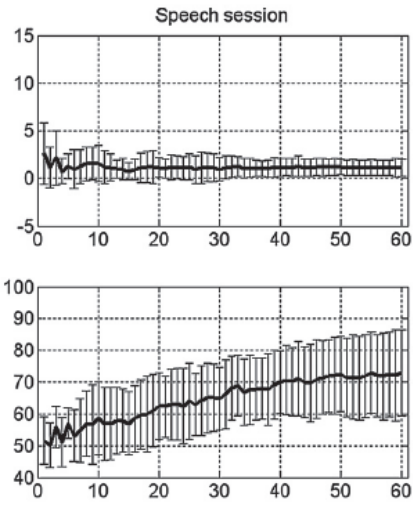


Figure 4(right). Mean performance (thick lines) and standard deviation (vertical bars) across subjects versus incremental number of stimuli for the spelling and speech sessions (left and right respectively).

Conclusions

- This intuitive and easy to use auditory BCI founds a novel approach in the field of BCIs
- This system may provide a clinically useful communication device to those that are in a locked-in state.(can not use visual stimuli)
- From the results, they can state that the use of the dichotic listening paradigm is a feasible approach for auditory BCIs.
- Beyond the performance, the results show that an auditory BCI based on natural speech is a reasonable approach indicated for users unable to use a visual BCI.
- The use of selective attention enhances the usability of the system, with a minimum of training (one run for calibration) and a minimum number of EEG channels (just one).
- In the future, simultaneous feedback will be allowed; thus, they expect an improvement in the performance.

Thank you

Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface

Sitaram et al.

NeuroImage (2007)

Presenter : Evgenii Kim

GIST, Dept. of Information and Communication

Gwangju Institute of Science and Technology

Outline

- Introduction
- Experimental procedure
- Signal acquisition
- Preliminary signal analysis
- Pattern classification
- Support Vector Machine
- Hidden Markov Model
- Result
- Conclusion

Introduction

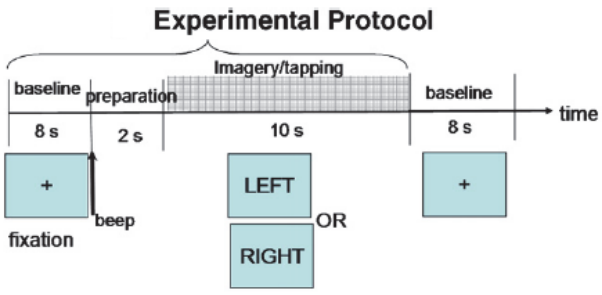
- NIRS is non-invasive optical method that allows us to detect brain activity.
- It possesses a number of advantages:
  - Portability
  - Good spatial resolution
  - Metabolic specificity
  - Low-cost

Disadvantages:
The signal is affected by hair;
Motion artifacts.

- The main goal of the study is to ascertain the feasibility of using NIR for developing a BCI.
- Motor imagery of left/right hand has been used as a paradigm for this work.
- In addition, two pattern recognition technique, SVM and Hidden Markov Model (HMM) were applied to the classification problem.

Experimental procedure

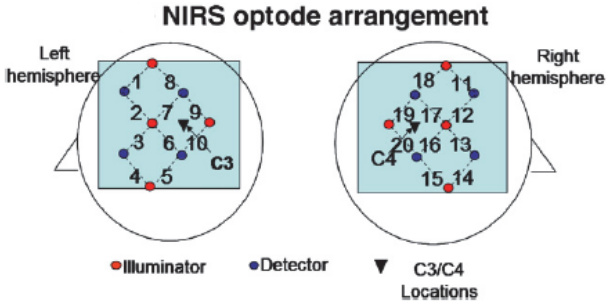
- Five healthy subjects (3 males and 2 females, mean age =30) participated in the study.
- NIRS signals were collected from both overt motor execution (finger tapping) and covert motor imagery



Data for finger tapping and imagery were collected in two separate session. Each tasks was carried out for 80 trials.

Signal acquisition

- The NIRS instrument (OMM-1000 from Shimadzu Co, Japan) was used in this study. The system operated at three different wavelength of 780, 805,and 830 nm, emitting an average power of 3 mW/mm^2
- The data were acquired at a sampling rate of 14 Hz and digitized by the 16-bit analog to digital converter



Preliminary signal analysis

- The reason of preliminary analysis was to observe the responses of HbO2 and Hb at different channels on both hemispheres due to left/right-hand imagery tasks.
- A custom Matlab NIRS data analysis program (HomER) was used
- Preprocessing
  - The raw intensity data from all channels being normalized

$$\text{Norm Intensity}(t) = \text{Intensity}(t)/\text{Mean Intensity}.$$

  - Chebyshev type II filter with cut-off frequency of 0.7 Hz and pass-band attenuation 0.5dB
  - The changed of optical density was calculated for each wavelenght

$$\Delta\text{OD} = -\log(\text{Norm Intensity}(t)).$$

  - PCA filters to remove head movement and to project out systemic physiology
  - Calculate HbO2 and Hb concentration changes from modified Beer-Lamber law.

$$\Delta\text{Hb}X = (e^T e)^{-1} e^T [\Delta\text{OD}].$$

Where e is the molar absorption coefficient for Hb and HbO2

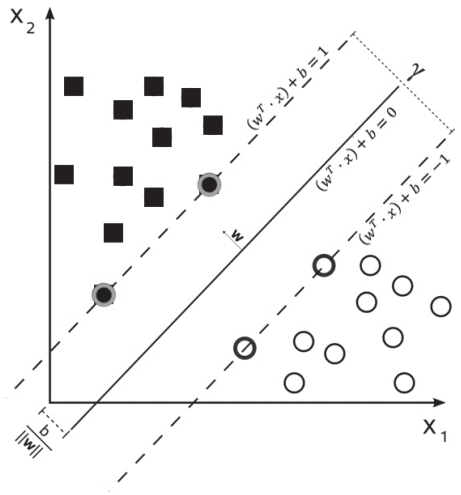
$$\begin{pmatrix} Hb \\ HbO2 \end{pmatrix} = \frac{1}{L} \overline{\begin{pmatrix} \varepsilon_{Hb}^{\lambda 1} & \varepsilon_{HbO2}^{\lambda 1} \\ \varepsilon_{Hb}^{\lambda 2} & \varepsilon_{HbO2}^{\lambda 2} \\ \vdots & \vdots \\ \varepsilon_{Hb}^{\lambda n} & \varepsilon_{HbO2}^{\lambda n} \end{pmatrix}^{-1}} \begin{pmatrix} A^{\lambda 1} \\ A^{\lambda 2} \\ \vdots \\ A^{\lambda n} \end{pmatrix}$$

- Image of brain activity was reconstructed

Pattern classification

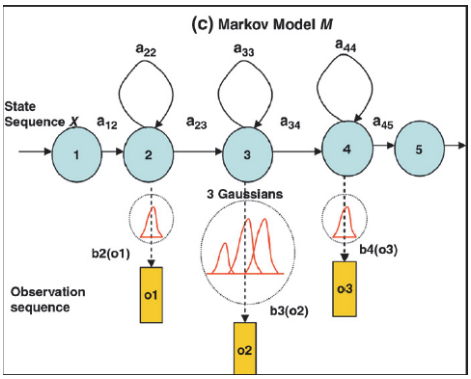
- the raw signals obtained from the signal acquisition process were used for the pattern classification.
- Only Chebyshev type II filter was utilized to remove artifacts from heart beat and high frequency noise
- Time series of amplitude changes of HbO2 and Hb in period 2-10 s after stimulation for the motor task were extracted and fed to the pattern classification system

SVM



y_t represents the type of task (left/right hand)
$\mathbf{X}_t$ represents the concentration values HbO2 and Hb from all the number of channels

HMM



At each time t that a state j is entered, an observation vector O is generated from the probability density $b_j(\mathbf{O}_j)$. Transition from state i to state j is also probabilistic and is governed by the discrete probability a_{ij}. The joint probability O generated by model moving through the state sequence X.

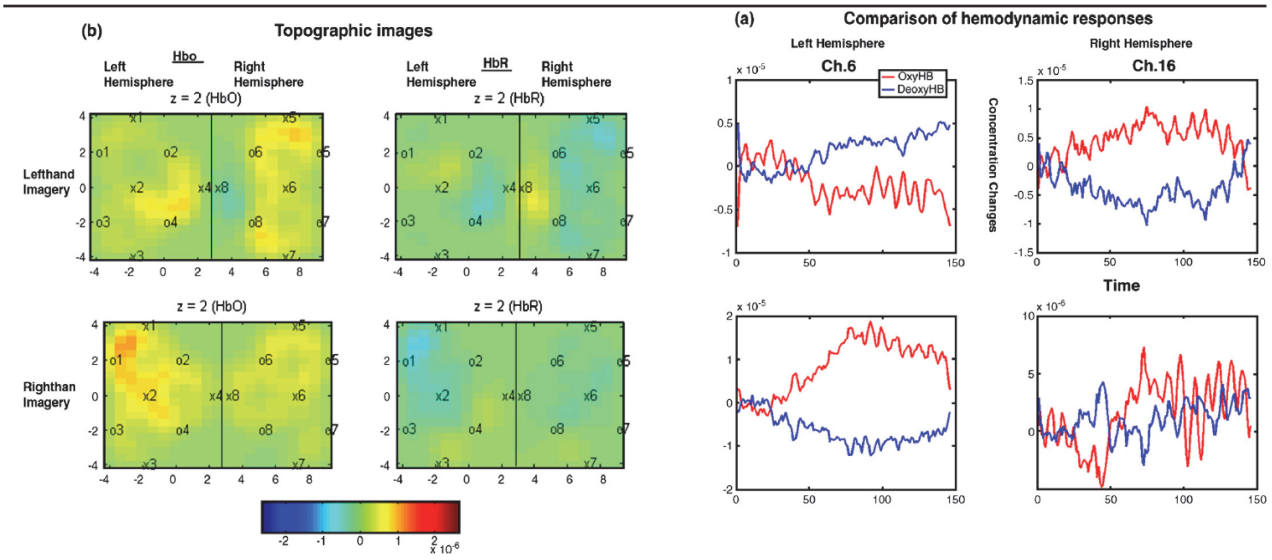
$$P(O, X|M) = \sum_x a_{x(o)x(1)} \prod_{t=1}^{T} b_{x(t)}(o_t) a_{x(t)x(t+1)}$$

X is unknown

$\{a_{ij}\}$ and $\{b_j(O_t)\}$ are determined by an estimation procedure. To determine Baum-Welch re-estimation procedure was used.

Result

Accuracy of Support Vector Machine (SVM) and Hidden Markov Model (HMM) classification of finger tapping and motor imagery tasks for 5 healthy volunteers



Conclusion

- The NIRS has been demonstrated as reliable technique for BCI application.

Thank you

INFONET Seminar Application Group 2013/10/22

A measurement-domain adaptive beamforming approach for ultrasound instrument based on distributed compressive sensing: Initial development

Quiong Zhang.
Ultrasounics, Elsevier
June 2012
Presenter Pavel Ni

Gwangju Institute of Science and Technology

Contents

Introduction

- Sensor arrays are used in ultrasound imaging, broad band wave transmitted and backscat tered signal is acquired
- The received signal is amplified LNA (Low Noise Amplifier) and VGA (Variable Gain Amplifier) filtered by AAF (Anti Alias Filter) and sampled by ADC (Analog Digital Con verter).
- Conventional Ultraosund systems use DAS beamformer (Delay And Sum) to process high-speed sampled data Fig. 1
- Conventional Beamformer delays received signals by array elements appropriate to their distances from the main target of imaging, weights them with predetermined coefficients and then sums the weighted signal to reconstruct the echo signal originating from main target.
- In conventional beamformer weights are independent of the array data therefore reducing the overall quality of the image
- In Adaptive Beamformer(ABF) delayed sensor signals are fed to the adaptive processor, which continuously updates the appodization weights for the specific sensor signal that provides significant increase in lateral resolution compared with DAS

Introduction

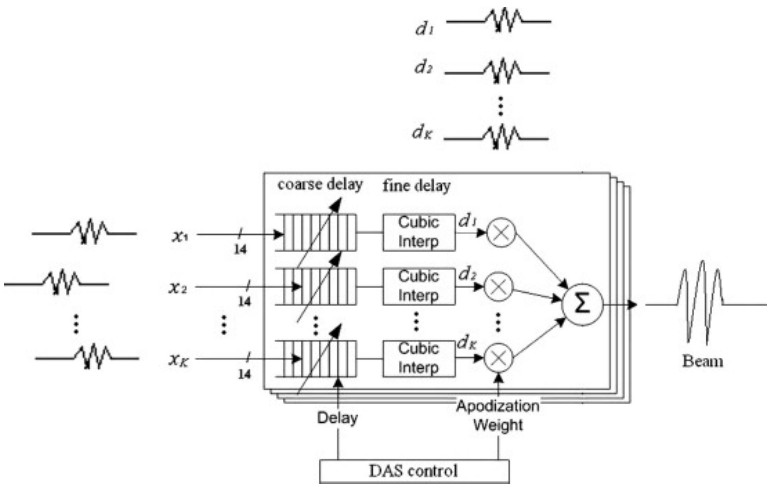


Fig. 1. The process diagram of traditional delay-and-sum (DAS) beamformer used in ultrasound imaging instrument.

- Author proposed framework for ultrasound sensor array data acquirement based on DCS and sample data with sub-Nyquist-rate
- High resolution MABF approach to reconstruct ultrasound images adaptively without recovering signal

Basis theory of distributed compressed sensing

Distributed compressed sensing (DCS) enables new distributing codding algorithms for multi-Signal ensembles that exploit both intra- and inter-signal correlation structures.

- In DCS scenario, a number of sensors measure signals that are each individually sparse in some basis and also correlated from sensor to sensor. Think about it as a multiple sensors acquire the same signal but with phase shift and attenuation caused by signal propagation (Joint Sparsity Model-2)
- Then each sensor independently encodes its received signal by projection it onto another, incoherent basis and then transmits just few essential coefficients.
- Under the right conditions decoder at the collecting point can jointly reconstruct all of the signals precisely.

Basis theory of distributed compressed sensing

In JSM-2 system all K sensor-array signals share a common support set but with different coefficients. Then signal on sensor i is:

$x_i = \Psi\theta_i$

Measurement vector y_i is computed as

$y_i = \Phi_i x_i.$

Where $\Phi_i \in \Re^{MixN}$ with $M_i << N$ and $y_i \in \Re^{M_i}$.

M_i is the number of measurements in the i-th sensor.

All the sensor array signals and measurements can be represented respectively as:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix} = \begin{bmatrix} \Psi & 0 & \cdots & 0 \\ 0 & \Psi & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Psi \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \\ \theta_K \end{bmatrix} = G\theta. \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_K \end{bmatrix} = \Phi x = \Phi G\theta = \Omega\theta \qquad (1)$$

The recovering of x can be achieved by solving a L1 minimization problem:

$\theta = \arg\min(\|\theta\|_1)$ s.t. $y = \Omega\theta$

In ultrasound, each sensor acquires a replica of the same frequency-sparse signal but with phase shift and attenuation caused by signal propagation

The measurement-domain adaptive beamforming (MABF) based on DCS

- The amount of acquired data is decreased by implementing DCS at the front-end.
- High-resolution ultrasound images can be obtained from few measurements of DCS using MABF directly without recovering raw sensor signals with complex convex optimization algorithm.

The author used a single unfocused transmission wave. K elements of transducer are exited by the identical wave at the same time, and entire region of interest is covered in a single transmission. The same transducer array is used to receive corresponding backpropagated acoustic echo signals.

The time signal received by the ith element is denoted by $s_i(t,\vec{\rho})$, assuming that there is only single scatterer $\vec{\rho}$.

$$y_i = \Phi_i x_i = \Phi_i \cdot \int_{ROI} s_i(t,\vec{\rho}) \cdot f(\vec{\rho})\vec{\rho} \qquad (2)$$

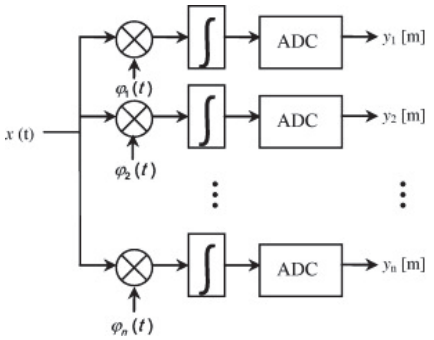


Fig. 2. The possible distribute compressive sensing implementation of one sensor signal at the ultrasound imaging instrument receiver. $\Phi=[\varphi_1,\varphi_2,\ldots,\varphi_n]$.

The measurement-domain adaptive beamforming (MABF) based on DCS

If Φ_i does not map two distinct S-sparse signals to the same measurement vector then it is possible to reconstruct image directly from measurements without signal reconstruction. Restricted Isometry Property (RIP) should hold for all S-sparse vectors x_i

To solve (2) the ROI must be discretized as shown in Fig. 3. The ROI lies in the product space $[x_s, x_e] \times [y_s, y_e]$, where (x_s, y_s) and (x_e, y_e) denote initial and final positions of the ROI to be imaged

The received measurements in the ith element can be expressed

The measurement-domain adaptive beamforming (MABF) based on DCS

$$y_i = \Phi_i \cdot \sum_{j=1}^{LxP}\left[s_i(t,\vec{\rho}_j) f(\vec{\rho}_j)\right] = \sum_{j=1}^{LxP} d_{ji} \cdot f(\vec{\rho}_j) = D_i f \quad (3)$$

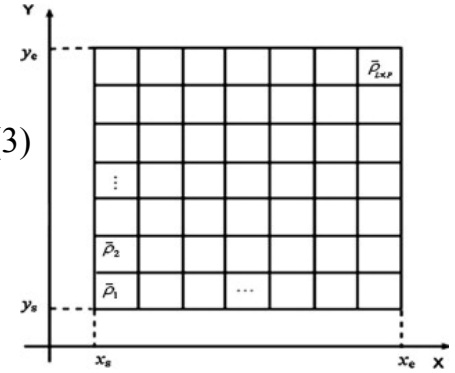


Fig. 3. RIO discretization. The ROI lies in the product space $[x_s, x_e] \times [y_s, y_e]$ and is discretized into in L by P scatterers $\vec{\rho}_1, \vec{\rho}_2, \ldots \vec{\rho}_{LxP}$

Consider all the elements, the measurements of actual backscattering signals can be represented as

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_K \end{bmatrix} = \begin{bmatrix} D_1 f \\ D_2 f \\ \vdots \\ D_K f \end{bmatrix} = V_{DCS} f \quad (4)$$

Where V_{DCS} is a dictionary matrix of dimensions MKxLP, $Di = [d_{1i}, d_{2i}, \ldots d_{(LxP)i}]$ and f is a target vector.

The measurement-domain adaptive beamforming (MABF) based on DCS

The maximum likelihood estimate of f is given by

$$\hat{f} = \max_f p(y, f) \text{ s.t. } y = V_{DCS} \cdot f \quad (5)$$

Where is the pdf. Then (5) becomes MAP optimization problem

$$\hat{f} = \max_f p(f \mid y) = \max_f [p(y \mid f) p(f)] \quad (6)$$

To find the MAP estimate of target vector f, we must maximize the priori PDF $p(f)$.

Results

To show efficiency of reduce numb er of samples and image reconstruc tion

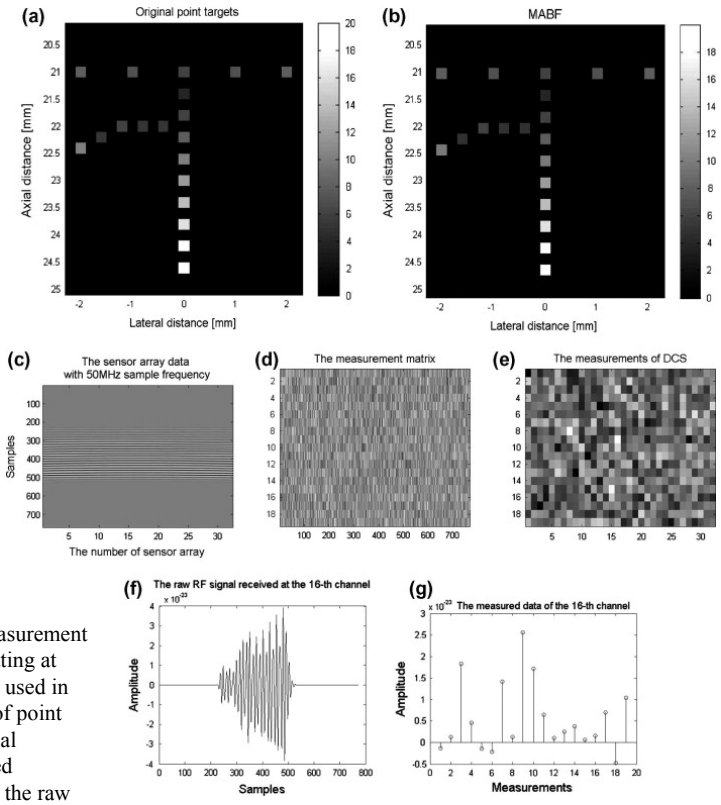


Fig. 4. The example of sensor array data acquirement, measurement and image reconstruction. A 32 element linear array operating at 5 MHz with 80% bandwidth in plane wave emission were used in simulation. (a) the positions and backscattering intensity of point targets, (b) the result of the proposed method, (c) the spatial response, (d) the linear measurement matrix, (e) distributed compressive measurements of the sensor array signals, (f) the raw RF signal received at the 16th channel, (g) the measured data corresponding to the 16th channel.

Results

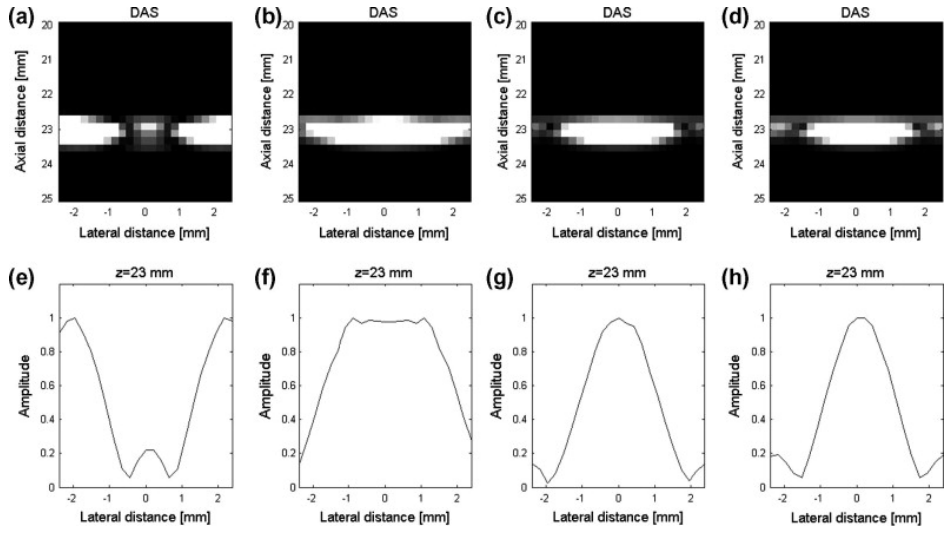


Fig. 5. Simulated lateral resolution of DAS for a 32 element linear array operating at 5 MHz with 80% bandwidth in plane wave emission. (a) shows the case, where the point targets are separated by 5 mm at depth of 23 mm; (b)–(d) correspond to point targets separations of 2, 1 and 0.2 mm, respectively. (e)–(h) show the normalized lateral variation of (a)–(d) at depth of 23 mm.

Results

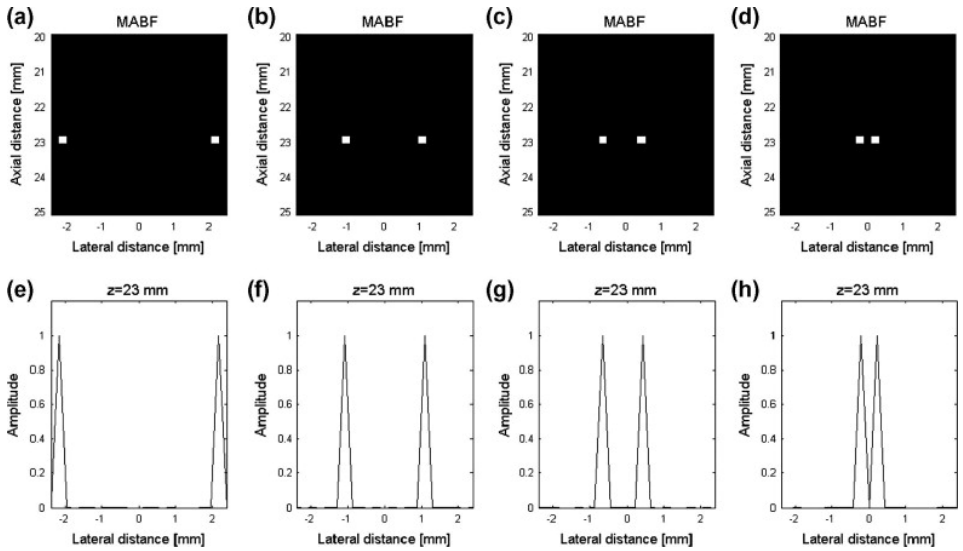
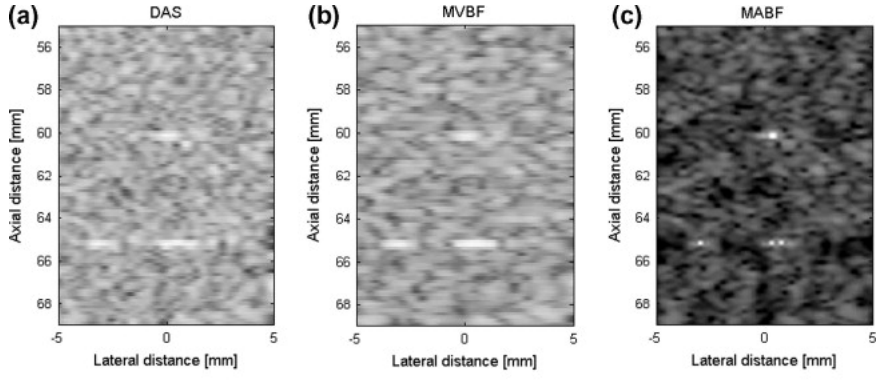


Fig. 6. Simulated lateral resolution of MABF for a 32 element linear array operating at 5 MHz with 80% bandwidth in plane wave emission. The hypothetical targets separated 0.2 mm in range and 0.2 mm in azimuth. (a) shows the case, where the point targets are separated by 4.5 mm at depth of 23 mm; (b)–(d) correspond to point targets separations of 2, 1 and 0.2 mm, respectively. (e)–(h) show the normalized lateral variation of (a)–(d) at depth of 23 mm.

Results



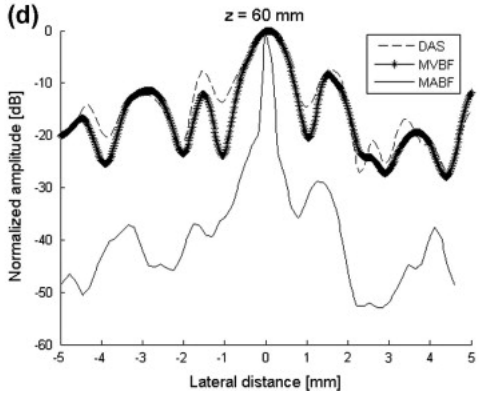


Fig. 7. Beamformed responses of four point targets in a speckle pattern with a 5-MHz, 80% bandwidth, 32-element, half-wavelength spacing linear array in the PW emission. The four point targets are located at $(x, z) = \{(0, 60), (-3.4, 65), (0, 65), (0.5, 65)\}$ mm. Both DAS and MVBF are implemented on the raw data, while the proposed MABF is implemented on the measurements. (a) DAS, (b) MVBF and (c) MABF. (d) shows the normalized lateral variation of (a)–(c) at depth of 60 mm. All images are shown with a dynamic range of 60 dB.

Conclusion

Using DCS few measurements are acquired by projecting the sensor array signals onto the s elected basis.
High resolution image is reconstructed by using MABF approach to the few measurements directly without recover the raw signals.

Further work
Computational time is not suitable for real-time applications

Q&A

A novel BCI based on ERP components sensitive to configural processing of human faces.

Yu Zhang et al. (Andrzej Cichocki*)

Journal of Neural Engineering (2012)

Presenter : SeungChan Lee

GIST, Dept. of Information and Communication, INFONET Lab.

Gwangju Institute of Science and Technology

Background

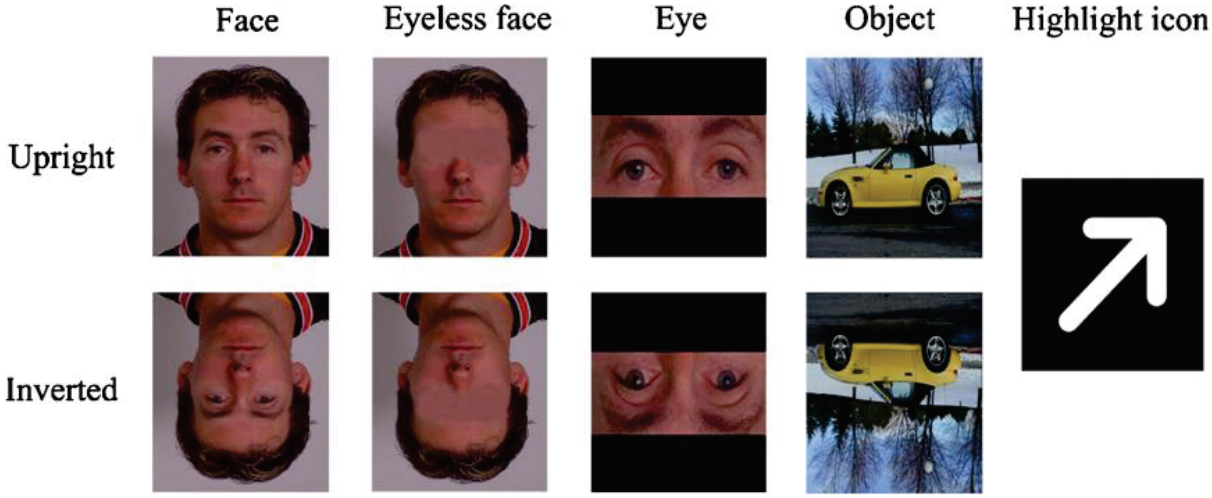
- Face-sensitive event related potentials(ERPs)
  - Related ERPs
    - N170 : a large negative component peaking at the lateral occipito-temporal sites between 140 ~ 200ms
    - VPP(Vertex Positive Potential) : a large positive component at the fronto-central sites with a similar latency to the N170
    - P1, N250
  - Previous study
    - Oddball paradigm-based BCI with stimuli of natural faces
    - Online accuracy reaches over 90% with two trials (better performance using facial images instead of using intensified icon stimuli)
    - The prominent features derived from the facial images at visual cortex, which may be associated with the cognitive components reflecting face perception.

Introduction

- Motivation
  - Face perception rely more on configural information rather than other visual object perception.
  - The inversion of a face can disrupt the configural face information, thereby making the face processing slower and more difficult.
  - The two components N170 and VPP are believed to reflect the configural processing of the face, their amplitudes and latencies can be modulated by the inversion of the face.
  - Could the signal modulation caused by the loss of configural face information be applied to the BCI using stimuli of facial images and improve the system performance?

Methods

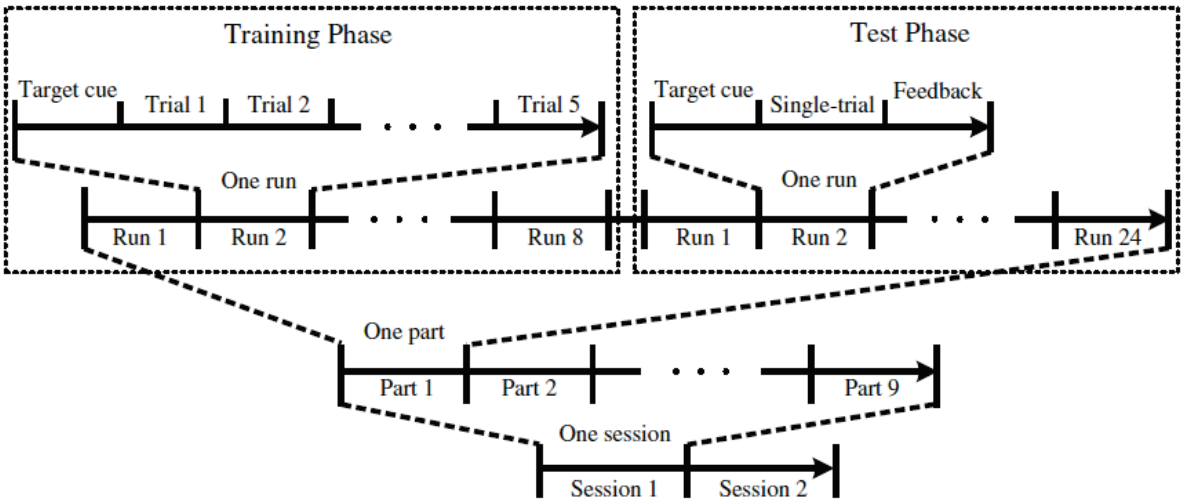
- Subjects
  - 7 healthy right-handed volunteers (aged from 24 to 49, all males)
- Stimuli
  - 9 types of stimuli on ERP components(N170, VPP, and P300)



  - 4 natural human faces(2 females) : face-related stimuli
  - 4 objects(car, ship, bicycle and house) : object stimuli

Methods

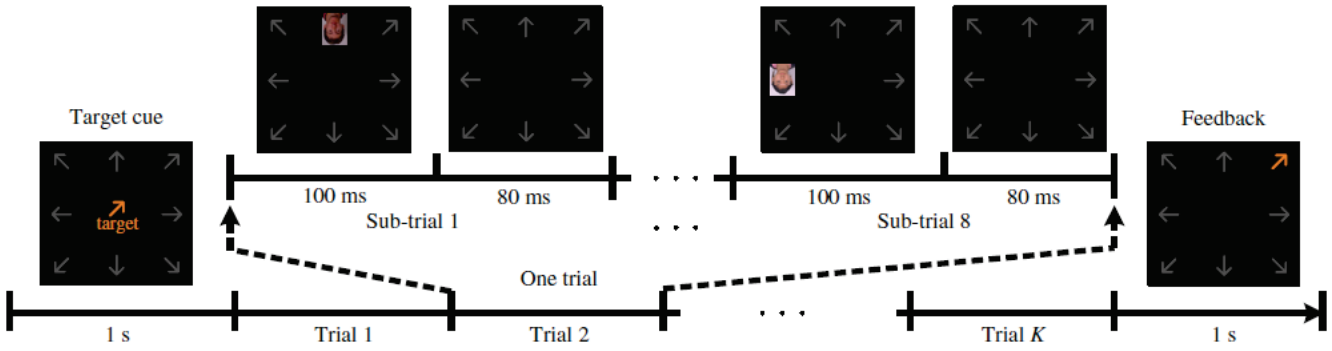
- Paradigm
  - Each subject completed two experimental sessions on two separate days. (interval : less than three days)
  - Each part being tested with same stimulus type.



  - Total 48 direction commands were implemented for each subject in the online test phases of the two sessions

Methods

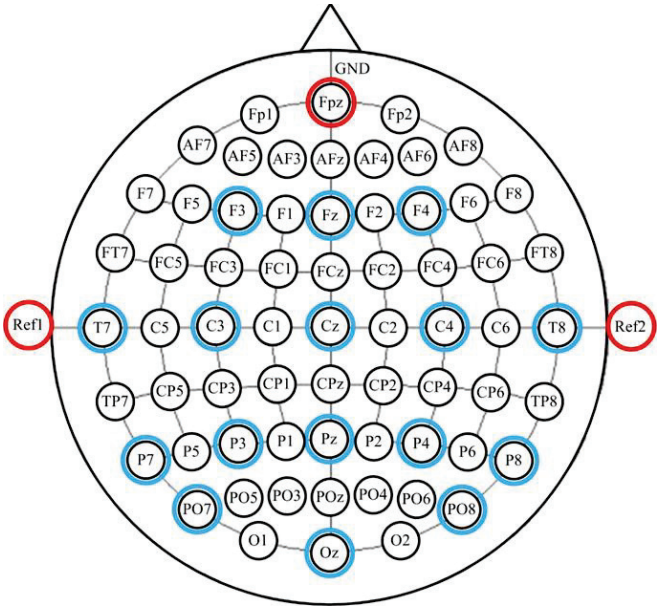
- Paradigm
  - The timing of one run



  - Training phase : K=5, each run consisted of 40 flash sub-trials (5 targets and 35 non-targets) with no feedback
  - Online test phase : k=1(single trial), feedback was provided.

Methods

- EEG acquisition
  - 256 Hz sampling rate with the g.USBamp amplifier (high-pass and low-pass filters 0.1Hz and 30 Hz; a notch filter 50 Hz)
  - 16 electrodes were used (F3, Fz, F4, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, PO7, PO8, Oz, two ear references, and one ground on the Fpz)



Methods

- Feature extraction
  - 700 ms data segment after baseline corrected (100 ms pre-stimulus interval was extracted)
  - Total 320 such data segments consisting of 40 targets and 280 non-targets were derived from each part during the training phase.
  - Each data segment was downsampled to 21 Hz after 12-point moving average.
  - A spatiotemporal feature vector with dimension of 240 (i.e. 16 channels × 15 sampling points)
  - 320 feature vectors were collected for each type of stimulus.
- Classification
  - Linear discriminant analysis(LDA) was used.
  - Procedure
    - Eight spatiotemporal feature vectors were extracted during the single trial.
    - Calculate their posterior probabilities belonging to the target class.
    - stimulus direction with the maximal posterior probability was detected and presented to the subject as feedback.
  - Classification accuracy was averaged over the two sessions.

Methods

- Evaluation
  - Information transfer rate(ITR)

$$ITR = M\left\{\log_2 N + P\log_2 P + (1-P)\log_2\left(\frac{1-P}{N-1}\right)\right\} bits/\min$$

    - N possible choices in which each choice is equally probable to be selected by the user.
    - The probability (P) that the desired choice will indeed be selected remains invariant.
    - Each error choice has the same probability of selection.
    - M denotes the number of commands per minute.
  - One-way analysis of variance(ANOVA)
    - ANOVA is a collection of statistical models used to analyze the differences between group means and their associated procedures (such as "variation" among and between groups)
    - ANOVAs are useful in comparing (testing) three or more means (groups or variables) for statistical significance.

Results

- Online accuracy and ITR

Stimulus	Performance	Subject S1	S2	S3	S4	S5	S6	S7	Average
Upright face	Acc	83.3	81.3	75.0	81.3	50.0	83.3	85.4	77.1 ± 12.4
	ITR	32.8	31.0	25.9	31.0	10.4	32.8	34.7	28.4 ± 8.39
Inverted face	Acc	**93.8**	**87.5**	**85.4**	**89.6**	**70.8**	**95.8**	**97.9**	**88.7** ± 9.08
	ITR	**43.4**	**36.7**	**34.7**	**38.8**	**22.8**	**45.9**	**48.7**	**38.7** ± 8.63
Upright eyeless face	Acc	85.4	79.2	81.3	79.2	58.3	87.5	93.8	80.7 ± 11.2
	ITR	34.7	29.3	31.0	29.3	14.8	36.7	43.4	31.3 ± 8.83
Inverted eyeless face	Acc	89.6	77.1	**85.4**	79.2	54.2	**95.8**	95.8	82.4 ± 14.5
	ITR	38.8	27.6	**34.7**	29.3	12.5	**45.9**	45.9	33.5 ± 11.8
Upright eye	Acc	91.7	70.8	75.0	72.9	45.8	87.5	79.2	74.7 ± 14.9
	ITR	41.1	22.8	25.9	24.4	8.43	36.7	29.3	26.9 ± 10.6
Inverted eye	Acc	89.6	68.8	77.1	70.8	41.7	81.3	87.5	73.8 ± 16.2
	ITR	38.8	21.4	27.6	22.8	6.69	31.0	36.7	26.4 ± 10.9
Upright object	Acc	70.8	75.0	64.6	66.7	37.5	77.1	81.3	67.6 ± 14.5
	ITR	22.8	25.9	18.6	20.0	5.08	27.6	31.0	21.6 ± 8.45
Inverted object	Acc	77.1	66.7	75.0	58.3	39.6	70.8	83.3	67.3 ± 14.6
	ITR	27.6	20.0	25.9	14.8	5.86	22.8	32.8	21.4 ± 8.92
Highlight icon	Acc	48.3	68.8	41.7	43.8	33.3	47.9	45.8	47.1 ± 10.8
	ITR	9.58	21.4	6.69	7.56	3.65	9.39	8.43	9.53 ± 5.61

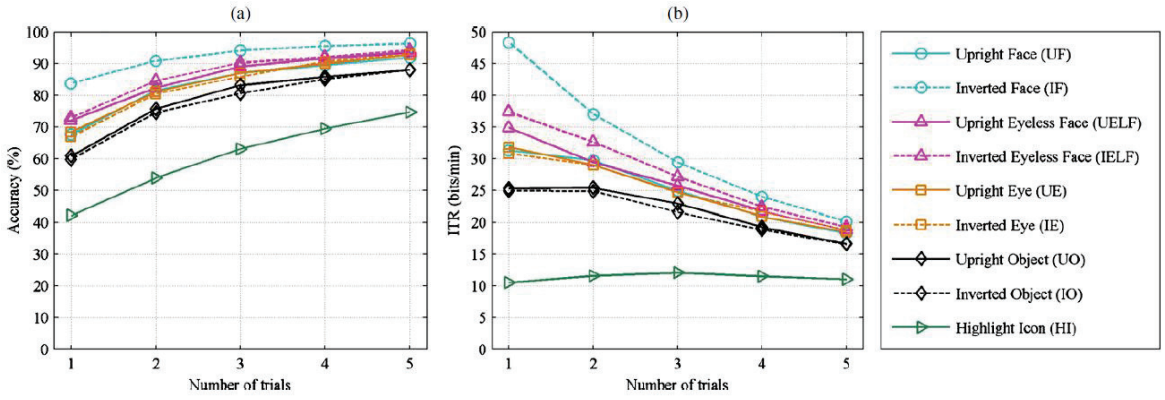
  - The best performance with accuracy of 88.7% and ITR of 38.7 bits min−1 was yielded by the inverted face.

Results

- Online accuracy and ITR
  - Accuracies
    - Compared with the highlight icon (accuracy of 47.1%), other stimuli achieved significantly higher accuracies.
    - While the accuracy had no significant difference between upright and inverted for all face-related stimuli and objects, the difference was marginally significant for the inverted face in contrast to the upright face.
    - Comparing the face-related stimuli with the object, only the inverted face generated significantly higher accuracy than that of the object.
  - ITRs
    - Both the face-related stimuli and the object achieved significantly higher ITRs than the ITR 9.53 bits min−1 of the highlight icon.
    - The inverted face yielded significantly higher ITR than that of the upright face while there was no significant difference between upright and inverted for the eyeless face, eye and object.
    - The inverted face also significantly improved the ITR in comparison to the object.

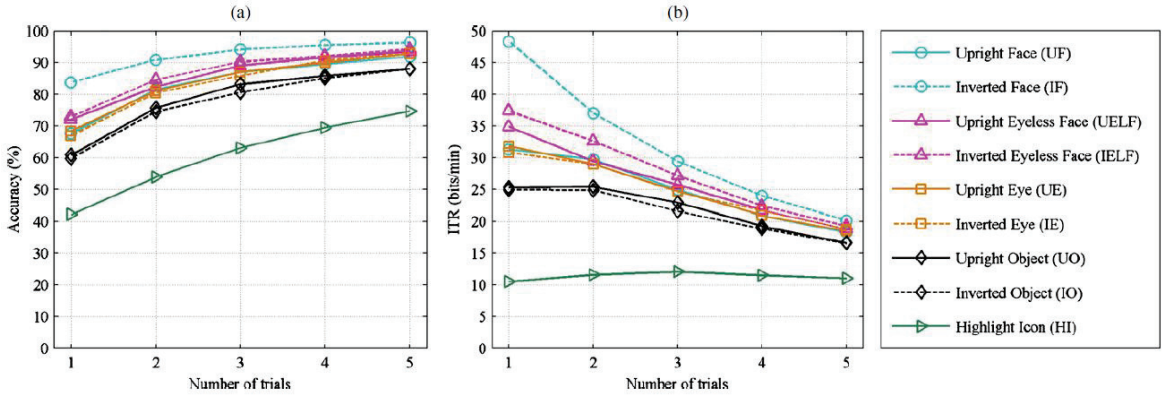
Results

- Offline analysis
  - Why performance improved for the inverted face?
  - Methods
    - For each type of stimulus, 8 runs were randomly selected from the 16 runs (5 targets and 35 non-targets in each run) of the two experimental sessions for the classifier training.
    - The remaining 8 runs were used as test data.
    - Such procedure was repeated 100 times and the average classification accuracy and ITR were then calculated.



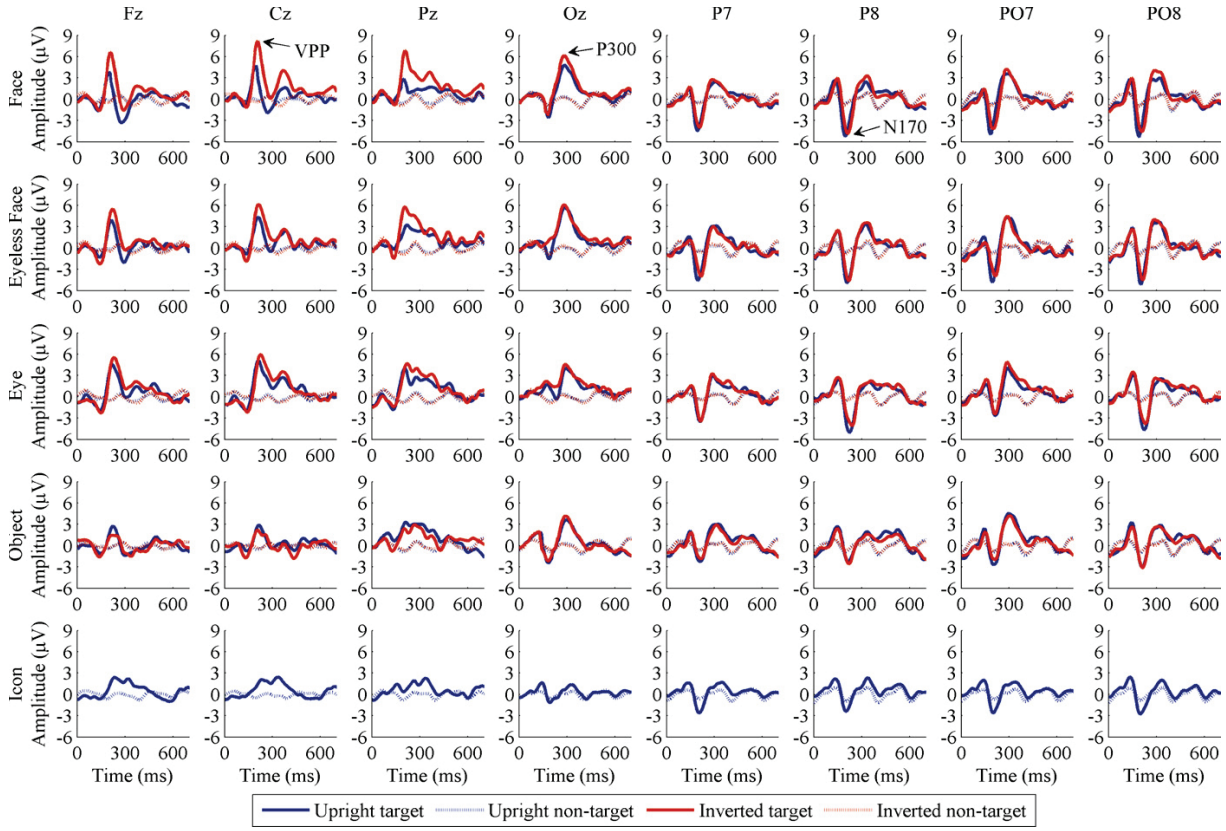
Results

- Offline analysis
  - Comments
    - The inverted face yielded higher accuracy and ITR than those of the other stimuli across various trials.
    - The face-related stimuli obtained a performance exceeding that of the object, while both of them performed better than the highlight icon.
    - There was no big difference between upright and inverted for the eyeless face, eye and object, whereas the inverted face was noticeably better than the upright face.



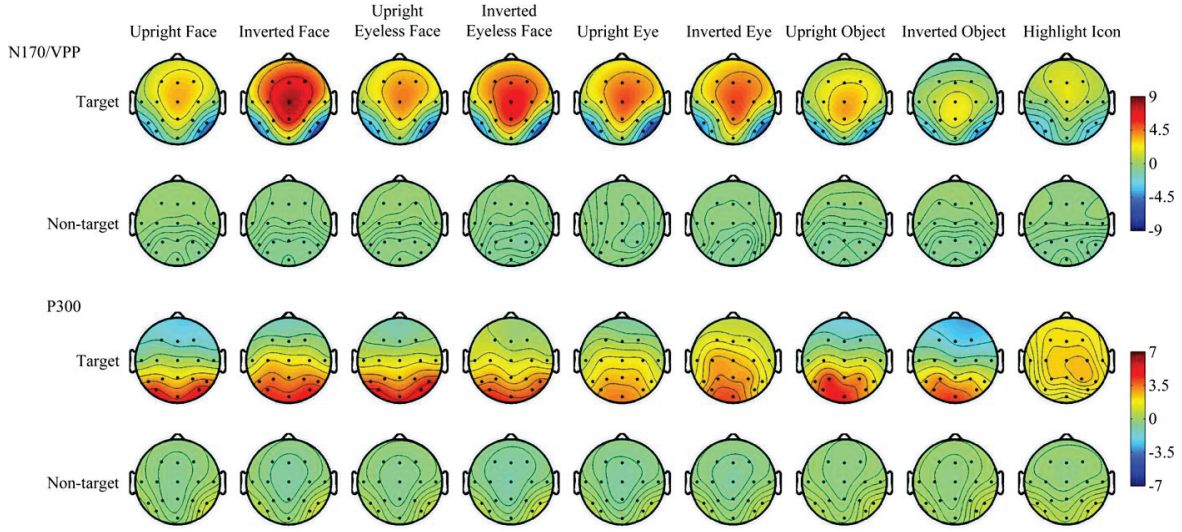
Results

- ERP analysis



Results

- ERP analysis



	N170		VPP		P300	
Paradigm	Amplitude	Latency	Amplitude	Latency	Amplitude	Latency
Upright face	−6.36 ± 2.64	200.9 ± 16.4	5.17 ± 1.10	199.2 ± 13.2	5.40 ± 1.15	274.1 ± 59.8
Inverted face	−6.18 ± 2.34	217.1 ± 11.7	8.37 ± 1.71	214.9 ± 7.81	6.68 ± 1.17	279.6 ± 44.6
Upright eyeless face	−6.19 ± 2.39	204.8 ± 12.9	5.36 ± 1.06	206.3 ± 12.7	6.44 ± 1.02	284.1 ± 30.1
Inverted eyeless face	−5.83 ± 2.13	218.8 ± 7.80	7.12 ± 1.94	218.3 ± 8.80	6.58 ± 1.24	281.6 ± 37.2
Upright eye	−5.60 ± 1.87	207.6 ± 9.30	6.22 ± 2.11	208.7 ± 9.10	5.15 ± 1.36	282.3 ± 45.8
Inverted eye	−5.22 ± 1.18	217.1 ± 10.7	7.06 ± 2.40	219.3 ± 8.10	5.20 ± 1.31	276.3 ± 25.2
Upright object	−3.44 ± 1.61	208.7 ± 11.0	3.43 ± 1.47	201.5 ± 26.1	5.19 ± 1.13	290.7 ± 35.9
Inverted object	−3.54 ± 1.24	209.8 ± 10.7	3.15 ± 1.79	199.2 ± 31.0	5.16 ± 1.03	294.0 ± 24.4
Highlight icon	−3.25 ± 1.79	199.8 ± 11.8	2.79 ± 1.86	205.4 ± 32.3	3.57 ± 1.44	301.5 ± 37.4

Results

- ERP analysis
  - N170
    - Larger N170 amplitudes evoked by the face-related stimuli than by the highlight icon.
    - No significant difference was found among the face-related stimuli and between the object (both upright and inverted) and highlight icon.
    - A longer N170 latency was observed for the inverted than the upright.
  - VPP
    - A larger VPP amplitudes evoked by the face-related stimuli than by the highlight icon and by the object.
    - The inverted face elicited significantly larger VPP than the upright face.
    - A longer VPP latency was observed for the inverted than the upright.
  - P300
    - A larger P300 amplitudes evoked by the face-related stimuli and the object than by the highlight icon, especially at the parietal-occipital and occipital sites.
    - The inverted face yielded higher P300 amplitude than that of the upright face.
    - the P300 amplitude evoked by the eyeless face (both upright and inverted) was higher than by the upright face.
    - the P300 amplitudes derived from the inverted face and eyeless face were significantly higher than that of the object.

Results

- Discriminative feature analysis
  - r^2-value(squared pointwise biserial correlation coefficients)
    - Pointwise biserial correlation coefficient
      - Definition

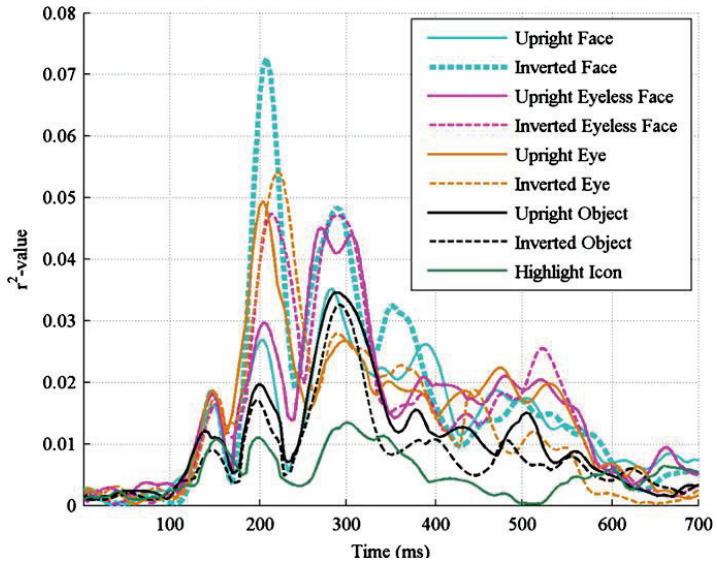
$$r(x) = \frac{\sqrt{N_1 N_2}}{N_1 + N_2} \frac{\text{mean}\{x_i | y_i = 1\} - \text{mean}\{x_i | y_i = 2\}}{\text{std}\{x_i | y_i = 1, 2\}}$$

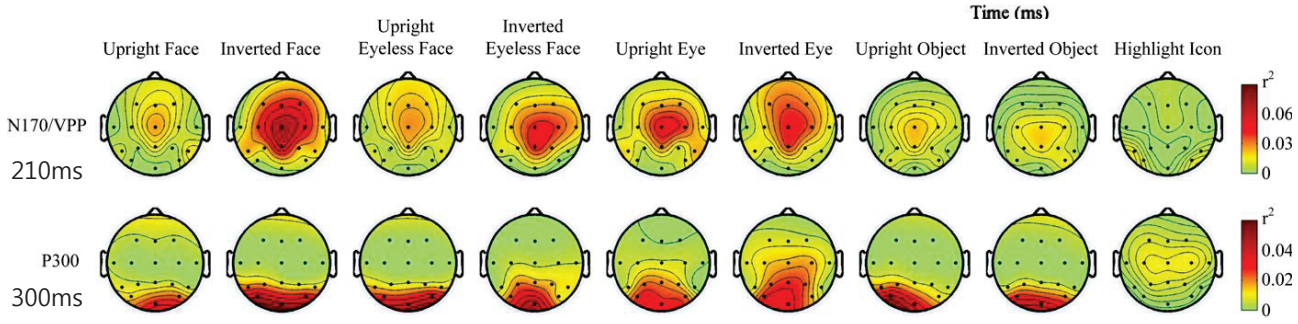
Where N1 and N2 are the numbers of variables belonging to class 1 (target) and class 2 (non-target), xi and yi are the value and class label of the ith variable.

    - The r2-value is equal to the squared of r(x).
    - Larger r2-value indicates higher separability of distributions.

Results

- Discriminative feature analysis
  - Temporal and spatial distributions of the most discriminative information for the nine stimuli





Results

- Discriminative feature analysis
  - Comments
    - Almost all of the face-related stimuli and the object yielded more discriminative features than the highlight icon from 200 to 500 ms after stimulus onset.
    - The most outstanding components in the features were found around 200 and 300ms, which just correspond to N170/VPP and P300.
    - The discriminative features around 200 ms for the face-related stimuli and the object were mainly located at the fronto-central sites(Cz)
    - the P300 distributions for the face-related stimuli and the object were mainly located at the parietal-occipital sites, compared with the centro-parietal distribution of P300s elicited by the highlight icon

Discussion

- Advantages of facial images based BCI
  - A high luminance contrast is usually required to elicit a prominent visual evoked potential for the visual stimuli, and this may cause visual fatigue and discomfort for the user.
  - The facial images are more vivid than icons, letters or symbols, they may resist fatigue and discomfort to improve the visual attention for subjects.
  - Loss of configural information makes face perception more difficult and associated with higher cognitive functions. This encourages subjects to focus attention on the target more actively.
- Performance
  - Both the face related stimuli and the object yielded significantly higher accuracies and ITRs than that of the highlight icon.
    - This implies that stimuli with higher cognitive task requirement, such as face and object perception, are more effective than the intensified stimuli of dull icons for the P300-based BCI system.
  - The ITR derived from the inverted face was significantly higher than that of the upright face.
    - This suggests that the loss of configural face information assists in improving the performance of the BCI system.

Conclusion

- They proposed a novel BCI system using multi-component ERPs sensitive to configural processing of human face with an oddball paradigm.
- The performance of the proposed BCI is significantly improved in comparison to the conventional P300-based BCI with stimuli of intensification pattern.
  - The online performance of classification accuracy 88.7% and ITR of 38.7 bits min−1 obtained by the LDA classification using only single trial without any optimization of algorithm for feature extraction.