Computational Neuroscience

Tactile and bone-conduction auditory brain computer interface for vision and hearing impaired users

Tomasz M. Rutkowski a, b, *, Hiromu Mori a

a Life Science Center of TARA, University of Tsukuba, Japan
b RIKEN Brain Science Institute, Japan

HIGHLIGHTS

• Novel multimodal tactile and bone-conduction BCI paradigm.
• EEG signal processing methods leading to improved single-trial classification by applying synchrosqueezing transform based frequency domain filtering.
• Application of logistic regression based classifier for improved single trial BCI results.

ARTICLE INFO

Article history:
Received 26 January 2014
Received in revised form 7 April 2014
Accepted 10 April 2014
Available online 21 April 2014

Keywords:
EEG
BCI
Somatosensory evoked potential
P300
Data-driven filtering

ABSTRACT

Background: The paper presents a report on the recently developed BCI alternative for users suffering from impaired vision (lack of focus or eye-movements) or from the so-called “ear-blocking-syndrome” (limited hearing). We report on our recent studies of the extents to which vibrotactile stimuli delivered to the head of a user can serve as a platform for a brain computer interface (BCI) paradigm.

New method: In the proposed tactile and bone-conduction auditory BCI novel multiple head positions are used to evoke combined somatosensory and auditory (via the bone conduction effect) P300 brain responses, in order to define a multimodal tactile and bone-conduction auditory brain computer interface (tbcBCI). In order to further remove EEG interferences and to improve P300 response classification synchrosqueezing transform (SST) is applied. SST outperforms the classical time–frequency analysis methods of the non-linear and non-stationary signals such as EEG. The proposed method is also computationally more effective comparing to the empirical mode decomposition. The SST filtering allows for online EEG preprocessing application which is essential in the case of BCI.

Results: Experimental results with healthy BCI-naive users performing online tbcBCI validate the paradigm, while the feasibility of the concept is illuminated through information transfer rate case studies.

Comparison with existing method(s): We present a comparison of the proposed SST-based preprocessing method, combined with a logistic regression (LR) classifier, together with classical preprocessing and LDA-based classification BCI techniques.

Conclusions: The proposed tbcBCI paradigm together with data-driven preprocessing methods are a step forward in robust BCI applications research.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The state of the art BCIs rely mostly on mental, visual and motor imagery paradigms, which require users to have healthy vision and often to participate in a long training. Recently alternative solutions have been proposed to utilize spatial, auditory (Rutkowski et al., 2009; Halder et al., 2010; Schreuder et al., 2010; Lelievre and Rutkowski, 2013) or tactile (somatosensory) (Muller-Putz et al., 2006; van der Waal et al., 2012; Mori et al., 2013a,b,c) modalities in order to enhance brain–computer interface comfort or to boost the information transfer rate (ITR) (Schreuder et al., 2010) achieved by users. The concept described in this paper of utilizing the brain’s somatosensory (tactile) modality opens up the attractive possibility of targeting the tactile sensory domain, which does not rely on visual stimuli to elicit evoked potentials during visual computer applications or operation of
robotic interfaces (mental speller, prosthetic arm, vehicular robot, smart house appliance, etc.) or visual computer applications. The first successful trial to utilize steady-state somatosensory responses (SSSR) to create a BCI (Müller-Putz et al., 2006) targeted a very low stimulus frequency in a range of 20–31 Hz to elucidate the users’ steady-state activity, which was then used to create BCI commands. A very recent report (van der Waal et al., 2012) proposed using a Braille stimulator with a 100 ms long static push stimulus delivered to six fingers to evoke a somatosensory response related P300. Very encouraging results were obtained with 7.8 bit/min on average and 27 bit/min for the best user. Here we propose to combine the two above-mentioned modalities in the tbcABI paradigm, which relies on P300 response evoked by the audio and tactile stimuli delivered simultaneously via the vibrotactile exciters attached to the head positions (see Fig. 1), thus benefiting from the bone-conduction effect for audio. This offers a viable alternative for individuals lacking somatosensory responses from the fingers or torso/chest body locations, which prevents them from utilizing tactile BCI solutions as proposed in Mori et al. (2013a,b), Severens et al. (2013), or for people suffering from the ear blocking syndrome (a middle ear effusion/negative pressure), which is a common condition of the amyotrophic lateral sclerosis (ALS) and especially in the late disease stages called locked in syndrome (LIS) (Gelinas, 2007; Kaufmann et al., 2013).

Our research hypothesis is summarized as follows. In the tactile and bone-conduction BCI, the use of spatially distributed bimodal (tactile and auditory) stimuli resolves the problem of bone-conducted spatial auditory localization problem (MacDonald et al., 2006) and it should lead to an increase in accuracy compared with classic tactile cases (Brouwer and Van Erp, 2010; Ortner et al., 2012; van der Waal et al., 2012; Kaufmann et al., 2013). The concept has already been tested by Mori et al. (2012, 2013a) without fully satisfactory results, due to the necessary averaging of EEG event related potentials (ERP) and the low classification accuracy obtained with a classical stepwise linear discriminant analysis (SWLDA) (Krusienski et al., 2006) methods. In order to improve the classification accuracy and the information transfer rate (ITR), we propose to use a logistic regression (LR) classifier as implemented by Fan et al. (2008) in a single trial (no response averaging) scenario. We also propose a data-driven filtering method capable of a successful filtering of the non-linear and non-stationary signals such as EEG and especially ERP responses. The method is based on synchronosqueezing transform application (Daubechies et al., 2011).

In order to test the hypothesis, a six digits head-locations–spatial speller task is proposed, as first developed and head positions optimized for the best bone-conduction and tactile stimuli by Mori et al. (2013c). We test and compare the new results offline with the state-of-the-art methods, using the experimental dataset collected within the projects by Mori et al. (2013c) and Rutkowski et al. (2014).

2. Materials and methods

In the experiments described in this paper, eleven BCI-naive users (mean age 21.82 with standard deviation of 0.87) took a part. All the experiments were performed at the Life Science Center of TARA, University of Tsukuba, Japan. The psychophysical and online EEG tbcABI paradigm experiments were conducted in accordance with WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects. The experimental procedures were approved and designed in agreement with guidelines of The Faculty of Engineering, Information and Systems Ethical Committee at the University of Tsukuba, Japan. The BCI-naive users performed the experiments for monetary compensation. The 100 ms long stimuli in the form of sinusoidal waves were delivered to each user’s head areas via the tactile exciters HiWaveHIAx19C01–8 working in the range of 300–20,000 Hz. The vibrotactile stimulators were arranged as follows. The pairs of exciters were attached on both sides of the forehead, chin, and behind the ears respectively.

The psychophysical experiments were conducted in order to test the tactile and bone-conduction stimuli based paradigm with BCI-naive users. The users were asked to respond using a computer keypad after the target stimulus was presented in a random series of distractors. The stimulus onset asynchrony was set to 1000 ms.

During the online tbcABI experiments, the EEG signals were captured with a portable wireless EEG amplifier system, g.MOBILab+, using eight dry g.SAHARA electrodes by g.tec Medical Instruments GmbH, Austria. The electrodes were attached to eight electrode sites, Cz, C4, F3, P3, P4, C3, C4, CP5, and CP6, as in the 10/20 extended international system (Jurcak et al., 2007) as depicted in form of a topographic plots of the top panels in Fig. 4. The ground and reference electrodes were attached behind the left and right ears respectively. In order to limit electromagnetic interference, the user’s hand was also grounded with a conductive armband connected to the amplifier’s ground. No electromagnetic interference was observed from the vibrotactile transducers attached to the EEG cap and positioned on the user’s head as depicted in Fig. 1. The recorded EEG signals were processed by the in-house enhanced BCI2000 application to preprocess EEG with SST. The classification was carried out by a logistic regression classifier based on the Mat- lab toolboxes (Fan et al., 2008; Donders Institute for Brain, 2014) with features drawn from 0 to 600 ms ERP intervals. The sampling rate was set to 256 Hz, with a high pass filter at 0.1 Hz, and low pass filter at 40 Hz. The stimulus onset asynchrony (SOA) was of 500 ms and each stimulus lasted 100 ms. The user was instructed to spell six digit long random sequences of numbers ranging from 1 to 6, which were represented by the locations of the vibrotactile exciters on the head in each experimental session. The single trial ERP responses were used for the classification in order to improve ITR results. The details of EEG signal recording settings are summarized in Table 1.

2.1. Synchronosqueezing transform-based EEG preprocessing

The empirical mode decomposition (EMD) class of algorithms is a technique that aims to decompose a given univariate (Huang et al., 1998) or multivariate (Rehman and Mandic, 2010) signal into its building block functions that are the superposition of a limited number of components. The EMD-based techniques have been already applied successfully in offline mode to artifacts
Table 1
Summary of the single trial-based online tbcABCI results of the eleven BCI-naive users performing online experiments.

<table>
<thead>
<tr>
<th>Experimental parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of commands</td>
<td>6</td>
</tr>
<tr>
<td>Stimulus onset asynchrony (SOA)</td>
<td>500 ms</td>
</tr>
<tr>
<td>Number of averaged ERPs</td>
<td>1</td>
</tr>
<tr>
<td>Mean accuracy</td>
<td>64.3%</td>
</tr>
<tr>
<td>The best accuracy</td>
<td>72.6%</td>
</tr>
<tr>
<td>Mean ITR</td>
<td>16.4 bit/min</td>
</tr>
<tr>
<td>The best ITR</td>
<td>22.3 bit/min</td>
</tr>
</tbody>
</table>

Fig. 2. Confusion matrix of the psychophysical experiment users’ grand mean averaged responses. The diagonal of the above matrix depicts the correct response rates, while the off-diagonal the mistakes. The last (seventh and marked with N) column illustrates the lack of misses (no response) in our experiment. The color-coding, summarized with the color-bar on the right of the plot, visualizes the response rates.

Fig. 3. Distribution boxplots depicting the all participating users psychophysical experiment grand mean response times for each direction plotted separately. The red lines in each plot depict the 25th, 50th, and 75th percentiles. No significant differences were observed among the medians as tested with the pairwise Wilcoxon rank sum tests. The color-coding, summarized with the color-bar on the right of the plot, visualizes the response numbers. (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

removal from EEG (Rutkowski et al., 2009; Molla et al., 2012, 2013; Mandic et al., 2013). Unfortunately the EMD algorithm due to its iterative decomposition nature is hard to apply in online ERP (P300 response especially) stimulus-driven BCI application due to limited time constrains (hundreds of millisecond stimulus onset asynchrony requirements). A good solution for this problem is the synchrosqueezing transform (SST) method proposed by Daubechies et al. (2011). The SST method is a combination of wavelet analysis and reallocation methods. In the case presented in the paper the computational speed of the proposed SST was of 160 ms for a single eight-channels decomposition in comparison to 10 s required by the MEMD.

Before explaining details of the proposed implementation of the SST to ERP responses preprocessing, let’s review briefly the wavelet and SST methods. Given the originally recorded EEG signal s(t), its classical wavelet transform \( W_s(a, t) \) is obtained as

\[
W_s(a, t) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s(u) \psi \left( \frac{u - t}{a} \right) du,
\]

where \( a \) sets the scale and \( \psi(u) \) is the chosen wavelet function (Morlet wavelet has been chosen in our project). The wavelet transform does not cause any loss of information, so the original signal can be reconstructed as

\[
s(u) = D_\psi^{-1} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} \psi \left( \frac{u - t}{a} \right) W_s(a, t) \frac{da}{a^2},
\]

with the constant \( D_\psi \) determined as \( D_\psi = \int_{-\infty}^{\infty} |\psi(\xi)|^2 d\xi \) with \( \psi(\xi) \) representing the Fourier transform of the chosen wavelet function. In real-life cases, both time and scale (frequency) take the discrete values (not continues as in the above equations). Time \( t_k = k0t = kf_s \) where \( f_s \) is the sampling frequency of the original, EEG in this case, signal. The scale values are usually chosen to be equi-log-spaced (dyadic convention, etc.). The wavelet transform frequency localization at \( f_0 = 1 \) is often not precise enough to distinguish the frequencies of different oscillatory components so common in the noisy EEG recordings. It is possible to increase \( f_0 \), but that would cause loss in time resolution, causing some oscillations in the
frequency of a given harmonic to be regarded as a set of independent harmonics, what is usually a case in the Fourier transform. The recently proposed SST method by Daubechies et al. (2011) allows for providing a time–frequency representation with much better frequency and time resolutions at the same time. The concept is based first on an identification of the frequencies \( f_i \) for which the phase of the wavelet coefficient grows for each scale and time:

\[
  f(a, t) = \frac{1}{2\pi} \int \frac{\delta}{M} \cdot \text{arg}(W_i(a, t)),
\]

(3)

where \( \text{arg}(\cdot) \) stands for the phase of the complex coefficient and the multiplier \( 1/2\pi \) is necessary to convert from circular to normal frequency. Once the \( f_i(a, t) \) have been determined from the analyzed signal, the frequencies \( f_i \) could be chosen to form the bins as \( [f_i - f_i^+; f_i^+] \) and the SST can be calculated as

\[
  T_i(f_i, t) = C_\Psi \sum_{j=1}^{N \times \Psi} W_i(a_j, t) a_j^{-3/2} \Delta a_j,
\]

(4)

where \( \Delta a_j \) are the distances between the adjacent scales. The constant \( C_\Psi \) having meaning of amplitude is defined as

\[
  C_\Psi = \frac{1}{2\pi} \int_0^\infty \Psi(\xi) d\xi
\]

with \( \Psi(\xi) \) being here again Fourier transform of the chosen wavelet transform (Morlet for examples in this paper). The single channel EEG responses after transformation to the SST frequency domain as in Eq. (4) are bandpass filtered only in the frequency range of 0–8 Hz, which allows us to keep only the ERP-related oscillations. The above very low and narrow frequency band is known to carry the majority of ERP components necessary for discrimination of P300 responses occurrences and it does not lower significantly the classification accuracy as shown by Bougrain et al. (2012).

The original EEG signal could be reconstructed from SST to its time domain form simply (Daubechies et al., 2011) as

\[
  s(t) = \text{real} \left( \sum_i T_i(f_i, t) \right) = \left| \sum_i T_i(f_i, t) \right| \cos \left( \text{arg} \sum_i T_i(f_i, t) \right) .
\]

(5)

The very fast implementation of the SST transform and its inverse as in Eqs. (4) and (5) allows for the very precise bandpass filtering of the non-stationary and non-linear ERPs in order to enhance the
very slow P300 responses. The very encouraging results of P300 classification improvement results are summarized in Fig. 6.

2.2. Logistic regression-based P300 responses classification

The logistic regression (LR) is a popular linear classification method in which given a set of ERP response label (target vs. non-target) pairs \((s_i, y_i), i = 1, \ldots, n\), where \(s_i \in \mathbb{R}^n\) and \(y_i \in \{-1, +1\}\) an unconstrained optimization problem is solved as

\[
\min_w \frac{1}{2} w^T w + C \sum_{i=1}^{n} \log \left(1 + e^{-y_i w^T s_i}\right),
\]

where \(C > 0\) is a penalty parameter. In the classifier testing phase (the actual usage in BCI application after the classifier has been trained) the resulting class membership is predicted for the ERP input features \(s\) as the target if \(w^T s > 0\), and the non-target otherwise.

The very encouraging, yet not fully perfect, implementation of LR classification to single trial P300 responses in tbcaBCI application with BCI-naive users is summarized in Fig. 6. The ERP responses with the proposed SST stage further improved the classification results.

3. Results

The results of the conducted psychophysical study to confirm the new paradigm validity as well the BCI offline data filtering with SST for further improvement of classification results are summarized in the following sections. The new tactile and bone-conduction auditory paradigm and the signal processing methods resulted with very encouraging outcomes supporting the original research hypothesis.

3.1. Psychophysical experiment results

The results of the psychophysical experiments conducted in order to validate the tbcaBCI paradigm with BCI-naive users are summarized in Figs. 2 and 3. The confusion matrix with averaged response accuracies to the instructed targets printed in Fig. 2 have confirmed the experimental paradigm of the novel bimodal (tactile and bone-conduction) interfacing modality. The mistakes made by
the users were marginal. The even more encouraging results are depicted in Fig. 3 where the comparison of response time distributions to the six targets has been summarized. The median responses resulted with almost the same (no significant differences observed) scores further confirming the same cognitive loads induced by the various stimulus points in the tactile and bone-conductance auditory perception and recognition, even in the case of BCI-naive users.

3.2. EEG BCI experiment results

The tbcaBCI paradigm EEG responses in form of grand mean averages for the all eleven participating users are depicted in Fig. 5 with very clear differences of attended versus ignored stimuli in P300 latencies. Also the area under the curve (AUC) of the ROC for feature plots marking the most discriminable latencies is depicted in Fig. 4 showing the same ranges. Topographic plots of the AUC distributions are also presented in Fig. 4, supporting the choice of the eight dry EEG electrodes covering the vertex and the parietal cortex locations. The results of offline BCI signal processing and classification improvements sessions are summarized in Fig. 6 in the form the classification accuracies using the classical SWLDA method with the proposed LR, combined with SST preprocessing. We present also offline concatenated and averaged trial results of 2–5 ERPs which did not improve significantly results further over the proposed combination of SST-preprocessed LR classification. The mean 6.6% classification boost after application of the SST has been observed for the single trial-based LR classification. The SST application improvement was less pronounced as depicted also in Fig. 6. Unfortunately in the case of BCI-naive users, who performed the experiments for the first time in their lives, nobody could reach a 100% accuracy even with five trials averaging scenario.

These preliminary yet encouraging results are a step forward in the search for the new “non-vision-based” BCI paradigms.

4. Conclusions

This study demonstrated results obtained with a novel sixcommands and the head locations based tbcaBCI paradigm developed, used and further in accuracy improved in experiments with eleven BCI-naive “body-able” users. The experiment results obtained in this study confirmed the validity of the tbcaBCI for interactive applications as well as the proposed signal preprocessing method, based on frequency domain filtering in SST space, allowed for successful classification even of the single-trial ERP responses which has been crucial for ITR scores boosting.
The results presented offer a step forward in the development of novel, and very much expected to improve life of ALS patients, neurotechnology applications. Since the online BCI did not yield maximum possible information transfer rates overall, the current paradigm would obviously need improvements and modifications. These needs determine the major lines of study for future research. However, even in its current form, the proposed tbsBCI can be regarded as a practical solution for ALS–US patients (locked into their own bodies despite often intact cognitive functioning), who cannot use vision or auditory based interfaces due to sensory disabilities.

Author contributions

TMR: Conceived the concept of the spatial tactile BCI and designed the EEG experiments. HM, TMR: Performed the psychophysical and EEG experiments, as well as analyzed the data. TMR: Wrote the paper.

Acknowledgements

This research was supported in part by the Strategic Information and Communications R&D Promotion Programme no. 121803027 of The Ministry of Internal Affairs and Communication in Japan, and by KAKENHI, the Japan Society for the Promotion of Science Grant No. 12010738. We also acknowledge the technical support from YAMADA Sound & IT Development Division in Hamamatsu, Japan.

References

Bourgain L, Saavedra C, Ranta R. Finally, what is the best filter for P300 detection? In: TOBI workshop III – tools for brain-computer interaction; 2012 [http://hal.inria.fr/hal-00756669].


Donders Institute for Brain, Cognition and Behaviour. Fieldtrip; 2014 [http://fieldtriptoolbox.org].


Molla MKI, Tanaka T, Rutkowski T. Multivariate EMD based approach to EEG artifacts separation from EEG. In: 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP); 2012. p. 653–6 [URL http://dx.doi.org/10.1109/ICASSP.2012.6287968].


