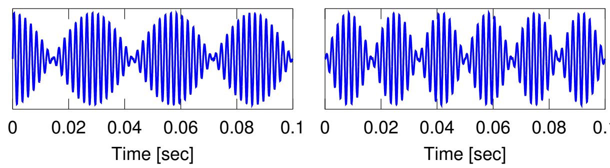
33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

EEG Auditory Steady State Responses Classification for the Novel BCI

Hiroshi Higashi, Tomasz M. Rutkowski, Yoshikazu Washizawa, Andrzej Cichocki, and Toshihisa Tanaka

***Abstract*— An auditory modality brain computer interface (BCI) is a novel and interesting paradigm in neurotechnology applications. The paper presents a concept of auditory steady state responses (ASSR) utilization for the novel BCI paradigm. Two EEG feature extraction approaches based on a bandpass filtering and an AR spectrum estimation are tested together with two classification schemes in order to validate the proposed auditory BCI paradigm. The resulting good classification scores of users intentional choices, of attending or not to the presented stimuli, support the hypothesis of the ASSR stimuli validity for a solid BCI paradigm.**



(a) 35 Hz amplitude modulation. (b) 60 Hz amplitude modulation.

Fig. 1: The stimuli signals designed to create ASSR response captured later in EEG.

I. INTRODUCTION

Brain computer/machine interface (BCI/BMI) is a device enabling to create an additional or independent communication channel between the brain/mind and a computer/machine without an involvement of peripheral nervous or muscular systems [1]–[3]. Non-invasive BCI is a challenging paradigm to achieve this goal. BCI tasks can be classified into two major categories: independent and dependent [2]. It has been known that the latter one can provide more commands and can be implemented in real-world settings. A number of studies in dependent BCI paradigms use visual stimuli and the corresponding brain responses. However, the use of auditory modality in BCI has several advantages for which belong: simple stimuli delivery via headphones or loudspeakers; no necessity for the user to direct a head toward the sound sources; minimal user's distraction; possible embedding of auditory stimuli within music; no evidence of possible danger of causing seizures; a possibility to utilize the paradigm for locked-in patients without any remaining muscle activity, vision or possibility to learn a movement imagery paradigm [2]. Auditory steady-state response (ASSR) is an established tool in objective hearing levels estimation [4]. ASSR is evoked by the periodic modulation, or turning *ON* and *OFF* of a tone [5]. The neural response is a brain potential that closely follows the time course of the amplitude modulation and it can be detected objectively at intensity levels close to behavioral threshold. In order to quantify brain responses evoked by attended ASSR stimuli we compare two feature extraction and two classification strategies. There have been reported studies on auditory modality based BCIs [6], [7], however, unlike the visual based paradigms, detailed stimulus types, effectiveness and user-friendly designs to achieve practical applications are still open questions.

The objective of this paper is to confirm a hypothesis that the ASSR based BCI paradigm is a satisfactory concept based on the offline testing results and the possibility to adopt to various users. The results presented at the end of the paper together with discussion corroborate the ASSR stimuli validity hypothesis.

H. Higashi and T. Tanaka are with the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Japan `higashi@sip.tuat.ac.jp, tanakat@cc.tuat.ac.jp`

T.M. Rutkowski is with the TARA Center, University of Tsukuba, Japan `tomek@tara.tsukuba.ac.jp`

T.M. Rutkowski, Y. Washizawa, A. Cichocki, H. Higashi, and T. Tanaka are with the RIKEN Brain Science Institute, Japan `{washizawa, cia}@brain.riken.jp`

This work is supported in part by KAKENHI, the Japan Society for the Promotion of Science grant no. 21360179.

II. METHODS

The EEG experiments were conducted in the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Tokyo, Japan, in accordance with the local institutional ethical committee guidelines. In the present study ten healthy male subjects labeled $S1$ to $S10$, age $22-30$ with mean 24 years old, were fully informed about the experimental procedure and they agreed voluntarily to participate by signing subject consent forms. All subjects had normal hearing levels. The recorded EEG data were next anonymized to protect subjects' privacy. The experimental procedure and data processing steps are described below.

A. EEG Measurement Procedure

For the ASSR experiments we use two types of stimuli. Both of them are based on a sinusoidal carrier tone with frequency of 500 Hz. To generate the ASSR response observable later in EEG the carrier tone is amplitude modulated with 35 Hz and 60 Hz sine waves separately as illustrated in Figures 1a and 1b. The modulated signal with 35 Hz (*stimulus L*) and the modulated signal with 60 Hz (*stimulus R*) are presented to the subjects through left and right earphone channels separately. During EEG measurement, the subjects sit in an armchair while focusing their sites on a computer display with a fixation mark. The subjects have been instructed to listen to two stimuli which are delivered to left and right earphones sequentially together with visual

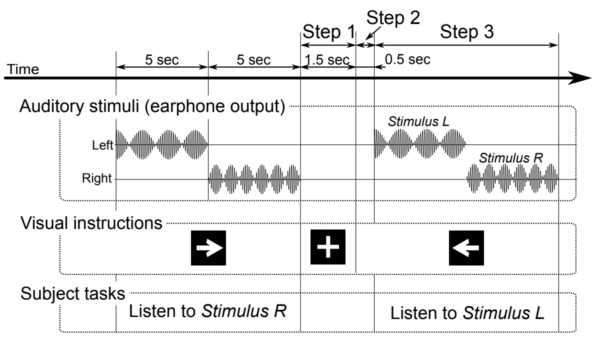


Fig. 2: Experimental procedure protocol.

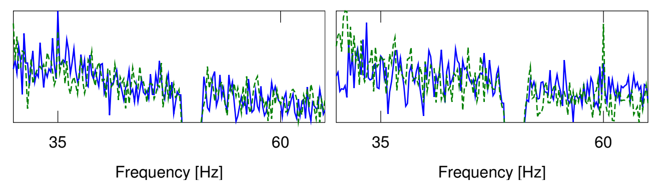


Fig. 3: DFT amplitude spectra for the EEG channel Cz of the *S5* subject's signals averaged over 50 trials. The solid (*L* class) and the dashed (*R* class) lines represent EEG spectra for both ASSR responses respectively. The instruction given to the subject in the right panel was to attend to *R* class and in the left panel to *L* class.

instructions as illustrated in Figure 2. The detailed experimental procedure steps are as follows:

Step 1: Subject is instructed to fixate eyes on a cross in a computer display in front of them for 1.5 seconds.

Step 2: A direction of stimulus is presented on the display in form of an arrow instructing the subject which direction of ASSR stimuli in next step to attend (left or right). The instructions with directions to attend are randimized.

Step 3: After 500 ms break, the ASSR stimulus is played via the single earphone channel (left or right). Subject attends to the modulated tone only if the instructed direction (displayed arrow) matches the earphone channel (left or right). Otherwise the subject shall ignore the stimulus.

Our experimental hypothesis is that an observed ASSR response corresponding to attended class of *L* and *R* stimuli should be reflected in EEG features. For a reference, we also extend the experiment with a third simple stimulus in form of a 100 Hz sinusoidal "through-bass" sound which is given to the subject in the Step 3. In the reference experiment with the through-bass, the subject attends to it when the ASSR stimulus shall be ignored (the instruction arrow left/right direction and the playback earphone channel mismatch case). The through-bass sound is played through the left and right channels of the earphone with the same volume level. The sound pressure level for the ASSR stimulus is about -14 dB. The protocol is repeated 100 times for each subject resulting with 50 trials for each of *L* and *R* classes. The length of each single trial is 10 seconds.

The EEG signals are recorded with 12 electrodes located on $T7$, $C5$, $C3$, $C1$, Cz, $C2$, $C4$, $C6$, $T8$, $TP7$, CPz, and $TP8$ as in international $10-10$ EEG recording system. A ground electrode is located on a forehead and a reference is attached to an earlobe ($A1$ position). The EEG signals are amplified and bandpass filtered in a frequency range of $0.08-100$ Hz using an amplifier MEG-6116 (NIHON KOHDEN). Moreover, we apply the hardware notch filter at 50 Hz to remove electrical power noise. The amplified signals are next digitized with 512 Hz sampling frequency by an A/D converter AIO-163202F-PE (CONTEC) and recorded on a computer hard disk using Data Acquisition Toolbox in MATLAB. Discrete Fourier transform (DFT) amplitude spectra averages over trials for a Cz channel (in later discussed classification results the single trial EEG signals are used) derived from a single subject EEG are presented in Figure 3. This figure visualizes strong DFT peaks related to attended ASSR response frequencies at 35 and 60 Hz respectively.

B. Classification Procedure

Responses to presented stimuli in EEG signals are tested with utilization of two feature extraction and two classification procedures to verify the usability of ASSR paradigm for BCI purposes, as described in following sections.

1) Feature extraction: We obtain each feature vector from an observed signal by bandpass filtering and autoregressive (AR) spectrum estimation, respectively. In this section, let $x_n(t,i)$ be the recorded signal of the i-th channel from the n-th trial where i represents an index of channel for $i = 1,\dots,12$, t stands for a time index of a discrete signal for $t = 1,\dots,5120$, and $n = 1,\dots,100$, and let $y_n \in \{L,R\}$ be the class label of the n-th trial.

a) Bandpass filtering: We apply two Butterworth bandpass filters F_1 and F_2 with pass-band of $35-f_b$ to $35+f_b$ Hz and $60-f_b$ to $60+f_b$ Hz respectively, where $2f_b$ represents the width of the pass-band of the chosen order-two filters. Then, we obtain two filtered signals from each channel and the energy of these signals as

$$p_n^j(i) = \sum_{t=1}^{2560}\{F_j(x_n(t,i))\}^2, \quad j=1,2, \tag{1}$$

$$q_n^j(i) = \sum_{t=2561}^{5120}\{F_j(x_n(t,i))\}^2, \quad j=1,2. \tag{2}$$

The feature vector is defined as

$$\begin{aligned} \boldsymbol{z}_n =& [p_n^1(1),\dots,p_n^1(12),p_n^2(1),\dots,p_n^2(12),\\ & q_n^1(1),\dots,q_n^1(12),q_n^2(1),\dots,q_n^2(12)]^\top. \end{aligned} \tag{3}$$

The labeled data are described in $\{\boldsymbol{z}_n, y_n\}_{n=1}^{100} \subset \mathbb{R}^{48} \times \{L,R\}$. The norm of the feature vector is normalized to 1.

b) AR spectrum estimation: We define AR models as;

$$x_n(t,i) = -\sum_{k=1}^{M} r_n(k,i)x_n(t-k,i) + v_n(t,i), \quad t = M+1,\dots,2560, \tag{4}$$

$$x_n(t,i) = -\sum_{k=1}^{M} s_n(k,i)x_n(t-k,i) + w_n(t,i), \quad t = 2561+M,\dots,5120, \tag{5}$$

where (4) and (5) are models in periods of playing *stimulus L* and *stimulus R*, respectively, M is an AR model order, $r_n(k,i)$ and $s_n(k,i)$ are coefficients of AR models, and $v_n(t,i)$ and $w_n(t,i)$ are white noises with zero time averages. From observation signal, $x_n(t,i)$, we estimate the coefficients of AR models by the Yule-Walker method [8]. A feature vector, $\boldsymbol{z}_n$, consists of the estimated coefficient as:

$$\begin{aligned}\boldsymbol{z}_n = [&r_n(1,1),\dots,r_n(M,1),r_n(1,2),\dots,r_n(M,2),\\ &\dots,r_n(1,12),\dots,r_n(M,12),\\ &s_n(1,1),\dots,s_n(M,1),s_n(1,2),\dots,s_n(M,2),\\ &\dots,s_n(1,12),\dots,s_n(M,12)]^\top.\end{aligned} \tag{6}$$

The labeled data is described in $\{\boldsymbol{z}_n, y_n\}_{n=1}^{100} \in \mathbb{R}^{24M} \times \{L,R\}$. Moreover, the norm of the feature vector is normalized to 1.

2) Classification: We adopt two classification methods: (i) principal component analysis (PCA) and linear discriminant analysis (LDA) [9]; (ii) linear support vector machine (SVM) with soft-margin [10]. Those are linear classification methods with a discriminant function,

$$f(\boldsymbol{z}) = \begin{cases} L & (\boldsymbol{w}^\top \boldsymbol{z} - b \geq 0) \\ R & (\boldsymbol{w}^\top \boldsymbol{z} - b < 0) \end{cases}, \tag{7}$$

where parameters ($\boldsymbol{w}$, b) are replaced by ($\boldsymbol{w}_{\mathrm{LDA}}$, b_{LDA}) or ($\boldsymbol{w}_{\mathrm{SVM}}$, b_{SVM}) obtained as follows. We define a set of the samples for learning as $\boldsymbol{z}_n \in \mathbb{R}^D$ where the sets of indexes corresponding to L and R classes given as $\Omega_L = \{n|y_n = L\}$ and $\Omega_R = \{n|y_n = R\}$ and D denotes the dimension of a sample vector. And we assume an unlabeled feature vector as $\boldsymbol{z} \in \mathbb{R}^D$.

a) PCA and LDA: PCA is performed for the both classes, so a set of eigenvectors $\boldsymbol{u}_1, \dots \boldsymbol{u}_r$ corresponds to the largest r eigenvalues of $\boldsymbol{R} = \frac{1}{|\Omega_L|+|\Omega_R|}\sum_{n\in\Omega_L\cup\Omega_R} \boldsymbol{z}_n\boldsymbol{z}_n^\top$. All sample vectors and test vector are transformed using a matrix of $\boldsymbol{U} = [\boldsymbol{u}_1|\boldsymbol{u}_2|\dots|\boldsymbol{u}_r] \in \mathbb{R}^{D\times r}$ such as $\hat{\boldsymbol{z}}_n = \boldsymbol{U}^\top \boldsymbol{z}_n$ and $\hat{\boldsymbol{z}} = \boldsymbol{U}^\top \boldsymbol{z}$. Next, we classify the test sample by using classification function based on LDA. Let mean vectors be

$$\boldsymbol{m}_L = \frac{1}{|\Omega_L|}\sum_{n\in\Omega_L} \hat{\boldsymbol{z}}_n, \quad \boldsymbol{m}_R = \frac{1}{|\Omega_R|}\sum_{n\in\Omega_R} \hat{\boldsymbol{z}}_n, \tag{8}$$

and the scatter matrix be

$$\begin{aligned}\boldsymbol{S}_w =& \frac{1}{|\Omega_L|}\sum_{n\in\Omega_L}(\hat{\boldsymbol{z}}_n - \boldsymbol{m}_L)(\hat{\boldsymbol{z}}_n - \boldsymbol{m}_L)^\top \\ &+ \frac{1}{|\Omega_R|}\sum_{n\in\Omega_R}(\hat{\boldsymbol{z}}_n - \boldsymbol{m}_R)(\hat{\boldsymbol{z}}_n - \boldsymbol{m}_R)^\top.\end{aligned} \tag{9}$$

Then the parameters of the classifier are given by $\boldsymbol{w}_{\mathrm{LDA}} = \boldsymbol{U}\hat{\boldsymbol{w}}$ and $b_{\mathrm{LDA}} = \frac{1}{2}\hat{\boldsymbol{w}}^\top(\boldsymbol{m}_L + \boldsymbol{m}_R)$, where $\hat{\boldsymbol{w}} = \boldsymbol{S}_w^{-1}(\boldsymbol{m}_L - \boldsymbol{m}_R)$.

b) Linear SVM: The parameters $\boldsymbol{w}_{\mathrm{SVM}}$ and b_{SVM} are obtained by solving the convex quadratic optimization problem of SVM [10]. Then a parameter for soft-margin is set to the problem as C. We use a package SVM^{light} [11] to solve the problem.

III. RESULTS

The very encouraging subject responses recognition results ranging from 62% to even 100% of presented experimental approach are presented in Table I, where classification accuracy and optimal parameters found for each subject are summarized. The parameters are chosen in $\{0.5, 1, \dots, 5\}$ for a bandwidth, f_b, $\{1, 2, \dots, 10\}$ for an order of AR model, $\{1, 2, \dots, 48\}$ for a rank of PCA in bandpass filtering, and $\{1, 2, \dots, \min(12M, 99)\}$ for a rank of PCA in AR spectrum estimation, and $\{10^1, 10^{1.2}, \dots, 10^5\}$ for soft-margin, C, of the SVM. The accuracy has been obtained by leave-one-out cross validation (L-CV) method. Because the results differ for the subjects, it is not clear whether the choice of the stimulus with through-bass has been effective.

The relations between averaged classification accuracy and parameters for each procedure are shown in Fig 4. Trends in a function of the parameters can be observed, for example, the parameter of the soft-margin, C, performs with better accuracy for $C > 10^2$ comparing to its lower values. However the results differ among the participants showing the sensitivity to subject variability.

The presented approach and very encouraging results based on the ten subjects sample are a step forward in creation of the new user friendly, auditory stimuli based BCIs. We have shown that the ASSR stimuli based BCI paradigm allows the users to switch their attention to one of the presented modulated tones and to generate a resulting stronger following response in EEG possible to rank with discussed in this paper feature extraction and classification approach.

We plan further multi-subject experimental trials to examine the through-bass inclusion on subjects performance as well to further validate the proposed BCI paradigm in an online multi-command application.

REFERENCES

[1] J. J. Vidal, "Toward direct brain-computer communication," *Annual Review of Biophysics and Bioengineering*, vol. 2, pp. 157–180, 1973.

[2] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 164–173, 2000.

[3] A. Cichocki, Y. Washizawa, T. Rutkowski, H. Bakardjian, A.-H. Phan, S. Choi, H. Lee, Q. Zhao, L. Zhang, and Y. Li, "Noninvasive BCIs: Multiway signal-processing array decompositions," *Computer*, vol. 41, no. 10, pp. 34–42, 2008.

[4] D. R. Stapells, D. Linden, J. B. Suffield, G. Hamel, and T. W. Picton, "Human auditory steady state potentials," *Ear and Hearing*, vol. 5, no. 2, 1984.

TABLE I: Classification accuracy and optimal parameters given by L-CV. Upper and lower columns of each table show the results of experiment "without" and "with" through-bass, respectively. The columns labeled "Soft-margin" show the value of $\log_{10}(C)$, where C is the parameter for soft-margin.

(a) Feature extraction: bandpass filtering, classifier: PCA and LDA

	Subject									
	S1	*S2*	*S3*	*S4*	*S5*	*S6*	*S7*	*S8*	*S9*	*S10*
Acc. [%]	66	82	76	99	79	84	66	75	74	71
f_b [Hz]	5.0	2.5	5.0	0.5	3.5	4.0	0.5	2.0	3.5	3.5
Rank	27	18	18	31	7	21	14	37	20	34
Acc. [%]	67	76	66	100	67	93	62	61	70	84
f_b [Hz]	5.0	4.0	2.0	0.5	2.5	5.0	1.0	4.5	1.5	3.5
Rank	22	41	5	31	7	14	13	20	26	21

(b) Feature extraction: bandpass filtering, classifier: linear SVM

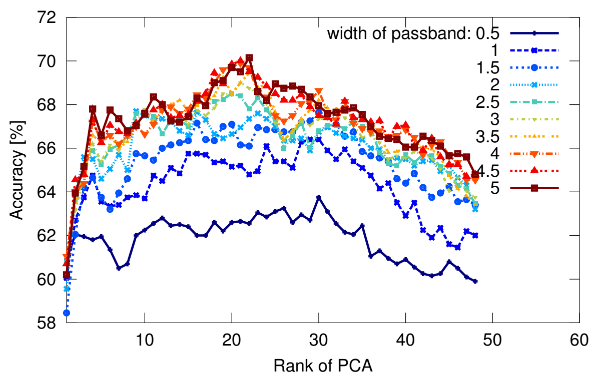
	Subject									
	S1	*S2*	*S3*	*S4*	*S5*	*S6*	*S7*	*S8*	*S9*	*S10*
Acc. [%]	68	83	79	97	84	82	68	74	72	82
f_b [Hz]	4.5	1.0	4.5	1.5	5.0	4.5	1.0	2.5	3.5	5.0
Soft-margin	3.4	2.8	2.4	1.0	1.0	1.0	3.0	4.0	4.2	3.4
Acc. [%]	65	74	65	99	64	94	63	62	71	90
f_b [Hz]	4.0	2.5	2.0	0.5	1.0	5.0	1.5	5.0	3.0	4.5
Soft-margin	4.0	3.0	2.0	1.0	1.0	1.6	2.6	1.0	3.6	2.6

(c) Feature extraction: AR spectrum estimation, classifier: PCA and LDA

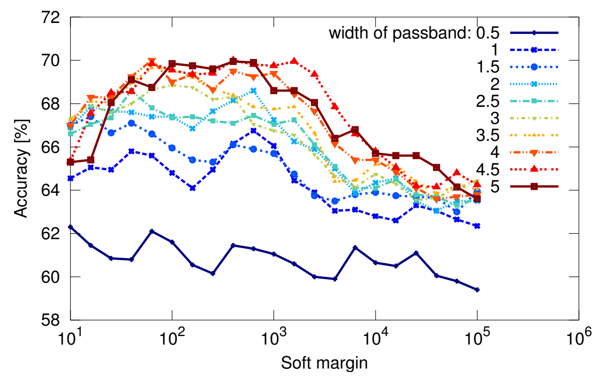
	Subject									
	S1	*S2*	*S3*	*S4*	*S5*	*S6*	*S7*	*S8*	*S9*	*S10*
Acc. [%]	86	90	74	97	96	65	74	87	81	90
AR order	4	8	2	6	2	2	3	7	4	7
Rank	26	91	13	3	12	8	44	57	96	72
Acc. [%]	67	74	72	100	70	82	74	73	72	95
AR order	7	6	8	2	4	3	4	3	2	7
Rank	62	20	21	3	20	32	37	16	26	22

(d) Feature extraction: AR spectrum estimation, classifier: linear SVM

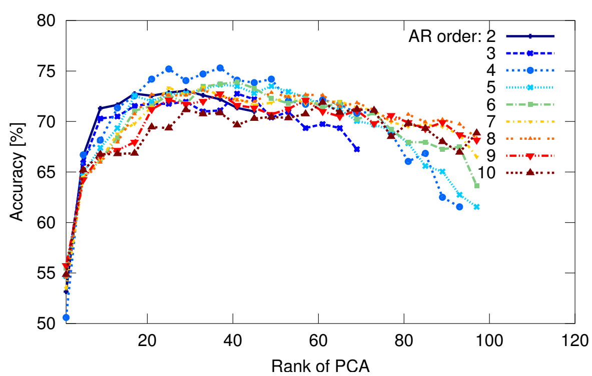
	Subject									
	S1	*S2*	*S3*	*S4*	*S5*	*S6*	*S7*	*S8*	*S9*	*S10*
Acc. [%]	83	89	75	96	98	69	74	86	79	87
AR order	7	9	3	2	2	3	2	10	4	6
Soft-margin	2.4	2.6	3.4	1.0	2.6	5.0	3.0	1.8	2.4	1.6
Acc. [%]	68	73	72	100	68	85	74	74	76	93
AR order	2	6	9	2	2	8	3	8	2	6
Soft-margin	3.4	2.6	2.2	1.0	2.4	2.6	3.8	1.2	3.2	2.8



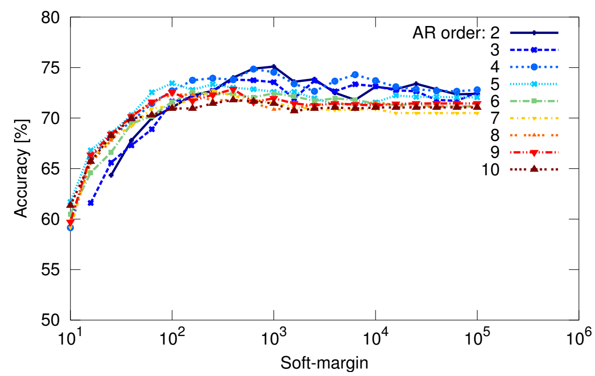
(a) Feature extraction: bandpass filtering, classifier: PCA and LDA



(b) Feature extraction: bandpass filtering, classifier: linear SVM



(c) Feature extraction: AR spectrum estimation, classifier: PCA and LDA



(d) Feature extraction: AR spectrum estimation, classifier: linear SVM

Fig. 4: Relations between classification accuracy and classification parameters.

[5] R. Galambos, S. Makeig, and P. J. Talmachoff, "A 40-Hz auditory potential recorded from the human scalp," *Proceedings of National Academy of Sciences*, vol. 78, no. 4, pp. 2643-2647, 1981.

[6] M. A. Lopez, H. Pomares, F. Pelayo, J. Urquiza, and J. Perez, "Evidences of cognitive effects over auditory steady-state responses by means of artificial neural networks and its use in brain-computer interfaces," *Neurocomputing*, vol. 72, no. 16-18, pp. 3617-3623, 2009.

[7] M. Schreuder, B. Blankertz, and M. Tangermann, "A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue," *PLoS ONE*, vol. 5, p. e9813, 04 2010.

[8] S. L. Marple, *Digital spectral analysis: with applications*, Upper Saddle River: Prentice-Hall, Inc., 1986.

[9] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

[10] V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

[11] T. Joachims. "Making large-scale SVM learning practical" in *Advances in Kernel Method - Support Vector Learning*. B. Schölkopf, C. Burges, and A. Smola, Ed. MIT-Press, 1999.